scientific reports

OPEN

Proper likelihood functions for parameter estimation in S-shaped models of unperturbed tumor growth

Erick E. Ramirez-Torres[1,2,5], Antonio R. Selva Castañeda[1], Luis Rández[1], Scott A. Sisson[3], Luis E. Bergues Cabrales[4,5] & Juan I. Montijano[1,5]

The analysis of unperturbed tumor growth kinetics, particularly the estimation of parameters for S-shaped equations used to describe growth, requires an appropriate likelihood function that accounts for the increasing error in solid tumor measurements as tumor size grows over time. This study aims to propose suitable likelihood functions for parameter estimation in S-shaped models of unperturbed tumor growth. Five different likelihood functions are evaluated and compared using three Bayesian criteria (the Bayesian Information Criterion, Deviance Information Criterion, and Bayes Factor) along with hypothesis tests on residuals. These functions are applied to fit data from unperturbed Ehrlich, fibrosarcoma Sa-37, and F3II tumors using the Gompertz equation, though they are generalizable to other S-shaped growth models for solid tumors or analogous systems (e.g., microorganisms, viruses). Results indicate that error models with tumor volume-dependent dispersion outperform standard constant-variance models in capturing the variability of tumor measurements, particularly the Thres model, which provides interpretable parameters for tumor growth. Additionally, constant-variance models, such as those assuming a normal error distribution, remain valuable as complementary benchmarks in analysis. It is concluded that models incorporating volume-dependent dispersion are preferred for accurate and clinically meaningful tumor growth modeling, whereas constant-dispersion models serve as useful complements for consistency and historical comparability.

Understanding tumor growth kinetics (TGK) is a challenging aspect of untreated cancer biology. Even though Bru et al.[1] reported that untreated tumors grow linearly, other studies suggest that tumor growth is inherently nonlinear. Most experimental studies reveal that TGK follows a sigmoidal (S-shaped) rather than exponential pattern, as observed across various tumor types in different hosts, independently of the tumor formation pathway (e.g., cell inoculation, solid tumor transplantation, or induction by carcinogenic agents)[2–6]. Different S-shaped mathematical models, particularly those based on ordinary differential equations (ODEs), have been used to describe TGK. These include power-law, Richards, Weibull, von Bertalanffy, von Bertalanffy-Pütter type, Verhulst (Logistic), and conventional Gompertz equations[1–9]. Additionally, models such as the Kolmogorov-John-Mell-Avrami[10], modified Kolmogorov-John-Mell-Avrami[10] and Montijano-Bergues-Bory-Gompertz[11] equations have also been applied to describe unperturbed TGK. Other models[12–14] have been proposed and recent advancements also explore fractional-order ODEs[15–19] to model tumor growth dynamics.

The consensus of authors is that the conventional Gompertz equation shows superior performance in fitting experimental data compared to other models[4,5,10,20–28]. Logistic and von Bertalanffy equations often fail to capture tumor growth data accurately, but several alternative models, including the Gompertz[14], have been effective for describing tumor growth dynamics in specific contexts. Nevertheless, those two models adequately described

[1]Instituto Universitario de Investigación de Matemáticas y Aplicaciones, Universidad de Zaragoza, Zaragoza, Spain. [2]Departamento de Biomédica, Facultad de Ingeniería en Telecomunicaciones, Informática y Biomédica, Universidad de Oriente, Santiago de Cuba, Cuba. [3]UNSW Data Science Hub, and School of Mathematics and Statistics, University of New South Wales, Sydney, Australia. [4]Departamento de Investigación e Innovación, Centro Nacional de Electromagnetismo Aplicado, Universidad de Oriente, Santiago de Cuba, Cuba. [5]Erick E. Ramirez-Torres, Luis E. Bergues Cabrales and Juan I. Montijano contributed equally. ✉email: berguesc@yahoo.com; monti@unizar.es

human tumor growths[4]. For instance, Ghaffari Laleh et al.[29] demonstrated that the general von Bertalanffy and conventional Gompertz equations yielded the lowest mean absolute errors in forecasting tumor growth across 1,472 patients with three or more data points.

Other studies have compared power-law, Logistic, conventional Gompertz, exponential, and von Bertalanffy equations across data from 10 tumor types[30]. The results revealed that the power-law, Logistic, and conventional Gompertz equations best described TGK in six, three, and one tumor types, respectively. Conversely, von Bertalanffy (in seven tumor types) and exponential (in three tumor types) models deviated significantly from experimental data. Furthermore, Valentim et al.[15] showed that non-integer order models (e.g. Conventional Gompertz, exponential, von Bertanlanffi and Guiott-West models) consistently described experimental data better than their integer-order counterparts in G1-101A human breast cancer cells transplanted into nude mice.

The performance of different models depends on factors such as the tumor growth phase and data range. For example, early tumor growth stages (exponential phase)[10,31] are best captured by fractional or integer exponential models[4,5,7,8,10,11,14,31–35], whereas other models perform poorly for time points beyond 80 days. Additionally, the fractality[11,32,36–39], heterogeneity[40], and viscoelasticity[41,42] of tumors may explain why fractional-order models often outperform integer-order counterparts. Notably, some models, like von Bertalanffy-Pütter, are more suitable for short time spans (e.g. $t < 10$days)[9].

These findings highlight three key points: 1) No single ODE-based model is universally accepted as the most appropriate for all cancer types or histological subtypes. 2) Incorrect model selection can lead to parameter estimation inaccuracies and unreliable oncological predictions. 3) Although statistical metrics often favor one model over others for specific experimental conditions, these differences are sometimes negligible[10,35], suggesting that multiple models can be used interchangeably for the same histological subtype in a given host.

From an experimental point of view, this may mean that these models may be used interchangeably to fit data for the histological variety growing in a given host. All S-shaped equations reported in the literature may be related to each other by establishing analytical relationships between their parameters, following the methodology reported in[32]. Therefore, any of these equations, in principle, may be used interchangeably in the fitting of experimental data. As Gompertz formulations are fractal in nature[10,11,32], this correspondence between S-shaped equations may suggest that all these equations are also fractal in nature.

In most of the above studies, the possible errors in the data are not considered or they are supposed to follow a Normal distribution. Nevertheless, in vitro[7,43], preclinical[7,8,10,11,33–35,44] and clinical[45] studies show that measurement errors tend to increase with tumor volume, especially during transitions from the avascular (spherical) to the vascular (ellipsoidal) phase of TGK. This increased error likely arises from irregularities in tumor edges, spatial and temporal variations in blood vessel distribution, and tumor heterogeneity and anisotropy[10,11,31,35,42,46,47]. These observations raise a critical question: Should Gompertz model parameter estimation explicitly account for the increasing measurement errors during TGK, or should the model aim to fit the data as closely as possible without considering error variability? Addressing this issue necessitates selecting an appropriate likelihood function for model fitting, whether using classical maximum likelihood estimation (MLE) or Bayesian approaches. Different strategies have been employed to address this challenge. Some studies rely on least-squares fitting without reference to a likelihood function, often implicitly assuming constant Gaussian errors, which also reduces to a least-squares fit under the MLE paradigm[48–50]. Others model error variability by using normal distributions with dispersion proportional to tumor volume[51,52], requiring algorithms like Nelder-Mead[7,53] for fitting instead of least-squares-based methods like Levenberg-Marquardt. The Student-t distribution, known for its robustness to outliers[54], also presents a promising alternative for modeling tumor measurement errors, as demonstrated in recent applications to COVID-19 data[55]. Despite differences between TGK and epidemic growth, several models are applicable to both domains.

In this paper, we analyze the impact of measurement errors on the estimation of standard Gompertz model parameters. Specifically, we compare five likelihood functions based on different error models: Normal with constant standard deviation (Norm), Normal with standard deviation proportional to tumor volume (Norm prop), Normal with a constant standard deviation below a threshold volume V_m and proportional standard deviation above it (Thres), Student-t with constant standard deviation (Stud) and Student-t with standard deviation proportional to tumor volume (Stud prop).

These likelihood functions are evaluated using Bayesian criteria, namely the Bayesian Information Criterion (BIC), Deviance Information Criterion (DIC), and Bayes Factor (BF). To our knowledge, this is the first comprehensive comparison of these error models, aiming to guide the selection of the most suitable approach for tumor growth analysis.

Methods
Models
As the main intention of this study is not to compare S-shaped models, but the different error models represented by likelihood functions, we select one of these S-shaped models to fit experimental data in solid tumors with different likelihood functions for the errors. We take the standard Gompertz equation based on the consensus of authors to describe unperturbed TGK[1,4,5,8,10,11,31,35]) and because this equation gives a very reasonable fitting in all cases, although it does not give the best fitting results for some tumor types. Note that the analysis we are presenting can be applied to any other tumor growth model.

The Gompertz growth model can be expressed in several parameterizations, classified as type I and type II based on their shape parameters[56]. We posit that the modified type-II model, which incorporates observable tumor volume, offers the most clinically interpretable parameters[11]. Although any parametrization is theoretically viable for likelihood comparison, the type-I Gompertz model is employed here for its simplicity. In this model, the tumor volume $V(t)$ at time t is a function of the negative logarithm of the initial tumor volume $V(0)$ divided by the carrying capacity (K), expressed as

$$V(t) = K \left(\frac{V(0)}{K} \right)^{e^{-rt}}, \tag{1}$$

where $r > 0$ represents the growth rate. This formulation of Gompertz growth has the practical advantage of simplifying the estimation of K due to its direct relationship with the data trends. As V(t) is a consequence of all the complex interactions that occur in the tumor and those between the tumor, as a whole, and the rest of the organism[8,10,11,31,35], equation 1 is sufficient for the intention of this study.

The individual observation model is defined as

$$y_i = f(t_i; \theta) + e_i, \qquad i = 1, \ldots, N, \tag{2}$$

where y_i denotes the measured tumor volume for the individual at time t_i, $f(t_i; \theta)$ is the Gompertz model evaluated at time t_i, $\theta = (r, K, V(0))$ is the vector of Gompertz model parameters, e_i is the model error for observation y_i, N is the number of measurements and M is the number of individuals in the experiment.

Since we are interested in the effect of the errors e_i on the computations of the parameters, we focus on individual datasets to isolate and accurately model tumor-specific measurement errors, which often depend on tumor size and shape irregularities. In this way the results will not be affected by the random inter-subject variability of the parameters of the equation, avoiding possible biases. Thus, we do not use the mixed effect model approach[57,58] in this study. We are also assuming that the errors in measurement follow the same distribution e for any individual, and depend only on the size of the tumor at time t_j, $e = e(t_j)$.

We now describe the five e_i expressions used to compare different likelihood functions. The first error model is the standard Normal error with constant standard deviation (σ), given by

$$e_i \sim \mathscr{N}(0, \sigma^2). \tag{3}$$

This error model and its corresponding likelihood function are referred to as Norm. The Norm model produces parameter estimates equivalent to the least-squares approach[59]..

The second error model, named Prop, posits that the standard deviation of the Normal distribution is proportional to the measured volume, formulated as

$$e_i \sim \mathscr{N}(0, (\sigma \times y_i)^2). \tag{4}$$

The third error model, designated as Thres, alters its law based on a threshold V_m of the measured volume[51], and is described by

$$e_i \sim \mathscr{N}(0, (\sigma \times E)^2), \qquad \text{with} \qquad E = \left\{ \begin{array}{ll} (y_i)^\alpha, & \text{if } y_i \geq V_m, \\ (V_m)^\alpha, & \text{if } y_i < V_m, \end{array} \right. \tag{5}$$

The rationale for this model is that, in a previous study, volume-independent standard deviation was statistically rejected, and proportional error demonstrated limited significance in Lewis lung carcinoma and breast cancer data[51]. Additionally, the Bayesian estimation of α and V_m in this model is an innovative aspect of our work.

The fourth error model, named Stud, employs the Student-t distribution with a fixed degree-of-freedom parameter (v) and constant dispersion, and can be expressed as

$$e_i \sim \text{StudentT}_{v=4}(0, \sigma^2). \tag{6}$$

Lastly, the fifth error model, Stud Prop, resembles the fourth but includes dispersion proportional to the measured volume as

$$e_i \sim \text{StudentT}_{v=4}(0, (\sigma \times y_i)^2). \tag{7}$$

The likelihood functions are obtained by treating each observation as independent realization of the error model and multiplying individual likelihood contributions.

Model comparison

We use three Bayesian criteria for model comparison. The first is the Bayesian Information Criterion (BIC)[60], defined as

3

$$BIC = k \log(N) - 2 \log(\widehat{L}), \tag{8}$$

where k is the number of model parameters, and $\widehat{L}$ is the maximized value of the likelihood function for the model.

Let us denote the likelihood of y_i under model j by $P(y_i \mid \theta_j, M_j)$ and the posterior expectation of θ by $\bar{\theta}$. The second criterion is the Deviance Information Criterion (DIC)[61], calculated as

$$DIC_j(y_i) = D(\bar{\theta}) + 2p_D, \tag{9}$$

where $D(\theta) = -2\log(P(y_i \mid \theta_j, M_j))$, $\overline{D(\theta)}$ is the posterior expectation of $D(\theta)$ and $p_D = \overline{D(\theta)} - D(\bar{\theta})$. Models with lower BIC and DIC values are preferable.

Although BIC is a consistent criterion and decisive when the correct model is among the candidates, it tends to heavily penalize free parameters[60,62]. Meanwhile, DIC is limited by its assumption of an approximately multivariate normal posterior distribution[61], which has not been verified for the models in this study. Furthermore, DIC inconsistency and reparametrization variance limit its applicability[61], as does its tendency to favor short-term prediction accuracy over model selection[63]. Therefore, DIC is accorded less weight in our model selection process.

The third criterion is the Bayes Factor (BF), representing the marginal likelihood (ML) ratio between two competing models (M_j, M_k), whose general expression is

$$BF_{jk}(y_i) = \frac{ML(M_j)}{ML(M_k)} = \frac{\int P(y_i \mid \theta_j, M_j)P(\theta_j \mid M_j)d\theta_j}{\int P(y_i \mid \theta_k, M_k)P(\theta_k \mid M_k)d\theta_k}, \tag{10}$$

where $P(\theta_j \mid M_j)$ is the prior distribution of the parameters for model j. The sensitivity of the BF to prior distributions, particularly with weak prior information, is a noted issue[64,65].

To mitigate this, we use the Fractional Bayes Factor (FBF)[66] with training fraction $0 < b < 1$ as

$$FBF_{jk}(y_i) = \frac{q_j(y_i; b)}{q_k(y_i; b)}, \tag{11}$$

where

$$q_j(y_i; b) = \frac{\int P(y_i \mid \theta_j, M_j)P(\theta_j \mid M_j)d\theta_j}{\int P(y_i \mid \theta_j, M_j)^b P(\theta_j \mid M_j)d\theta_j}. \tag{12}$$

Here, $q_j(y_i; b)$ represents a quantity similar to the marginal likelihood $ML(y_i; M_j)$, but where a fraction of the data, b, is used to reduce the sensitivity of the marginal likelihood to the choice of prior distribution[66]. We focus on the marginal likelihood (and $q_j(y; b)$) in our analysis. To compute the integrals in equation (12), we employ thermodynamic integration[67] using 30 temperature values between 0 and 1, implemented via the delayed-rejection adaptive Metropolis (DRAM) algorithm[68]. The Monte Carlo chain initialization for DRAM is based on results from the dual simulated annealing algorithm[69]. Uniform distributions $P(\theta_j \mid M_j) \sim U(0, \infty)$ are used as improper priors for the parameters, except for σ, which follows a normal distribution centered on the value computed by the dual simulated annealing algorithm with 10% standard deviation[69].

The fractional Bayes factor (FBF) approach does not impose assumptions about the prior and instead adjusts it using a subset of the data, offering greater robustness compared to methods applying penalizing terms in the mathematical formulation[70]. Nevertheless, the selection of b can introduce bias[71]. In this study, $b = 0.04$ was selected based on the following considerations:

- **Balancing Prior Influence and Data Sensitivity:** A goal in the choice of b is to use as small proportion of the data as possible to adequately modify the prior, while leaving as much data as possible to compute the marginal likelihood. Through an exploration of various b values across individuals, $b = 0.04$ was identified as an effective compromise, yielding robust marginal likelihood estimates without sacrificing too much data.
- **Empirical Evaluation Across Models:** Empirical assessments supported the choice of $b = 0.04$, which produced stable marginal likelihood estimates and consistent model comparisons across various error structures. This value effectively facilitated meaningful evaluations of competing likelihood models for most individuals in the dataset.
- **Mitigating Computational Bias:** Larger b values can overemphasize prior information, in that less data is used to compute the marginal likelihood, introducing artifacts, especially in cases of high data variability. Conversely, extremely small b values may lead to unreliable model assessments, due to not adequately modifying the prior. The selected $b = 0.04$ avoided these pitfalls, ensuring stability and reliability during thermo-

dynamic integration.Cases where $b = 0.04$ failed to produce reliable marginal likelihood estimates (less than 5% of the total cases) were excluded from the analysis. This exclusion did not affect the final conclusions, as the analysis focused exclusively on individuals with reliably calculated marginal likelihoods. Additionally, we performed Kolmogorov-Smirnov (KS) tests[72] on the posterior-drawn residuals to check how well they align with each model's implied distribution. Let $\theta^{(\ell)}$ denote the ℓ-th sample from the posterior distribution $(\ell = 1, \ldots, N_\ell)$, and define

$$R_i^{(\ell)} = f\left(t_i;\, \theta^{(\ell)}\right) - y_i.$$

For each model, we test the following sets of residuals against the corresponding theoretical distribution:

- **Norm model:** $\left\{R_i^{(\ell)}\right\}$ against the normal distribution.
- **Prop model:** $\left\{R_i^{(\ell)}/y_i\right\}$ against the normal distribution.
- **Thres model:** $\left\{R_i^{(\ell)}/E_i\right\}$ against the normal distribution, where E_i is given in equation (5).
- **Stud model:** $\left\{R_i^{(\ell)}\right\}$ against the Student-t distribution.
- **Stud Prop model:** $\left\{R_i^{(\ell)}/y_i\right\}$ against the Student-t distribution. For each individual, we then compute the fraction of KS tests (across all posterior draws) that do not reject the null hypothesis for that model, denoted as $\mathrm{fr(KS)}$.

For model comparison, we use experimental volume data from datasets of unperturbed Ehrlich and fibrosarcoma Sa-37 tumors growing in male BALB/c/Cenp mice[8], and unperturbed F3II mammary carcinoma (highly aggressive and metastatic) growing in both female and male BALB/c/Cenp mice[33]. Datasets in F3II mammary carcinoma-bearing male and female BALB/c/Cenp mice are analyzed separately.

The data were provided by Dr. Luis Enrique Bergues Cabrales, head of the Bioelectricity group, Centro Nacional de Electromagnetismo Aplicado, Universidad de Oriente and are available at https://github.com/era mireztorres/tumor_error_model. Simulations were conducted using Python 3[73], specifically the *scipy.optimize*[74] and *pymcmcstat*[75] packages, on a 4-core high-performance computing (HPC) system with 256 GB of RAM at the Instituto Universitario de Matemáticas y Aplicaciones (IUMA), Universidad de Zaragoza, Spain.

Results

In Tables 1, 2, 3 and 4 (see Appendix), we present the model fit indicators for each individual in the unperturbed Ehrlich tumor dataset, highlighting the best results in bold. The Stud Prop model consistently provides the best results for the BIC indicator across all individuals. For the DIC indicator, the Thres model performs best for seven individuals, whereas the Norm model leads for the remaining three. Regarding ML, the Prop model is optimal for one individual, the Norm model for five, and the Stud model for four. For fr(KS), the Prop, Thres, and Stud Prop models excel for 3, 5, and 2 individuals, respectively.

Even though these results prevent us from concluding that one model is superior to the others, they show clearly that models with variable dispersion (Prop, Thres and Stud Prop) are markedly better than constant dispersion models (Norm and Stud), particularly for fr(KS). This suggests that likelihoods with variable dispersion are better suited to modeling errors in tumor growth studies. No significant gender-based differences were observed in the results.

For the unperturbed fibrosarcoma Sa-37 tumor dataset (Tables 5, 6, 7 and 8), variable dispersion models show clear superiority in BIC and fr(KS). Nevertheless, DIC and ML results are split evenly across individuals. In the unperturbed F3II mammary carcinoma dataset for male BALB/C/Cenp mice (Tables 9, 10, 11 and 12), variable dispersion models outperform in DIC and fr(KS), whereas results are close for BIC and ML (6 individuals vs. 5). Similarly, in female BALB/C/Cenp mice (Tables 13, 14, 15 and 16), variable dispersion models yield better outcomes for all indicators: BIC, DIC, ML, and fr(KS).

Tables 17, 18, 19 and 20 summarize average results across all individuals. These tables were constructed by taking the arithmetic mean of each model criterion (BIC, DIC, ML, fr(KS)) across all individuals in the respective dataset. The averages reveal varied performance across models. For the unperturbed Ehrlich (Table 17) and fibrosarcoma Sa-37 datasets (Table 18): The Thres model demonstrates the best mean DIC and fr(KS) values. The Stud Prop and Norm models excel in mean BIC and ML, respectively. For the unperturbed F3II mammary carcinoma dataset (Tables 19 and 20): The Prop model excels in mean BIC and fr(KS). The Thres model leads in mean DIC. The Norm model (male mice) and Prop model (female mice) are best in mean ML. These results support the conclusion that adding variable dispersion likelihood to mixed-effects models is an excellent approach for estimating tumor growth model parameters.

In terms of prediction accuracy, all models demonstrated comparable performance to individual 1 from the unperturbed fibrosarcoma Sa-37 tumor dataset, a trend consistent across all four histological tumor types. The Prop, and Thres models exhibited greater variability in the prediction cone, represented by the 95% credible intervals of posterior trajectories, yet the maximum *a posteriori* (MAP) trajectory closely aligned with the observed data in all five models (Figure 1).

Discussion

The comparison criteria and visualisations of model fits reveal no definitive best model. To determine priorities, we analyze the advantages and limitations of each criterion and explore clinical implications.
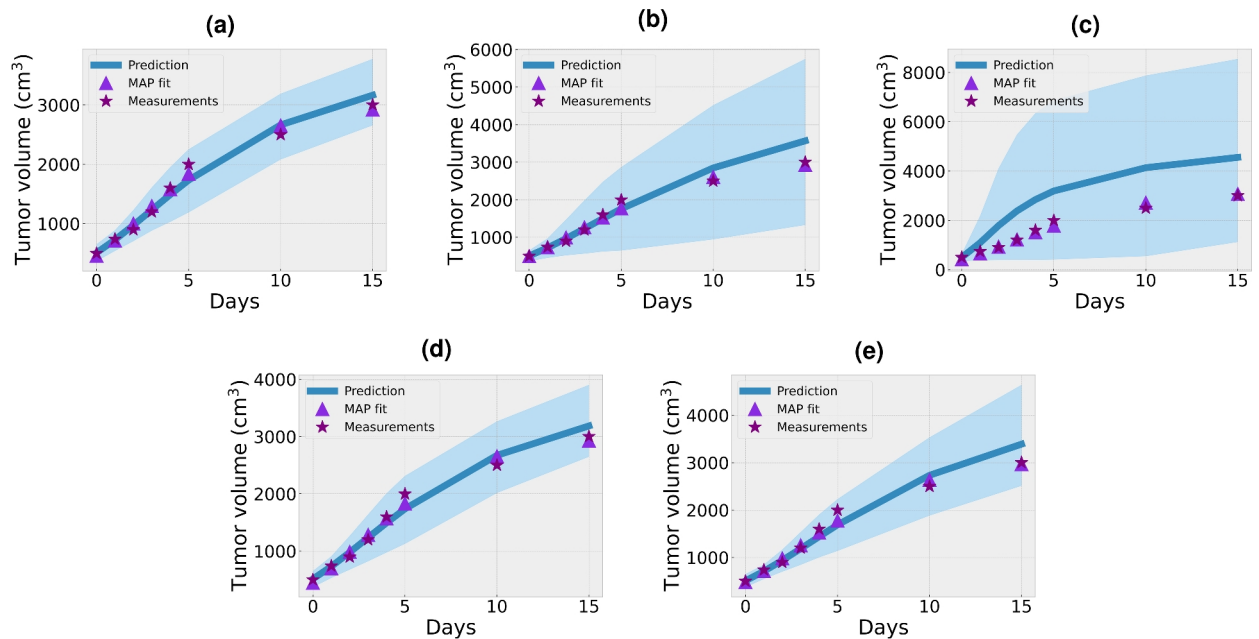
Fig. 1. Gompertz fits with five error models for the mouse (individual 1) bearing the unperturbed fibrosarcoma Sa-37 tumor. Norm (a), Prop (b), Thres (c), Stud (d), and Stud Prop (e). Shaded areas indicate the 95% credible intervals of posterior trajectories.

Since the BIC heavily penalizes the number of model parameters, it naturally favors the simpler models. Its value would be decisive if for some data it favored the more complex model. This is not the case here, because a simple model is always favored by BIC, as expected, and never the Thres model, which is the most complex one. Nevertheless, for some particular datasets, the Thres model yields better BIC values than other simpler models. For instance, the third individual from the Ehrlich tumor data (Table 1) demonstrates this. It indicates that, for the BIC, the complexity introduced by the additional parameters of the Thres model (α and V_m) is justified by the quality of the fit. This finding, along with the relatively good DIC values (Table 2) and fr(KS) values (Table 4) it achieves, supports recommending the Thres model. Another advantage of the Thres model is in the clinical interpretability of it's parameters. For example, the parameter V_m could represent the threshold tumor volume at which the surface/contour begins to become irregular, or where tumor growth becomes irreversible, with or without anticancer therapy (either oncospecific or under study). Experimental observations indicate that tumor surfaces become irregular from 2.0 cm^3[8,32,37,76]. Nevertheless, regulatory and ethical considerations for laboratory experiments with animals[10,11,34,35] prevent the measurement of tumor volume when it exceeds 10% of body weight. This must be considered in all experimental and similar studies. The differences in ML values obtained may be categorized as anecdotal[77].

A major concern is *fr(KS)*, for which very small values are observed for constant-dispersion models. The Norm and Stud models show zero values for several individuals (see Tables 4, 8, 12 and 16). This indicates that very few or no trajectories drawn from the posterior generate residuals with non-rejected KS tests. This surprising result challenges the use of the two constant-dispersion models assessed and suggests revisiting the entire constant-dispersion paradigm in the absence of information to the contrary. Based solely on quantitative criteria, our results suggest that the Norm model is not the best to fit Gompertz growth.

This raises an important question: what are the clinical implications of introducing variable-dispersion errors instead of traditional least-squares fitting? The first consequence is that predictions accounting for variable-dispersion errors are more likely to indicate actual tumor volumes exceeding measured volumes. Second, standard cancer therapies based on closely fitting models to data may require theoretical reformulation. Nevertheless, we do not recommend discarding the Norm model entirely, as it is widely used in existing research and serves as a baseline for evaluating cancer therapy improvements. We believe that the best option is always to apply the Norm model in the model fits, either with the least-squares approach or the Bayesian approach, and complement it with an estimation of variable dispersion.

Although we argue that the Gompertz model remains preferable for modeling most solid tumors, it's appropriateness or otherwise does not undermine the applicability of our results. This is because Gompertz growth is used only in the deterministic component of the model. Measurement challenges, such as unsharp boundaries and irregular tumor shapes, persist in clinical practice regardless of the deterministic growth model.

To resolve the likelihood dilemma, we propose that studies should include the Norm model to ensure compatibility with previous results. Nevertheless, at least one model with variable dispersion should also be included for a more realistic representation of tumor size. Both approaches should complement each other. We recommend the Thres model but suggest the Prop or Stud Prop models for researchers seeking reduced computational complexity.

Although the Gompertz growth model was used to fit experimental tumor data from mice, its generalizability to clinical practice requires careful evaluation. Studies have shown that Gompertzian growth can describe human tumor kinetics in specific cases, such as the foundational work by Skipper et al.[78] and patient-specific findings by Grassberger et al.[79]. Nevertheless, alternative growth patterns, such as near-linear growth for certain tumor types and stages[80], have been observed. Kidney tumors under active surveillance exemplify this, with smaller tumors exhibiting near-linear growth[81,82] and Gompertzian dynamics emerging in long-term studies[6]. This variability supports including models like the Bertalanffy model, which provides biologically motivated alternatives that converge to similar growth trajectories[83].

Importantly, even though the Gompertz model was used as the deterministic component in our analysis, the conclusions of our study are not model-dependent. The deterministic growth function interacts with the likelihood and error models only to generate predicted data for comparison with observations. Measurement challenges-such as irregular tumor shapes, unsharp boundaries, and imaging variability-are pervasive in both experimental and clinical settings and are independent of the underlying deterministic model. These challenges underscore the critical role of appropriate likelihood models, as our results demonstrate that error model selection significantly affects parameter estimation and predictive accuracy.

To address these challenges, we propose a dual likelihood modeling approach. The Norm model should remain a baseline to ensure compatibility with previous studies, but incorporating a likelihood model that accounts for variable dispersion is essential for capturing tumor heterogeneity. Our results highlight the suitability of Thres model, as it balances robustness and interpretability. For researchers seeking less computational complexity, the Prop or Stud Prop models provide viable alternatives, improving realism over the Norm model. This complementary use of baseline and advanced likelihood models ensures reproducibility and accommodates the increasing complexity of tumor growth dynamics.

Advantages of the proposed methodology

The methodology presented in this study offers significant advancements over existing approaches for modeling tumor growth dynamics. Unlike many published methods that rely on fixed error models or likelihood functions, our approach systematically evaluates multiple likelihood models, including those that account for variable dispersion. This flexibility allows for a more accurate representation of tumor heterogeneity and measurement variability, which are critical factors in understanding cancer progression[84].

In addition, the incorporation of diverse growth models, such as Gompertz and Bertalanffy, enables the application of our methodology to a wide range of tumor types and stages. This adaptability is essential for addressing the variability observed in experimental and clinical tumor growth data[51].

The method also addresses practical challenges in tumor measurement, such as irregular shapes, unsharp boundaries, and imaging inconsistencies, making it particularly relevant for real-world applications. Furthermore, the detailed computational pipeline and open-source implementation ensure reproducibility, allowing other researchers to replicate and extend our findings effectively.

These features establish the proposed methodology as a robust and versatile tool for studying tumor growth dynamics, complementing and extending the capabilities of existing approaches in the field.

Limitations of the proposed methodology

Despite the significant advancements it brings, the proposed methodology does have some limitations. First, the selection of growth and likelihood models depends on the availability and quality of data. In scenarios where limited data points are available, model overfitting may occur, potentially affecting the reliability of predictions.

Second, although our approach integrates Bayesian and Frequentist (by KS test) methods to enhance robustness, the computational complexity of Bayesian analysis, particularly with multiple b-values, can be resource-intensive. This may limit its applicability in high-throughput scenarios without access to sufficient computational resources.

Finally, the approach has been only validated on specific tumor types and datasets; nevertheless, its adaptability to other datasets and generalizability to other cancer types, particularly those with non-solid growth dynamics, should be fully explored. Future studies could address these limitations by incorporating broader datasets, testing additional tumor types, and improving computational efficiency through optimized algorithms.

Conclusions

Our analysis highlights the advantages of variable dispersion error models over constant variance models. Although none of the three models (Prop, Thres, Stud Prop) consistently excels, the Thres model emerges as a robust choice for most scenarios, with solid performance across criteria and meaningful clinical interpretability. In any case, our results indicate that the standard constant-variance Normal error model should remain part of the analysis toolkit, to provide a complementary analytic reference in addition to more flexible error models.

Finally, although the Gompertz model was employed as the deterministic component, our conclusions remain valid regardless of the specific growth model used. Incorporating both constant and variable dispersion likelihoods offers a balanced approach for future tumor growth modeling studies.

Data availability

The datasets used and/or analyzed during the present study as well as the phyton code are available at https://github.com/eramireztorres/tumor_error_model.

Appendix

Individual	Norm	Prop	Thres	Stud	Stud prop
1	1.0727×10^2	1.0642×10^2	1.0736×10^2	1.0727×10^2	$\mathbf{1.0320 \times 10^2}$
2	9.7720×10^1	1.0420×10^2	9.2940×10^1	9.7718×10^1	$\mathbf{8.8783 \times 10^1}$
3	8.7728×10^1	1.0331×10^2	8.4857×10^1	8.7739×10^1	$\mathbf{8.0731 \times 10^1}$
4	1.0848×10^2	1.0576×10^2	1.0607×10^2	1.0849×10^2	$\mathbf{1.0191 \times 10^2}$
5	9.6301×10^1	1.0343×10^2	9.3257×10^1	9.6300×10^1	$\mathbf{8.9101 \times 10^1}$
6	8.8998×10^1	1.0369×10^2	9.0639×10^1	8.9014×10^1	$\mathbf{8.8342 \times 10^1}$
7	9.4826×10^1	1.0403×10^2	9.4389×10^1	9.4828×10^1	$\mathbf{9.0952 \times 10^1}$
8	9.0340×10^1	1.0344×10^2	9.0425×10^1	9.0351×10^1	$\mathbf{8.7231 \times 10^1}$
9	1.0644×10^2	1.0487×10^2	1.0204×10^2	1.0643×10^2	$\mathbf{9.7865 \times 10^1}$
10	1.0166×10^2	1.0416×10^2	9.8188×10^1	1.0166×10^2	$\mathbf{9.4033 \times 10^1}$

Table 1. Model comparison criteria: BIC for the unperturbed Ehrlich tumor data.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	1.5023×10^2	1.4454×10^2	$\mathbf{1.4140 \times 10^2}$	1.5193×10^2	1.4661×10^2
2	1.4192×10^2	1.4461×10^2	$\mathbf{1.3979 \times 10^2}$	1.4011×10^2	1.4662×10^2
3	1.4534×10^2	1.4462×10^2	$\mathbf{1.4091 \times 10^2}$	1.4132×10^2	1.4482×10^2
4	$\mathbf{1.3691 \times 10^2}$	1.4538×10^2	1.3758×10^2	1.4798×10^2	1.4509×10^2
5	1.4015×10^2	1.4392×10^2	$\mathbf{1.3814 \times 10^2}$	1.4651×10^2	1.4328×10^2
6	$\mathbf{1.3389 \times 10^2}$	1.4354×10^2	1.3952×10^2	1.3631×10^2	1.4444×10^2
7	$\mathbf{1.3800 \times 10^2}$	1.4464×10^2	1.3853×10^2	1.4151×10^2	1.4725×10^2
8	1.4667×10^2	1.4331×10^2	$\mathbf{1.3855 \times 10^2}$	1.3970×10^2	1.4305×10^2
9	1.4455×10^2	1.4446×10^2	$\mathbf{1.4114 \times 10^2}$	1.4259×10^2	1.4589×10^2
10	1.4193×10^2	1.4520×10^2	$\mathbf{1.3688 \times 10^2}$	1.4486×10^2	1.4272×10^2

Table 2. Model comparison: DIC for the unperturbed Ehrlich tumor data.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	$\mathbf{9.8104 \times 10^{-1}}$	9.7912×10^{-1}	8.5000×10^{-1}	9.0190×10^{-1}	6.0503×10^{-1}
2	9.5885×10^{-1}	9.5971×10^{-1}	8.0709×10^{-1}	$\mathbf{9.6392 \times 10^{-1}}$	5.0000×10^{-1}
3	9.6500×10^{-1}	$\mathbf{9.6540 \times 10^{-1}}$	7.8951×10^{-1}	9.4998×10^{-1}	1.4112×10^{-1}
4	$\mathbf{9.7002 \times 10^{-1}}$	9.5009×10^{-1}	9.5344×10^{-1}	9.5500×10^{-1}	5.2574×10^{-1}
5	$\mathbf{9.9567 \times 10^{-1}}$	9.7469×10^{-1}	8.3813×10^{-1}	9.7026×10^{-1}	4.5033×10^{-1}
6	$\mathbf{9.6170 \times 10^{-1}}$	9.4000×10^{-1}	8.2500×10^{-1}	9.4009×10^{-1}	9.3654×10^{-1}
7	$\mathbf{9.7164 \times 10^{-1}}$	8.2125×10^{-1}	8.0345×10^{-1}	9.3366×10^{-1}	3.0000×10^{-1}
8	9.6327×10^{-1}	8.3279×10^{-1}	5.4741×10^{-1}	$\mathbf{9.6890 \times 10^{-1}}$	3.5770×10^{-1}
9	9.6800×10^{-1}	9.2536×10^{-1}	8.1936×10^{-1}	$\mathbf{9.8101 \times 10^{-1}}$	6.2238×10^{-1}
10	9.6935×10^{-1}	9.5010×10^{-1}	8.6951×10^{-1}	$\mathbf{9.9764 \times 10^{-1}}$	4.0441×10^{-1}

Table 3. Model comparison: ML for the unperturbed Ehrlich tumor data.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	0.0000	3.0612×10^{-1}	$\mathbf{3.1912 \times 10^{-1}}$	8.0000×10^{-5}	3.1520×10^{-1}
2	6.0000×10^{-4}	$\mathbf{2.6700 \times 10^{-1}}$	2.6612×10^{-1}	2.4000×10^{-4}	2.5576×10^{-1}
3	2.0000×10^{-4}	2.5188×10^{-1}	$\mathbf{2.7088 \times 10^{-1}}$	8.0000×10^{-5}	2.6776×10^{-1}
4	3.6000×10^{-4}	2.6288×10^{-1}	2.5996×10^{-1}	8.0000×10^{-5}	$\mathbf{2.7116 \times 10^{-1}}$
5	2.4000×10^{-4}	$\mathbf{2.7832 \times 10^{-1}}$	2.6468×10^{-1}	1.2000×10^{-4}	2.6648×10^{-1}
6	1.2000×10^{-4}	$\mathbf{2.7336 \times 10^{-1}}$	2.6656×10^{-1}	0.0000	2.6528×10^{-1}
7	8.0000×10^{-5}	2.6392×10^{-1}	$\mathbf{2.6708 \times 10^{-1}}$	5.6000×10^{-4}	2.6340×10^{-1}
8	1.2000×10^{-4}	2.6816×10^{-1}	$\mathbf{2.8264 \times 10^{-1}}$	2.8000×10^{-4}	2.6424×10^{-1}
9	1.6000×10^{-4}	2.7136×10^{-1}	2.7088×10^{-1}	4.8000×10^{-4}	$\mathbf{2.7704 \times 10^{-1}}$
10	1.2000×10^{-4}	2.7336×10^{-1}	$\mathbf{2.8328 \times 10^{-1}}$	4.8000×10^{-4}	2.8272×10^{-1}

Table 4. Model comparison: fr(KS) for the unperturbed Ehrlich tumor data.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	1.0444×10^{2}	1.0428×10^{2}	1.0561×10^{2}	1.0444×10^{2}	$\mathbf{1.0145 \times 10^{2}}$
2	1.1210×10^{2}	1.0754×10^{2}	1.1150×10^{2}	1.1210×10^{2}	$\mathbf{1.0734 \times 10^{2}}$
3	9.8797×10^{1}	1.0314×10^{2}	1.0229×10^{2}	9.8798×10^{1}	$\mathbf{9.8137 \times 10^{1}}$
4	1.0643×10^{2}	1.0565×10^{2}	1.0916×10^{2}	1.0644×10^{2}	$\mathbf{1.0444 \times 10^{2}}$
5	1.0730×10^{2}	1.0468×10^{2}	1.0641×10^{2}	1.0731×10^{2}	$\mathbf{1.0226 \times 10^{2}}$
6	1.0006×10^{2}	1.0339×10^{2}	1.0348×10^{2}	1.0007×10^{2}	$\mathbf{9.9324 \times 10^{1}}$
7	1.0637×10^{2}	1.0514×10^{2}	1.0755×10^{2}	1.0638×10^{2}	$\mathbf{1.0340 \times 10^{2}}$
8	1.0610×10^{2}	1.0438×10^{2}	1.0579×10^{2}	1.0611×10^{2}	$\mathbf{1.0165 \times 10^{2}}$
9	1.0542×10^{2}	1.0442×10^{2}	1.0627×10^{2}	1.0543×10^{2}	$\mathbf{1.0214 \times 10^{2}}$
10	1.0312×10^{2}	1.0419×10^{2}	1.0542×10^{2}	1.0312×10^{2}	$\mathbf{1.0126 \times 10^{2}}$

Table 5. Model comparison: BIC for the unperturbed fibrosarcoma Sa-37 tumor data.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	1.3919×10^{2}	1.4263×10^{2}	$\mathbf{1.3485 \times 10^{2}}$	1.4461×10^{2}	1.4275×10^{2}
2	1.4604×10^{2}	1.4306×10^{2}	$\mathbf{1.3880 \times 10^{2}}$	1.4444×10^{2}	1.4482×10^{2}
3	$\mathbf{1.3385 \times 10^{2}}$	1.4089×10^{2}	1.3557×10^{2}	1.3906×10^{2}	1.4260×10^{2}
4	1.3967×10^{2}	1.4191×10^{2}	1.3780×10^{2}	$\mathbf{1.3557 \times 10^{2}}$	1.4168×10^{2}
5	1.4113×10^{2}	1.4147×10^{2}	$\mathbf{1.3759 \times 10^{2}}$	1.3956×10^{2}	1.4338×10^{2}
6	1.3472×10^{2}	1.4213×10^{2}	$\mathbf{1.3436 \times 10^{2}}$	1.3913×10^{2}	1.4054×10^{2}
7	$\mathbf{1.3671 \times 10^{2}}$	1.4099×10^{2}	1.3738×10^{2}	1.4300×10^{2}	1.4205×10^{2}
8	$\mathbf{1.3422 \times 10^{2}}$	1.4181×10^{2}	1.3907×10^{2}	1.3513×10^{2}	1.4229×10^{2}
9	$\mathbf{1.3479 \times 10^{2}}$	1.4158×10^{2}	1.3965×10^{2}	1.3486×10^{2}	1.4164×10^{2}
10	1.4309×10^{2}	1.3946×10^{2}	$\mathbf{1.3532 \times 10^{2}}$	1.4040×10^{2}	1.4254×10^{2}

Table 6. Model comparison: DIC for the unperturbed fibrosarcoma Sa-37 tumor data.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	9.1053×10^{-1}	9.4307×10^{-1}	$\mathbf{9.6001 \times 10^{-1}}$	9.6000×10^{-1}	9.1080×10^{-1}
2	9.6030×10^{-1}	9.3006×10^{-1}	$\mathbf{9.6529 \times 10^{-1}}$	9.5500×10^{-1}	8.6810×10^{-1}
3	9.3881×10^{-1}	$\mathbf{9.8643 \times 10^{-1}}$	9.5324×10^{-1}	9.5906×10^{-1}	9.8331×10^{-1}
4	$\mathbf{9.9649 \times 10^{-1}}$	9.1477×10^{-1}	9.5803×10^{-1}	9.7014×10^{-1}	9.1399×10^{-1}
5	$\mathbf{9.8051 \times 10^{-1}}$	9.2111×10^{-1}	9.7750×10^{-1}	9.6500×10^{-1}	9.4022×10^{-1}
6	9.2955×10^{-1}	$\mathbf{9.7836 \times 10^{-1}}$	9.6145×10^{-1}	9.4851×10^{-1}	9.5061×10^{-1}
7	9.5123×10^{-1}	8.7347×10^{-1}	9.6504×10^{-1}	$\mathbf{9.9152 \times 10^{-1}}$	9.4500×10^{-1}
8	9.4456×10^{-1}	9.2135×10^{-1}	$\mathbf{9.6423 \times 10^{-1}}$	9.6025×10^{-1}	9.2365×10^{-1}
9	9.4500×10^{-1}	9.0990×10^{-1}	9.6205×10^{-1}	$\mathbf{9.7127 \times 10^{-1}}$	9.3321×10^{-1}
10	$\mathbf{9.9794 \times 10^{-1}}$	8.1350×10^{-1}	9.6793×10^{-1}	9.5800×10^{-1}	9.0130×10^{-1}

Table 7. Model comparison criteria: ML for the unperturbed fibrosarcoma Sa-37 tumor data.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	2.0000×10^{-4}	2.7420×10^{-1}	$\mathbf{2.9508 \times 10^{-1}}$	2.4000×10^{-4}	2.8296×10^{-1}
2	8.0000×10^{-5}	$\mathbf{2.8636 \times 10^{-1}}$	2.8208×10^{-1}	2.8000×10^{-4}	2.8384×10^{-1}
3	5.2000×10^{-4}	2.6648×10^{-1}	$\mathbf{2.7996 \times 10^{-1}}$	6.0000×10^{-4}	2.6416×10^{-1}
4	1.6000×10^{-4}	2.7260×10^{-1}	2.7180×10^{-1}	2.0000×10^{-4}	$\mathbf{2.8528 \times 10^{-1}}$
5	3.6000×10^{-4}	2.7668×10^{-1}	$\mathbf{3.0132 \times 10^{-1}}$	8.4000×10^{-4}	2.6724×10^{-1}
6	8.0000×10^{-5}	2.7208×10^{-1}	$\mathbf{2.9612 \times 10^{-1}}$	2.4000×10^{-4}	2.6764×10^{-1}
7	2.8000×10^{-4}	2.7660×10^{-1}	$\mathbf{2.7960 \times 10^{-1}}$	2.4000×10^{-4}	2.6412×10^{-1}
8	1.2000×10^{-4}	$\mathbf{2.8528 \times 10^{-1}}$	2.5116×10^{-1}	5.6000×10^{-4}	2.6692×10^{-1}
9	0.0000	$\mathbf{2.9320 \times 10^{-1}}$	2.6760×10^{-1}	1.2000×10^{-4}	2.7468×10^{-1}
10	2.0000×10^{-4}	$\mathbf{2.8400 \times 10^{-1}}$	2.8304×10^{-1}	1.6000×10^{-4}	2.6556×10^{-1}

Table 8. Model comparison criteria: fr(KS) for the unperturbed fibrosarcoma Sa-37 tumor data.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	$\mathbf{1.0727 \times 10^{2}}$	1.2827×10^{2}	1.3178×10^{2}	1.3282×10^{2}	1.3137×10^{2}
2	1.3002×10^{2}	1.2532×10^{2}	1.2388×10^{2}	$\mathbf{1.2023 \times 10^{2}}$	1.2855×10^{2}
3	$\mathbf{1.1789 \times 10^{2}}$	1.2718×10^{2}	1.3096×10^{2}	1.3451×10^{2}	1.3037×10^{2}
4	1.3137×10^{2}	9.8670×10^{1}	1.0193×10^{2}	1.0947×10^{2}	$\mathbf{9.8264 \times 10^{1}}$
5	$\mathbf{1.0700 \times 10^{2}}$	1.3978×10^{2}	1.4017×10^{2}	1.4418×10^{2}	1.4145×10^{2}
6	1.6338×10^{2}	$\mathbf{1.2054 \times 10^{2}}$	1.3039×10^{2}	1.2969×10^{2}	1.2404×10^{2}
7	1.2652×10^{2}	$\mathbf{1.0322 \times 10^{2}}$	1.0761×10^{2}	1.1800×10^{2}	1.0680×10^{2}
8	1.1481×10^{2}	1.2208×10^{2}	$\mathbf{1.1397 \times 10^{2}}$	1.1470×10^{2}	1.2060×10^{2}
9	$\mathbf{1.1218 \times 10^{2}}$	1.1844×10^{2}	1.2283×10^{2}	1.3247×10^{2}	1.2186×10^{2}
10	1.3684×10^{2}	$\mathbf{1.1142 \times 10^{2}}$	1.2437×10^{2}	1.1762×10^{2}	1.1262×10^{2}
11	$\mathbf{1.1527 \times 10^{2}}$	1.1752×10^{2}	1.2576×10^{2}	1.2596×10^{2}	1.2105×10^{2}

Table 9. Model comparison criteria: BIC for the unperturbed F3II mammary carcinoma growing in male BALB/C/Cenp mice.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	1.5023×10^2	1.6450×10^2	$\mathbf{8.4985 \times 10^1}$	1.2774×10^2	1.5630×10^2
2	1.6303×10^2	1.6177×10^2	$\mathbf{9.9921 \times 10^1}$	1.1922×10^2	1.3766×10^2
3	1.9594×10^2	1.6284×10^2	$\mathbf{3.6005 \times 10^1}$	1.2687×10^2	1.4369×10^2
4	1.2218×10^2	1.2447×10^2	1.2182×10^2	1.1401×10^2	$\mathbf{1.1041 \times 10^2}$
5	2.5272×10^2	1.6855×10^2	1.8744×10^2	1.5096×10^2	$\mathbf{1.5066 \times 10^2}$
6	1.6869×10^2	1.4198×10^2	1.5358×10^2	1.3024×10^2	$\mathbf{1.2607 \times 10^2}$
7	1.4673×10^2	1.4553×10^2	1.3480×10^2	$\mathbf{1.1656 \times 10^2}$	1.2915×10^2
8	1.5271×10^2	1.4685×10^2	1.2228×10^2	$\mathbf{1.1610 \times 10^2}$	1.3409×10^2
9	2.2668×10^2	1.5631×10^2	$\mathbf{1.1825 \times 10^2}$	1.2891×10^2	1.2990×10^2
10	1.6661×10^2	1.2886×10^2	$\mathbf{1.0485 \times 10^2}$	1.1597×10^2	1.1721×10^2
11	1.6153×10^2	1.5749×10^2	$\mathbf{7.0909 \times 10^1}$	1.1097×10^2	1.3294×10^2

Table 10. Model comparison criteria: DIC for the unperturbed F3II mammary carcinoma growing in male BALB/C/Cenp mice.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	$\mathbf{9.8104 \times 10^{-1}}$	9.4286×10^{-1}	9.0037×10^{-1}	9.1177×10^{-1}	9.5231×10^{-1}
2	9.7665×10^{-1}	9.6634×10^{-1}	9.4567×10^{-1}	$\mathbf{9.9971 \times 10^{-1}}$	9.3595×10^{-1}
3	$\mathbf{9.9542 \times 10^{-1}}$	9.6478×10^{-1}	9.0111×10^{-1}	9.0001×10^{-1}	9.6798×10^{-1}
4	9.8289×10^{-1}	$\mathbf{9.9193 \times 10^{-1}}$	8.2038×10^{-1}	8.8880×10^{-1}	9.8229×10^{-1}
5	9.9500×10^{-1}	$\mathbf{9.9928 \times 10^{-1}}$	9.9677×10^{-1}	7.1568×10^{-1}	9.6570×10^{-1}
6	$\mathbf{9.9573 \times 10^{-1}}$	9.7007×10^{-1}	9.0251×10^{-1}	8.0235×10^{-1}	9.7043×10^{-1}
7	9.9000×10^{-1}	$\mathbf{9.9420 \times 10^{-1}}$	8.9145×10^{-1}	9.4039×10^{-1}	9.4545×10^{-1}
8	9.6276×10^{-1}	$\mathbf{9.9654 \times 10^{-1}}$	8.2234×10^{-1}	9.5016×10^{-1}	9.6134×10^{-1}
9	9.9867×10^{-1}	$\mathbf{9.9920 \times 10^{-1}}$	8.7001×10^{-1}	8.9893×10^{-1}	9.6678×10^{-1}
10	9.8818×10^{-1}	9.8551×10^{-1}	8.7890×10^{-1}	$\mathbf{9.9033 \times 10^{-1}}$	9.8515×10^{-1}
11	$\mathbf{9.9308 \times 10^{-1}}$	9.9137×10^{-1}	9.1094×10^{-1}	9.2335×10^{-1}	9.7494×10^{-1}

Table 11. Model comparison criteria: ML for the unperturbed F3II mammary carcinoma growing in male BALB/C/Cenp mice.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	0.0000	$\mathbf{3.2484 \times 10^{-1}}$	2.0104×10^{-1}	1.6000×10^{-4}	1.6264×10^{-1}
2	0.0000	$\mathbf{3.2112 \times 10^{-1}}$	2.4064×10^{-1}	0.0000	1.9604×10^{-1}
3	0.0000	$\mathbf{3.0220 \times 10^{-1}}$	2.5896×10^{-1}	0.0000	1.5552×10^{-1}
4	1.5600×10^{-3}	$\mathbf{3.2900 \times 10^{-1}}$	1.1592×10^{-1}	1.0000×10^{-3}	2.3020×10^{-1}
5	0.0000	$\mathbf{5.2788 \times 10^{-1}}$	3.8756×10^{-1}	0.0000	2.6836×10^{-1}
6	0.0000	$\mathbf{8.6772 \times 10^{-1}}$	1.6388×10^{-1}	0.0000	4.5332×10^{-1}
7	1.6000×10^{-4}	$\mathbf{1.4812 \times 10^{-1}}$	8.3040×10^{-2}	0.0000	7.8520×10^{-2}
8	0.0000	$\mathbf{4.9628 \times 10^{-1}}$	1.5340×10^{-1}	0.0000	2.6296×10^{-1}
9	3.6000×10^{-4}	2.1572×10^{-1}	$\mathbf{2.6780 \times 10^{-1}}$	4.0000×10^{-4}	1.2556×10^{-1}
10	9.2000×10^{-4}	$\mathbf{4.3476 \times 10^{-1}}$	1.6496×10^{-1}	8.0000×10^{-4}	3.5620×10^{-1}
11	0.0000	$\mathbf{2.4992 \times 10^{-1}}$	1.8936×10^{-1}	0.0000	1.4524×10^{-1}

Table 12. Model comparison criteria: fr(KS) for the unperturbed F3II mammary carcinoma growing in male BALB/C/Cenp mice.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	1.1228×10^2	$\mathbf{1.1019 \times 10^2}$	1.1247×10^2	1.1281×10^2	1.1289×10^2
2	1.2542×10^2	$\mathbf{1.0563 \times 10^2}$	1.1002×10^2	1.2790×10^2	1.0806×10^2
3	1.3505×10^2	$\mathbf{1.2278 \times 10^2}$	1.2754×10^2	1.3433×10^2	1.2665×10^2
4	1.2785×10^2	$\mathbf{1.2199 \times 10^2}$	1.2567×10^2	1.3092×10^2	1.2578×10^2
5	1.1956×10^2	$\mathbf{1.0805 \times 10^2}$	1.2472×10^2	1.2193×10^2	1.1057×10^2
6	1.6513×10^2	$\mathbf{1.1802 \times 10^2}$	1.2243×10^2	1.4217×10^2	1.2112×10^2
7	1.1853×10^2	$\mathbf{1.0583 \times 10^2}$	1.0942×10^2	1.2019×10^2	1.0738×10^2
8	1.2057×10^2	$\mathbf{9.3843 \times 10^1}$	1.2496×10^2	1.2146×10^2	9.6140×10^1
9	1.2959×10^2	$\mathbf{1.1672 \times 10^2}$	1.3399×10^2	1.3265×10^2	1.1967×10^2

Table 13. Model comparison criteria: BIC for the unperturbed F3II mammary carcinoma growing in female BALB/C/Cenp mice.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	1.4894×10^2	1.6061×10^2	$\mathbf{9.2930 \times 10^1}$	1.1043×10^2	1.4012×10^2
2	1.6089×10^2	1.4280×10^2	$\mathbf{1.2098 \times 10^2}$	1.2501×10^2	1.2468×10^2
3	2.1178×10^2	1.5028×10^2	$\mathbf{8.1812 \times 10^1}$	1.3812×10^2	1.3030×10^2
4	1.8125×10^2	1.3978×10^2	1.3329×10^2	1.2131×10^2	$\mathbf{1.1506 \times 10^2}$
5	1.5623×10^2	1.4522×10^2	1.1596×10^2	$\mathbf{1.1405 \times 10^2}$	1.2332×10^2
6	2.8666×10^2	1.3520×10^2	$\mathbf{5.8875 \times 10^1}$	1.3187×10^2	1.2956×10^2
7	1.6586×10^2	1.3600×10^2	1.2370×10^2	$\mathbf{1.1264 \times 10^2}$	1.1699×10^2
8	1.5675×10^2	1.1702×10^2	1.3681×10^2	1.2048×10^2	$\mathbf{1.0207 \times 10^2}$
9	1.9100×10^2	1.3398×10^2	1.2959×10^2	1.2679×10^2	$\mathbf{1.0606 \times 10^2}$

Table 14. Model comparison criteria: DIC for the unperturbed F3II mammary carcinoma growing in female BALB/C/Cenp mice.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	9.7803×10^{-1}	$\mathbf{9.9077 \times 10^{-1}}$	9.4578×10^{-1}	9.0740×10^{-1}	8.5190×10^{-1}
2	9.9043×10^{-1}	$\mathbf{9.9444 \times 10^{-1}}$	8.3303×10^{-1}	9.6814×10^{-1}	9.7283×10^{-1}
3	$\mathbf{9.9675 \times 10^{-1}}$	9.9652×10^{-1}	8.0320×10^{-1}	9.6004×10^{-1}	8.9394×10^{-1}
4	9.9569×10^{-1}	9.9542×10^{-1}	9.7649×10^{-1}	9.6750×10^{-1}	$\mathbf{9.9832 \times 10^{-1}}$
5	9.9012×10^{-1}	$\mathbf{9.9324 \times 10^{-1}}$	8.1280×10^{-1}	9.6062×10^{-1}	9.0304×10^{-1}
6	9.9842×10^{-1}	9.8355×10^{-1}	7.3480×10^{-1}	9.8724×10^{-1}	$\mathbf{9.9959 \times 10^{-1}}$
7	9.8376×10^{-1}	$\mathbf{9.9303 \times 10^{-1}}$	7.9099×10^{-1}	9.4567×10^{-1}	9.1903×10^{-1}
8	9.8220×10^{-1}	$\mathbf{9.9719 \times 10^{-1}}$	8.0786×10^{-1}	9.1221×10^{-1}	9.5558×10^{-1}
9	9.9033×10^{-1}	$\mathbf{9.9614 \times 10^{-1}}$	8.3555×10^{-1}	9.6237×10^{-1}	9.9609×10^{-1}

Table 15. Model comparison criteria: ML for the unperturbed F3II mammary carcinoma growing in female BALB/C/Cenp mice.

Individual	Norm	Prop	Thres	Stud	Stud prop
1	0.0000	$\mathbf{1.4504 \times 10^{-1}}$	1.2172×10^{-1}	0.0000	7.6000×10^{-2}
2	0.0000	$\mathbf{2.5184 \times 10^{-1}}$	1.6776×10^{-1}	2.4000×10^{-4}	1.4496×10^{-1}
3	0.0000	$\mathbf{4.0764 \times 10^{-1}}$	2.4096×10^{-1}	0.0000	2.2628×10^{-1}
4	0.0000	$\mathbf{7.9928 \times 10^{-1}}$	1.8980×10^{-1}	0.0000	4.6372×10^{-1}
5	1.6000×10^{-4}	$\mathbf{2.5396 \times 10^{-1}}$	1.0672×10^{-1}	4.0000×10^{-5}	1.4080×10^{-1}
6	2.4000×10^{-4}	$\mathbf{4.1912 \times 10^{-1}}$	2.7104×10^{-1}	8.0000×10^{-5}	3.6080×10^{-1}
7	2.0000×10^{-4}	$\mathbf{3.3660 \times 10^{-1}}$	1.3504×10^{-1}	2.4000×10^{-4}	2.2240×10^{-1}
8	8.3600×10^{-3}	$\mathbf{3.1356 \times 10^{-1}}$	1.6284×10^{-1}	1.2720×10^{-2}	2.2648×10^{-1}
9	0.0000	$\mathbf{7.6516 \times 10^{-1}}$	1.6940×10^{-1}	0.0000	5.2548×10^{-1}

Table 16. Model comparison criteria: fr(KS) for the unperturbed F3II mammary carcinoma growing in female BALB/C/Cenp mice.

Model	$\overline{BIC}$	$\overline{DIC}$	$\overline{ML}$	$\overline{fr(KS)}$
Norm	9.7976×10^{1}	1.4196×10^{2}	$\mathbf{9.7207 \times 10^{-1}}$	2.0000×10^{-4}
Prop	1.0433×10^{2}	1.4442×10^{2}	9.2043×10^{-1}	2.7164×10^{-1}
Thres	9.6017×10^{1}	$\mathbf{1.3925 \times 10^{2}}$	8.0212×10^{-1}	$\mathbf{2.7512 \times 10^{-1}}$
Stud	9.7980×10^{1}	1.4328×10^{2}	9.6507×10^{-1}	2.4000×10^{-4}
StudProp	$\mathbf{9.2215 \times 10^{1}}$	1.4498×10^{2}	5.6668×10^{-1}	2.7290×10^{-1}

Table 17. Model comparison criteria: Mean BIC, mean DIC, mean ML, and mean fraction of non-rejected normality tests with KS for the unperturbed Ehrlich-tumor data.

Model	$\overline{BIC}$	$\overline{DIC}$	$\overline{ML}$	$\overline{fr(KS)}$
Norm	1.0502×10^{2}	1.3834×10^{2}	$\mathbf{9.6570 \times 10^{-1}}$	2.0000×10^{-4}
Prop	1.0468×10^{2}	1.4159×10^{2}	9.1871×10^{-1}	2.7875×10^{-1}
Thres	1.0635×10^{2}	$\mathbf{1.3704 \times 10^{2}}$	9.6499×10^{-1}	$\mathbf{2.8078 \times 10^{-1}}$
Stud	1.0502×10^{2}	1.3957×10^{2}	9.6637×10^{-1}	3.4800×10^{-4}
StudProp	$\mathbf{1.0214 \times 10^{2}}$	1.4243×10^{2}	9.1757×10^{-1}	2.7224×10^{-1}

Table 18. Model comparison criteria: Mean BIC, mean DIC, mean ML, and fraction of non-rejected normality tests with KS for the unperturbed fibrosarcoma Sa-37 tumor data.

Model	$\overline{BIC}$	$\overline{DIC}$	$\overline{ML}$	$\overline{fr(KS)}$
Norm	1.2538×10^{2}	1.7413×10^{2}	$\mathbf{9.8491 \times 10^{-1}}$	2.7273×10^{-4}
Prop	$\mathbf{1.1931 \times 10^{2}}$	1.5083×10^{2}	9.7011×10^{-1}	$\mathbf{3.8342 \times 10^{-1}}$
Thres	1.2306×10^{2}	$\mathbf{1.1226 \times 10^{2}}$	9.0571×10^{-1}	2.0242×10^{-1}
Stud	1.2542×10^{2}	1.2342×10^{2}	9.1925×10^{-1}	2.1455×10^{-4}
StudProp	1.2154×10^{2}	1.3346×10^{2}	9.6843×10^{-1}	2.2132×10^{-1}

Table 19. Model comparison criteria: Mean BIC, mean DIC, mean ML, and fraction of non-rejected normality tests with KS for the unperturbed F3II mammary carcinoma growing in male BALB/C/Cenp mice.

Model	$\overline{BIC}$	$\overline{DIC}$	$\overline{ML}$	$\overline{fr(KS)}$
Norm	1.2822×10^2	1.8437×10^2	9.8674×10^{-1}	9.9556×10^{-4}
Prop	$\mathbf{1.1145 \times 10^2}$	1.4010×10^2	$\mathbf{9.9317 \times 10^{-1}}$	$\mathbf{4.1024 \times 10^{-1}}$
Thres	1.2125×10^2	$\mathbf{1.1044 \times 10^2}$	8.5866×10^{-1}	1.7392×10^{-1}
Stud	1.2715×10^2	1.2230×10^2	9.6542×10^{-1}	1.4800×10^{-3}
StudProp	1.1425×10^2	1.2091×10^2	9.5630×10^{-1}	2.6521×10^{-1}

Table 20. Model comparison criteria: Mean BIC, mean DIC, mean ML, and mean fraction of non-rejected normality tests with KS for the unperturbed F3II mammary carcinoma growing in male BALB/C/Cenp mice.

References

1. Brú, A., Albertos, S., Subiza, J. L., García-Asenjo, J. L. & Brú, I. The universal dynamics of tumor growth. *Biophysical journal* **85**, 2948–2961 (2003).
2. Steel, G. *Growth kinetics of tumours. Cell Population Kinetics in Relation to the Growth and Treatment of Cancer* (Clarendon press, Oxford, 1977).
3. Fu, K. K. & Steel, G. Growth kinetics of a rat mammary tumour transplanted into immune-suppressed mice. *Cell Proliferation* **12**, 493–499 (1979).
4. Vaidya, V. G. & Alexandro, J. F. Evaluation of some mathematical models for tumor growth. *Int J Biomed Comput* **13**, 19–35 (1982).
5. Marušić, M. Mathematical models of tumor growth. *Mathematical Communications* **1**, 175–88 (1996).
6. Crispen, P. L. et al. Natural history, growth kinetics, and outcomes of untreated clinically localized renal tumors under active surveillance. *Cancer* **115**, 2844–2852 (2009).
7. Benzekry, S., Beheshti, A., Hahnfeldt, P. & Hlatky, L. Capturing the driving role of tumor-host crosstalk in a dynamical model of tumor growth. *Bio-protocol* **5** (2015).
8. Cabrales, L. et al. Mathematical modeling of tumor growth in mice following low-level direct electric current. *Math. Simul. Comp.* **78**, 112–120 (2008).
9. Kühleitner, M., Brunner, N., Nowak, W. G., Renner-Martin, K. & Scheicher, K. Best fitting tumor growth models of the von Bertalanffy-PütterType. *BMC cancer* **19**, 1–11 (2019).
10. González, M. M. et al. Is cancer a pure growth curve or does it follow a kinetics of dynamical structural transformation?. *BMC cancer* **17**, 1–14 (2017).
11. Castañeda, A. R. S. et al. New formulation of the Gompertz equation to describe the kinetics of untreated tumors. *PloS one* **14**, e0224978 (2019).
12. Ribatti, D. & Pezzella, F. Overview on the different patterns of tumor vascularization. *Cells* **10**, 639 (2021).
13. Sachs, R. K., Hlatky, L. R. & Hahnfeldt, P. Simple ode models of tumor growth and anti-angiogenic or radiation treatment. *Mathematical and Computer Modelling* **33**, 1297–1305 (2001).
14. Marušić, M., Vuk-Pavlovic, S. & Freyer, J. P. Tumor growth in vivo and as multicellular spheroids compared by mathematical models. *Bulletin of mathematical biology* **56**, 617–631 (1994).
15. Valentim, J. C. A., Oliveira, N. A., Rabi, J. A. & David, S. A. Can fractional calculus help improve tumor growth models?. *Journal of Computational and Applied Mathematics* **379**, 112964 (2020).
16. Frunzo, L., Garra, R., Giusti, A. & Luongo, V. Modeling biological systems with an improved fractional Gompertz law. *Commun Nonlinear Sci Numer Simulat* **74**, 260–267 (2019).
17. Sabir, Z. et al. Numerical investigations of the fractional-order mathematical model underlying immune-chemotherapeutic treatment for breast cancer using the neural networks. *Fractal Fract.* **6**, 184 (2022).
18. Valentim, C. A., Rabi, J. A., A., D. S. & Machado, J. A. T. On multistep tumor growth models of fractional variable-order. *BioSystems* **199**, 104294 (2021).
19. Hassani, H., Tenreiro Machado, J. A., Avazzadeh, Z., Safari, E. & Mehrabi, S. Optimal solution of the fractional order breast cancer competition model. *Scientifc Reports* **11**, 15622 (2021).
20. Durbin, P. W. et al. Construction of a growth curve for mammary tumors of the rat. *Cancer Research* **27**, 1341–1347 (1967).
21. Simpson-Herren, L. & Lloyd, H. H. Kinetic parameters and growth curves for experimental tumor systems 1, 2. *Cancer Chemotherapy Reports* **54**, 143 (1970).
22. Steel, G. & Lamerton, L. The growth rate of human tumours. *British journal of cancer* **20**, 74 (1966).
23. Xu, X. The biological foundation of the Gompertz model. *International journal of bio-medical computing* **20**, 35–39 (1987).
24. Speer, J. F., Petrosky, V. E., Retsky, M. W. & Wardwell, R. H. A stochastic numerical model of breast cancer growth that simulates clinical data. *Cancer Research* **44**, 4124–4130 (1984).
25. Vaghi, C. et al. Population modeling of tumor growth curves and the reduced gompertz model improve prediction of the age of experimental tumors. *PLoS computational biology* **16**, e1007178 (2020).
26. Oberg, A. L. et al. Statistical analysis of comparative tumor growth repeated measures experiments in the ovarian cancer patient derived xenograft (PDX) setting. *Scientific reports* **11**, 8076 (2021).
27. Voulgarelis, D., Bulusu, K. C. & Yates, J. W. Comparison of classical tumour growth models for patient derived and cell-line derived xenografts using the nonlinear mixed-effects framework. *Journal of Biological Dynamics* **16**, 160–185 (2022).
28. Jarrett, A. et al. Mathematical models of tumor cell proliferation: a review of the literature. *Expert Review of Anticancer Therapy* **18**, 1271–1286 (2018).
29. Ghaffari, L. N. et al. Classical mathematical models for prediction of response to chemotherapy and immunotherapy. *PLoS computational biology* **18**, e1009822 (2022).
30. Sarapata, E. & de Pillis, L. G. A comparison and catalog of intrinsic tumor growth models. *Bull Math Biol* **76**, 2010–2024 (2014).
31. Cabrales, L. et al. Modified Gompertz equation for electrotherapy murine tumor growth kinetics: Predictions and new hypotheses. *BMC Cancer* **10**, 589–599 (2010).
32. Goris, N. et al. Correspondence between formulations of Avrami and Gompertz for untreated tumor growth kinetics. *Rev Mex FĀs* **66**, 632–636 (2020).
33. Ciria, H. et al. Antitumor effectiveness of different amounts of electrical charge in Ehrlich and fibrosarcoma Sa-37 tumors. *BMC cancer* **4**, 1–10 (2004).
34. González, M. M. et al. Dose-response study for the highly aggressive and metastatic primary F3II mammary carcinoma under direct current. *Bioelectromagnetics* **39**, 460–475 (2018).

35. Goris, N. et al. Efficacy of direct current generated by multiple-electrode arrays on F3II mammary carcinoma: experiment and mathematical modeling. *Journal Translational Medicine* **18**, 1–17 (2020).
36. Molski, M. & Konarski, J. Tumor growth in the space-time with temporal fractal dimension. *Chaos, Solitons & Fractals* **36**, 811–818 (2008).
37. Waliszewski, P. Complexity, diversity, homogeneity and topology measures re-define biological tumor aggressiveness as a function of intercellular connectivity. *J Biomed Res Environ Sci.* **15**, 1290–1307 (2023).
38. Waliszewski, P. *Complementarity, Complexity and the Fokker–Planck Equation; from the Microscale Quantum Stochastic Events to Fractal Dynamics of Cancer*, 19–54 (Springer International Publishing, 2022).
39. Szasz, A. Time-fractal modulation-possible modulation effects in human therapy. *Open journal Biophysics* **12**, 38–87 (2021).
40. Megyesfalvi, Z. *et al.* Clinical insights into small cell lung cancer: Tumor heterogeneity, diagnosis, therapy, and future directions. *CA: a cancer journal for clinicians* **73**, 620–652 (2023).
41. Helisaz, H., Belanger, E., Black, P., Bacca, M. & Chiao, M. Quantifying the impact of cancer on the viscoelastic properties of the prostate gland using a quasi-linear viscoelastic model. *Acta Biomaterialia* **173**, 184–198 (2024).
42. Castañeda, A. et al. Spatio temporal dynamics of direct current in treated anisotropic tumors. *Math Compt Simul* **203**, 609–632 (2023).
43. Zimmermann, M., Box, C. & Eccles, S. A. Two-dimensional vs. three-dimensional in vitro tumor migration and invasion assays. In *Target Identification and Validation in Drug Discovery*, 227–252 (Springer, 2013).
44. Ayers, G. D. et al. Volume of preclinical xenograft tumors is more accurately assessed by ultrasound imaging than manual caliper measurements. *Journal of Ultrasound in Medicine* **29**, 891–901 (2010).
45. Friberg, S. & Mattson, S. On the growth rates of human malignant tumors: implications for medical decision making. *Journal of surgical oncology* **65**, 284–297 (1997).
46. Davies, P. C., Demetrius, L. & Tuszynski, J. A. Cancer as a dynamical phase transition. *Theoretical Biology and Medical Modelling* **8**, 1–16 (2011).
47. Zhang, Q. & Austin, R. H. Physics of cancer: the impact of heterogeneity. *Annu. Rev. Condens. Matter Phys.* **3**, 363–382 (2012).
48. Murphy, H., Jaafari, H. & Dobrovolny, H. M. Differences in predictions of ODE models of tumor growth: a cautionary example. *BMC cancer* **16**, 1–10 (2016).
49. Patmanidis, S., Charalampidis, A. C., Kordonis, I., Mitsis, G. D. & Papavassilopoulos, G. P. Comparing methods for parameter estimation of the Gompertz tumor growth model. *IFAC-PapersOnLine* **50**, 12203–12209 (2017).
50. Geng, C., Paganetti, H. & Grassberger, C. Prediction of treatment response for combined chemo-and radiation therapy for non-small cell lung cancer patients using a bio-mathematical model. *Scientific reports* **7**, 1–12 (2017).
51. Benzekry, S. et al. Classical mathematical models for description and prediction of experimental tumor growth. *PLoS computational biology* **10**, e1003800 (2014).
52. Koziol, J. A., Falls, T. J. & Schnitzer, J. E. Different ODE models of tumor growth can deliver similar results. *BMC cancer* **20**, 1–10 (2020).
53. Moré, J. J. The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical analysis*, 105–116 (Springer, 1978).
54. Lange, K. L., Little, R. J. & Taylor, J. M. Robust statistical modeling using the t distribution. *Journal of the American Statistical Association* **84**, 881–896 (1989).
55. Dehning, J. *et al.* Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* **369** (2020).
56. Tjørve, K. M. & Tjørve, E. The use of Gompertz models in growth analyses, and new Gompertz-model approach: An addition to the unified-Richards family. *PloS one* **12**, e0178691 (2017).
57. Agur, Z., Elishmereni, M. & Kheifetz, Y. Personalizing oncology treatments by predicting drug efficacy, side-effects, and improved therapy: mathematics, statistics, and their integration. *WIREs Syst Biol Med* **6**, 239–253 (2014).
58. Davidian, M., Giltinan, David M. & D. M.,. Nonlinear models for repeated measurement data: An overview and update. *Journal of Agricultural, Biological, and Environmental Statistics* **8**, 387–419 (2003).
59. Motulsky, H. & Christopoulos, A. *Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting* (Oxford University Press, 2004).
60. Neath, A. A. & Cavanaugh, J. E. The Bayesian information criterion: background, derivation, and applications. *Wiley Interdisciplinary Reviews: Computational Statistics* **4**, 199–203 (2012).
61. Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & Van der Linde, A. The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **76**, 485–493 (2014).
62. Vrieze, S. I. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological methods* **17**, 228 (2012).
63. Pooley, C. & Marion, G. Bayesian model evidence as a practical alternative to deviance information criterion. *Royal Society open science* **5**, 171519 (2018).
64. De Santis, F. & Spezzaferri, F. Alternative Bayes factors for model selection. *Canadian Journal of Statistics* **25**, 503–515 (1997).
65. Conigliani, C. & O'hagan, A. Sensitivity of the fractional Bayes factor to prior distributions. *Canadian Journal of Statistics* **28**, 343–352 (2000).
66. O'Hagan, A. Fractional Bayes factors for model comparison. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**, 99–118 (1995).
67. Lartillot, N. & Philippe, H. Computing Bayes factors using thermodynamic integration. *Systematic biology* **55**, 195–207 (2006).
68. Haario, H., Laine, M., Mira, A. & Saksman, E. Dram: efficient adaptive MCMC. *Statistics and computing* **16**, 339–354 (2006).
69. Sahin, K. H. & Ciric, A. R. A dual temperature simulated annealing approach for solving bilevel programming problems. *Computers & chemical engineering* **23**, 11–25 (1998).
70. Raftery, A. E. Bayes factors and bic: Comment on "a critique of the bayesian information criterion for model selection". *Sociological Methods & Research* **27**, 411–427 (1999).
71. Berger, J. O. & Mortera, J. Default Bayes factors for nonnested hypothesis testing. *Journal of the american statistical association* **94**, 542–554 (1999).
72. Berger, V. W. & Zhou, Y. *Kolmogorov-Smirnov test: Overview* (Statistics reference online, Wiley statsref, 2014).
73. Van Rossum, G. & Drake, F. L. *Python 3 Reference Manual* (CreateSpace, Scotts Valley, CA, 2009).
74. Fundamental Algorithms for Scientific Computing in Python. Virtanen, P. et al. SciPy 1.0. *Nature Methods* **17**, 261–272 (2020).
75. Miles, P. R. pymcmcstat: A python package for bayesian inference using delayed rejection adaptive metropolis. *Journal of Open Source Software* **4**, 1417 (2019).
76. Cabrales, L. et al. Electrochemical treatment of mouse Ehrlich tumor with direct electric current. *Bioelectromagnetics* **22**, 316–322 (2001).
77. Jarosz, A. F. & Wiley, J. What are the odds? a practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving* **7**, 2 (2014).
78. Skipper, H. E. Kinetics of mammary tumor cell growth and implications for therapy. *Cancer* **28**, 1479–1499 (1971).
79. Grassberger, C. et al. Patient-specific tumor growth trajectories determine persistent and resistant cancer cell populations during treatment with targeted therapies. *Cancer research* **79**, 3776–3788 (2019).
80. Friberg, S. On the growth rates of human malignant tumors: implications for medical decision making. *Nowotwory. Journal of Oncology* **55**, 1–1 (2005).

81. Uzosike, A. C. et al. Growth kinetics of small renal masses on active surveillance: variability and results from the DISSRM registry. *The Journal of urology* **199**, 641–648 (2018).
82. McIntosh, A. G. et al. Active surveillance for localized renal masses: tumor growth, delayed intervention rates, and> 5-yr clinical outcomes. *European urology* **74**, 157–164 (2018).
83. Gerlee, P. The model muddle: in search of tumor growth laws. *Cancer research* **73**, 2407–2411 (2013).
84. Zhang, C., Jiang, X., Guo, S. & Li, Q. Mathematical modeling of tumor growth in mouse models. *Cancer Research* **79**, 4613–4613 (2019).
85. Held, L. & Sabanés Bové, D. *Likelihood and Bayesian Inference: With Applications in Biology and Medicine.* Statistics for Biology and Health (Springer, Berlin, Heidelberg, 2020), edition2nd edn.

Acknowledgements

Author contributions

The authors E.R., L.B. and J.M. contributed equally to this work, doing the research and writing the manuscript. L.R. , S.S. and A.R.S.C. reviewed the manuscript

Declarations

Competing interests

The authors declare no competing interests

Additional information

Correspondence and requests for materials should be addressed to L.E.B.C. or J.I.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.