A Framework for Performing Data-Driven Modeling of Tumor Growth with Radiotherapy Treatment

Heyrim Cho, Allison L. Lewis, Kathleen M. Storey, Rachel Jennings, Blerta Shtylla, Angela M. Reynolds, and Helen M. Byrne

Abstract Recent technological advances make it possible to collect detailed information about tumors, and yet clinical assessments about treatment responses are typically based on sparse datasets. In this work, we propose a workflow for choosing an appropriate model, verifying parameter identifiability, and assessing the amount of data necessary to accurately calibrate model parameters. As a proof-of-concept, we compare tumor growth models of varying complexity in an effort to determine

Authors Heyrim Cho, Allison L. Lewis, and Kathleen M. Storey have equally contributed to this chapter.

H. Cho
University of California Riverside, Riverside, CA, USA
e-mail: heyrimc@ucr.edu

A. L. Lewis
Lafayette College, Easton, PA, USA
e-mail: lewisall@lafayette.edu

K. M. Storey
University of Michigan, Ann Arbor, MI, USA
e-mail: storeyk@umich.edu

R. Jennings
United Health Group, Minnetonka, MN, USA
e-mail: rljennings@ara.com

B. Shtylla
Department of Mathematics, Pomona College, Claremont, CA, USA
e-mail: blerta.shtylla@pomona.edu

A. M. Reynolds (✉)
Virginia Commonwealth University, Richmond, VA, USA
e-mail: areynolds2@vcu.edu

H. M. Byrne
University of Oxford, Oxford, UK
e-mail: helen.byrne@maths.ox.ac.uk

© The Association for Women in Mathematics and the Author(s) 2021

R. Segal et al. (eds.), *Using Mathematics to Understand Biological Complexity*,
Association for Women in Mathematics Series 22,
https://doi.org/10.1007/978-3-030-57129-0_8

the level of model complexity needed to accurately predict tumor growth dynamics and response to radiotherapy. We consider a simple, one-compartment ordinary differential equation model which tracks tumor volume and a two-compartment model that accounts for tumor volume and the fraction of necrotic cells contained within the tumor. We investigate the structural and practical identifiability of these models, and the impact of noise on identifiability. We also generate synthetic data from a more complex, spatially-resolved, cellular automaton model (CA) that simulates tumor growth and response to radiotherapy. We investigate the fit of the ODE models to tumor volume data generated by the CA in various parameter regimes, and we use sequential model calibration to determine how many data points are required to accurately infer model parameters. Our results suggest that if data on tumor volumes alone is provided, then a tumor with a large necrotic volume is the most challenging case to fit. However, supplementing data on total tumor volume with additional information on the necrotic volume enables the two compartment ODE model to perform significantly better than the one compartment model in terms of parameter convergence and predictive power.

Keywords Systems biology · Mathematical oncology · Parameter identifiability · Bayesian sequential calibration · Model selection

1 Introduction

Cancer remains one of the leading causes of death in the world, second only to cardiac disease. As such, it represents a significant global public health and socio-economic problem. Of particular interest, given the unpleasant side-effects that accompany many cancer treatments, is being able to establish as early as possible whether a patient will respond (or is responding) to a particular treatment and, based on this assessment, whether treatment should be continued or a new treatment started. Mathematical modeling provides a natural framework within which to answer such questions. In more detail, mechanistic models that describe the growth dynamics of a tumor and its response to treatment may be fit to patient data collected during treatment and used to predict how the tumor's size (and possibly composition) will change if treatment is continued. Model fits, parameter estimates, and predictions may be revised as treatment progresses and more patient data become available. These predictions can then be used to inform decisions about whether to continue with the current treatment.

The approach outlined above relies upon the availability of time-dependent mathematical models of tumor growth and patient data to which the models can be fit. Advances in technology mean that it is now possible to collect detailed information about tumors (e.g., their size, spatial composition, mutational status, vascularity and degree of immune infiltration). Even so, decisions about treatment options (e.g., surgery, radiotherapy, chemotherapy and immunotherapy), and assessments about treatment responses are often based on statistical analyses of sparse and noisy data relating to a small number of quantities of interest (e.g., tumor volumes at three time-points: at diagnosis, at treatment start, and at treatment end).

The availability of detailed information about tumors has undoubtedly stimulated the development of a large number of mathematical and computational models of tumor growth. These range from spatially-averaged, phenomenological models formulated using differential equations (e.g., logistic growth, Gompertzian growth) [8, 10], to multiphase models based on mixture theory [4, 20] or phase field theory [22], and multiscale models that couple subcellular, cellular, and tissue scale phenomena [14]. In addition to simulating tumor growth, these models have also been used to study tumor responses to treatments including radiotherapy, chemotherapy, immunotherapy and combinations thereof [3, 19]. Unfortunately, the absence of suitable experimental data coupled with the complexity of many of the theoretical models means that few of them have been validated and/or parameterized. Additionally, the sheer number and variety of available models makes it difficult to determine which model may be most appropriate for a given scenario and available data [9, 21, 24].

In this investigation, we propose a workflow for determining an appropriate model to be used under the constraints of a given scenario (i.e. when data is scarce or noisy). We begin with an identifiability analysis in two parts: structural identifiability, which determines whether the model parameters can be uniquely estimated in an "ideal world", and practical identifiability, which reassesses parameter identifiability in a real-world scenario, in which data is noisy and potentially sparse.

After establishing structural and practical identifiability, we tackle the question of which model to fit to the available data. We perform a parameter sweep to measure error in model predictions across different parameter regimes, establishing conditions under which certain models can be used to make accurate predictions about tumor growth. While we desire model simplicity whenever possible, we recognize that in some scenarios (for example, when the tumor comprises a large portion of necrotic tissue), a model that tracks only tumor volume (for instance) may not be able to describe the tumor growth dynamics.

Once we have chosen an appropriate model for a given scenario and verified that its parameters are structurally and practically identifiable, we examine how the model calibration is affected by the availability of data. In this investigation, we perform sequential calibration, adding one data point at a time to determine how much data is necessary to uniquely calibrate the model parameters. At each step, we use the current parameter values to predict the future tumor volume after treatment has concluded, and compare this prediction to our known "truth" to assess whether the current amount of data is adequate for making future predictions. This investigation helps us to determine the extent to which additional data will increase the predictive power of our mathematical models.

The approach outlined above serves as a "proof-of-concept" for our proposed framework; we test our procedure on two simple compartmental models for tumor growth: one that describes only total tumor volume, and a second that also accounts for the proportion of necrotic cells. We compare their dynamics to data generated from a more complex cellular automaton model (CA), which we use as our "truth". Additionally, we combine our growth models with the linear-quadratic model [12] to simulate tumour responses to a radiotherapy treatment regimen and also to test our model predictions in the presence of an intervention to tumor growth.

The remainder of this paper is structured as follows. In Sect. 2 we introduce the one- and two-compartment models that we use to predict tumor growth and response to radiotherapy. We also outline the structure of the cellular automaton model that we use to generate synthetic data for fitting the compartment models. We present typical CA simulation results that illustrate how a tumor's spatial composition and growth dynamics may change in response to radiotherapy. In Sect. 3, we explain how structural and practical identifiability analysis methods can be used to determine whether it is possible to infer the parameters associated with a particular model when the data is perfect (structural identifiability) and noisy (practical identifiability). We include several case studies that investigate how the addition of noise to the CA data affects the ability to recover model parameters, and compare the ability of the one- and two-compartment models to fit data generated by the CA under a variety of conditions, including a range of necrotic heterogeneity levels. In Sect. 4, we study the goodness of fit of the one-compartment model to synthetic data generated from the two-compartment model. Through an extensive search of parameter space, we determine how parameters in the one- and two-compartment models are related, and discuss how these results can be used to select an appropriate model that will yield accurate predictions while still maintaining model simplicity whenever possible. Finally, in Sect. 5 we perform sequential model calibration to determine how much data is needed to accurately infer model parameters. The paper concludes in Sect. 6 with a summary of our key results and directions for future work.

2 The Mathematical Models

Here we introduce the three mathematical models that will be used throughout this investigation. The first is a one-compartment ODE model that tracks tumor volume over time. It is the most basic model that we use to describe tumor growth. The second model is a two-compartment ODE model that incorporates a state variable for tracking the portion of tumor volume that is composed of necrotic tissue, thereby allowing for tumor heterogeneity. This model will be used both as a data generator to test the capabilities of the one-compartment ODE model as well as a model to be calibrated against "true" data. Our final and most complex model is an cellular automaton model (CA) that we use to generate our "truth" data, as it is assumed to most accurately reflect reality by including the cell cycle, quiescent cells, and oxygen levels [14]. In all three cases, we will incorporate treatment via radiation using the linear-quadratic model for radiotherapy.

2.1 The One-Compartment Model

Our one-compartment model describes the time evolution of the tumor volume, $V(t)$, using a logistic growth model with growth rate λ and carrying capacity K:

$$\frac{dV}{dt} = \underbrace{\lambda V\left(1 - \frac{V}{K}\right)}_{\text{logistic growth}} - \underbrace{\eta V.}_{\text{natural cell death}} \tag{1}$$

Natural cell death, at rate η, is incorporated via the term $-\eta V$. In what follows, it will be convenient to re-parameterize Equation (1) to obtain the simpler form

$$\frac{dV}{dt} = AV\left(1 - \frac{B}{A}V\right), \tag{2}$$

where $A = \lambda - \eta$ and $B = \frac{\lambda}{K}$. From here on, we will refer to the one-compartment model in its re-scaled form, in Equation (2).

This simple model views the tumor as a homogeneous mass in which all cells are identical. In practice, however, as the tumor grows, regions at a distance from oxygen and nutrient sources (e.g., blood vessels for tumors growing in vivo) may undergo necrotic cell death in response to sustained oxygen and/or nutrient deprivation. In the one-compartment model, such dead or necrotic cells are assumed to be removed from the tumor instantaneously.

2.2 The Two-Compartment Model

In order to account for some aspects of tumor heterogeneity, we study a two-compartment model that tracks the time evolution of the volume of viable tumor cells ($V(t)$) and the volume of the necrotic core ($N(t)$), and that was originally developed in [18]. The population of proliferating (i.e., viable) cells is assumed to grow logistically with growth rate λ and carrying capacity K. Additionally, we assume that viable cells convert to necrotic cells at a constant rate η, and that necrotic material undergoes natural decay at a constant rate ζ. Combining these processes, we arrive at the following ODE system for $V(t)$ and $N(t)$:

$$\frac{dV}{dt} = \lambda V\left(1 - \frac{V}{K}\right) - \eta V, \tag{3a}$$

$$\frac{dN}{dt} = \eta V - \zeta N. \tag{3b}$$

To facilitate comparison with the one-compartment model, we reformulate (3a)–(3b) in terms of $Y(t)$, the total tumor volume ($Y = V + N$), and $\Phi(t)$, the fraction of the total volume that comprises necrotic cells ($\Phi = N/Y$). Using this notation, Equations (3a) and (3b) can be rewritten in the form

$$\frac{dY}{dt} = \lambda(1 - \Phi)Y\left(1 - (1 - \Phi)\frac{Y}{K}\right) - \zeta\Phi Y, \tag{4a}$$

$$\frac{d\Phi}{dt} = (1-\Phi)\left[\eta - \lambda\Phi\left(1-(1-\Phi)\frac{Y}{K}\right) - \zeta\Phi\right]. \tag{4b}$$

We note that in the limit as $\zeta \to \infty$ (i.e., if the timescale for degradation of necrotic material is much shorter than the timescale for other processes included in the model), Equations (4a) and (4b) reduce to the one-compartment model defined by Equation (1).

2.3 The Cellular Automaton Model

We use a spatially-explicit, hybrid cellular automaton model (CA) to generate synthetic tumor volume data. Our cellular automaton model is adapted from the one developed in [14]. In the CA model, cells are arranged on a 200×200 grid, which represents a two-dimensional cross-section of size $0.36 \times 0.36\,\text{cm}^2$ through a three-dimensional spheroid in vitro. Each automaton can be occupied either by a tumor cell or culture medium. The CA couples the dynamics of automaton elements arranged on the grid to the oxygen concentration. We identify with each automaton $\mathbf{x} = (x, y)$ at time t a dynamical variable with a state and a neighborhood. The four possible states are proliferating $\mathscr{P}$, quiescent $\mathscr{Q}$, necrotic $\mathscr{N}$, and empty $\mathscr{E}$, determined by the oxygen level at each site: if $c > c_Q$ then the cells proliferate, if $c_N < c < c_Q$ then the cells stop proliferating and halve the rate at which they consume oxygen, and if $c \leq c_N$ then the cells become necrotic (see Table 1 for details). Each cell communicates with cells within its Moore neighborhood, i.e., its eight nearest neighbors. Proliferating cells are assigned counters that describe

Table 1 A summary of the parameters used in the CA and their default values. Parameter values are estimated using experimental data from the prostate cancer cell line, PC3, in [14]

Parameter	Description	Value	Units
l	Cell size	0.0018	cm
L	Domain length	0.36	cm
$\bar{\tau}_{cycle}(\sigma_{cycle})$	Mean (standard deviation) cell cycle time	18.3 (1.4)	h
c_∞	Background O_2 concentration	2.8×10^{-7}	mol cm^{-3}
D	O_2 diffusion constant	1.8×10^{-5}	$cm^2 s^{-1}$
c_Q	O_2 concentration threshold for proliferating cells	1.82×10^{-7}	mol cm^{-3}
c_N	O_2 concentration threshold for quiescent cells	1.68×10^{-7}	mol cm^{-3}
κ_P	O_2 consumption rate of proliferating cells	1.0×10^{-8}	mol $cm^{-3} s^{-1}$
κ_Q	O_2 consumption rate of quiescent cells	5.0×10^{-9}	mol $cm^{-3} s^{-1}$
p_{NR}	Rate of lysis of necrotic cells	Range: 0.004–0.016	hr^{-1}

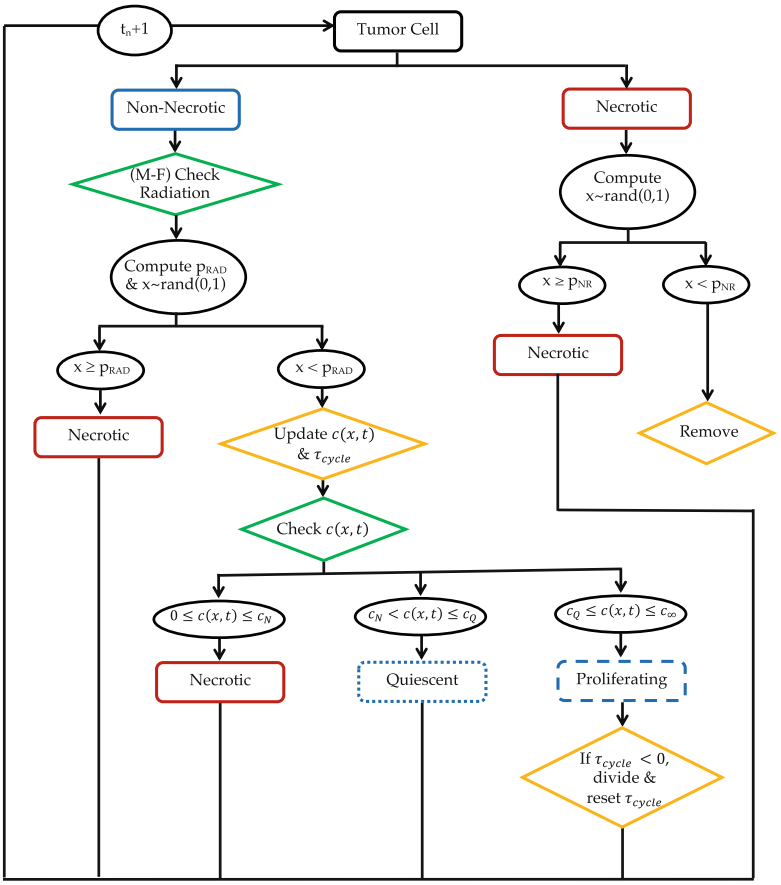

Fig. 1 Flow diagram describing the cell movement between necrotic, quiescent and proliferating states. The parameter p_{RAD} denotes the probability that a cell becomes necrotic following radiotherapy and $p_{RAD} = 1 - e^{-\alpha d - \beta d^2}$. The parameter τ_{cycle} denotes the specific cell's assigned cycle length

where they are in the cell cycle. These counters are initially drawn from a normal distribution with mean $\bar{\tau}_{cycle}$ and standard deviation σ_{cycle}. After each time step, the cell cycle counter of each cell decreases by an amount that depends on the number of neighboring cells; a smaller reduction in cell cycle time occurs with a larger number of neighbors, to model the regulatory process known as contact inhibition of proliferation. Figure 1 summarizes how a cell can transition between quiescent, proliferating and necrotic states.

We model the single growth-rate-limiting nutrient, oxygen, explicitly via a reaction-diffusion equation. In particular, the evolution of the oxygen concentration $c(\mathbf{x}, t)$ (mol cm^{-3}) at location $\mathbf{x}$ for time t is described by:

$$\frac{\partial c(\mathbf{x}, t)}{\partial t} = D\nabla^2 c(\mathbf{x}, t) - \Gamma(\mathbf{x}, t), \tag{5}$$

where D is the oxygen diffusion coefficient (cm^2 s^{-1}). and $\Gamma(\mathbf{x}, t)$ is the oxygen consumption rate (mol cm^{-3} s^{-1}), defined as follows:

$$\Gamma(\mathbf{x}, t) = \begin{cases} \kappa_P & \text{if } \mathbf{x} \text{ is occupied by a proliferating cell} \\ \kappa_Q & \text{if } \mathbf{x} \text{ is occupied by a quiescent cell} \\ 0 & \text{otherwise.} \end{cases}$$

The parameters κ_P and κ_Q are the rates at which quiescent and proliferating cells consume oxygen, respectively, and are in mol cm^{-3} s^{-1} with $\kappa_P >= \kappa_Q$. We also use the following initial and boundary conditions to describe the situation in which oxygen diffuses from the boundaries of a square Petri dish into the culture medium:

$$c(x, y, 0) = c_\infty,$$
$$c(0, y, t) = c(L, y, t) = c(x, 0, t) = c(x, L, t) = c_\infty,$$

where L is the domain length, and c_∞ is the background O_2 concentration. See Table 1 for a list of the parameter values used to simulate the CA.

When a cell cycle counter reaches 0, the cell divides to produce two identical cells, one located at the same site as the parent, and one placed in an empty neighboring site, if available. If more than one neighboring site is empty, the site with the maximum number of neighbors is chosen to maintain cell-cell adhesion. If no adjacent sites are empty, then in an effort to simulate the mechanical stress exerted on neighboring cells during spheroid expansion, we find the shortest chain of cells connecting the dividing cell to the spheroid's boundary and shift this chain outward to create space for the daughter cell. Figure 2a, b displays this process.

A cell becomes necrotic if the oxygen concentration at its location falls below a specified threshold, or if the cell is irradiated, as discussed in Sect. 2.4. Necrotic cells are lysed at rate p_{NR}. Lysis involves removing the necrotic cell and then shifting inward a chain of cells starting from the boundary of the spheroid to fill in the removed cell's site. The spheroidal shape of the tumor is preserved by choosing the boundary cell that is farthest from the spheroid center, and then shifting a chain of adjacent cells inward. Figure 2c, d displays this process.

All model simulations are initialized by placing a circular cluster of cells in the center of the grid: this imitates seeding a spheroid in a Petri dish. The cells consume oxygen as it diffuses from the culture medium and this enables them to progress through the cell cycle and to divide. As the spheroid grows, oxygen levels at its center fall. When the oxygen concentration drops below a threshold value cells exit the cell cycle and become quiescent. As the spheroid grows further, and oxygen levels decrease further, quiescent cells die via necrosis and the resulting necrotic debris is transported away from the spheroid.

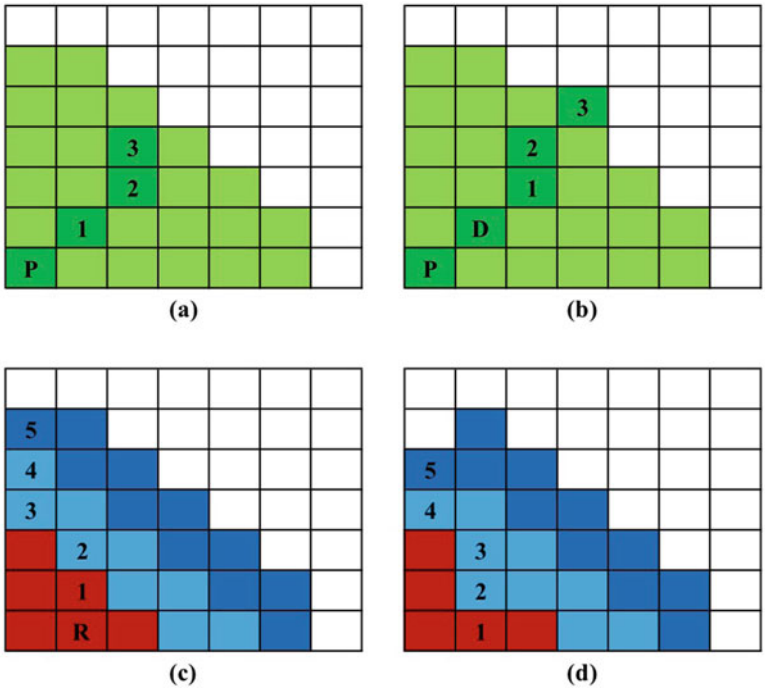

Fig. 2 Figures adapted from [14]. The spheroidal shape of the tumor is preserved by shifting a chain of cells outward when cell division occurs, shown (**a**) before and (**b**) after division, and by pulling a chain of cells inward when a cell lyses, shown (**c**) before and (**d**) after lysis. In (**a**)–(**b**), P denotes the dividing parent cell, and D denotes the daughter cell. The dividing cell pushes the chain of cells in (**a**), labeled by 1,2,3, outward to occupy the sites shown in (**b**). In (**c**)–(**d**), R denotes the cell that is removed. The numbered chain of cells, 1–5, are shifted in order to take the place of cell R

We use the CA to generate a series of synthetic spheroids which differ in their growth rates, sizes and spatial composition. Parameters are set to baseline values determined using experimental data from the prostate cancer cell line, PC3, in [14]. The results presented in Fig. 3 show how the size and composition of a spheroid change over 60 days of growth as we vary p_{NR}, the rate at which necrotic cells are removed. We use $p_{NR} = 0.015\,\mathrm{hr}^{-1}$ and $p_{NR} = 0.004\,\mathrm{hr}^{-1}$ to generate control (untreated) spheroids with approximately 20% and 50% necrotic volume, respectively, at steady state.

2.4 Radiotherapy Treatment

We now explain how we incorporate treatment with radiotherapy (RT) in all three models. We consider a typical tumor treatment regimen in which daily doses of 2 Gy

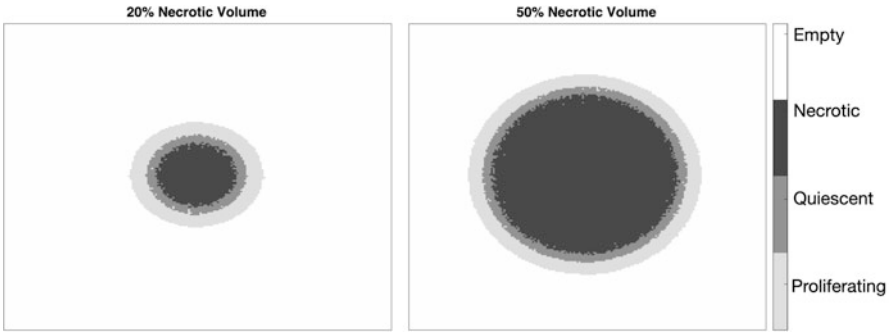

Fig. 3 CA simulations after 60 days in the absence of treatment, representing two-dimensional cross-sections of tumor spheroids. On the left, we used $p_{NR} = 0.015$ to generate a spheroid in which the necrotic cells occupy approximately 20% of total tumor volume. On the right, we used $p_{NR} = 0.004$ to generate a spheroid in which necrotic cells occupy approximately 50% of total tumor volume

are administered Monday through Friday for 6 weeks. We use the linear-quadratic model [12] to account for the effect of RT. This model assumes that the fraction of cells that survive exposure to a dose d of RT is given by

$$\text{Survival fraction,} \quad SF = e^{-\alpha d - \beta d^2}, \tag{6}$$

where α and β are tissue specific radiosensitivity parameters that model single and double strand breaks of the DNA [17]. We assume that the effect of RT is instantaneous, with the non-surviving cell fraction immediately removed when therapy is administered. Under these assumptions, the one-compartment model becomes

$$\begin{aligned} \frac{dV}{dt} &= AV\left(1 - \frac{B}{A}V\right) \quad \text{for } t_i^+ < t < t_{i+1}^-, \\ V(t_i^+) &= \exp(-\alpha d - \beta d^2)\, V(t_i^-). \end{aligned} \tag{7}$$

where t_i (for $i = 1, 2, \ldots, N$) denote the times at which radiotherapy is delivered, and $V(t_i^{\pm})$ denote the tumour volume just before and after radiotherapy is administered.

Treatment in the two-compartment model is modeled analogously, except that the sink of irradiated cells from the viable tumor volume will have an equal and opposite source term in the ODE for the necrotic component. Similarly in the cellular automaton model, each living cell becomes necrotic with probability $1 - e^{-\alpha d - \beta d^2}$ when radiotherapy is administered.

2.5 Typical CA Simulation Results

We generate synthetic spheroids using the CA described in Sect. 2.3, using the default parameter values listed in Table 1. Snapshots of typical CA tumor spheroids in the absence of treatment at days 6, 17, 30 and 70 are presented in Fig. 4a. Similar results showing the response to radiotherapy of tumors with low radiosensitivity are presented in Fig. 4b and for tumors with high radiosensitivity in Fig. 4c. Tumors with faster necrotic decay ($p_{NR} = 0.015$) are shown in the four-panels on the left, and tumors with slower necrotic decay ($p_{NR} = 0.004$) are shown on the right. We note that when radiotherapy is applied, treatment begins on Day 15. We observe that, in most cases, the size of the tumor on Day 70 is larger when the rate of necrotic decay is low than when it is high. However, if the tumor has high radiosensitivity, then the situation reverses: all living cells are eliminated when the necrotic cells decay slowly but not when they decay rapidly.

For each set of parameter values, we simulate 100 realizations of the CA and determine how the mean total tumor volume and mean necrotic volume change over time. Figure 5a presents the averaged results when no treatment is applied, with results corresponding to the high rate of necrotic decay ($p_{NR} = 0.015$) on the left, and results corresponding to the low rate ($p_{NR} = 0.004$) on the right. Figures 5b, c summarize the results for both cases when radiotherapy is applied. In Fig. 5b, the tumor has a lower radiosensitivity level ($\alpha/\beta = 9$), while Fig. 5c shows the results for tumors with higher radiosensitivity ($\alpha/\beta = 1$). The plots of mean tumor volume and mean necrotic volume confirm that the trends we observe in a single realization (see Fig. 4a–c) are representative of the average behavior in each case. We use the synthetic data from these representative simulations to calibrate the ODE models in Sects. 3, 4, and 5.

3 Identifiability Analysis

3.1 Structural Identifiability Analysis

The concept of structural identifiability was first introduced in 1970 by Bellman and Astrom [2]; they asked whether, given perfect input data and a measured output signal that relates to available experimental data, it is possible to determine parameters associated with a dynamical systems model. A model identification question then asks whether it is possible to uniquely recover all model parameters given sufficient, error-free data about one (or more) model outputs. We note here that this type of analysis is sometimes referred to as structural identifiability analysis, as it relies solely on the properties of the dynamical system and respective model observable outputs. It should not be confused with practical identifiability analysis, which is concerned with the ability to recover parameter values from error-prone experimental data and depends on the computational approach used to parameterize

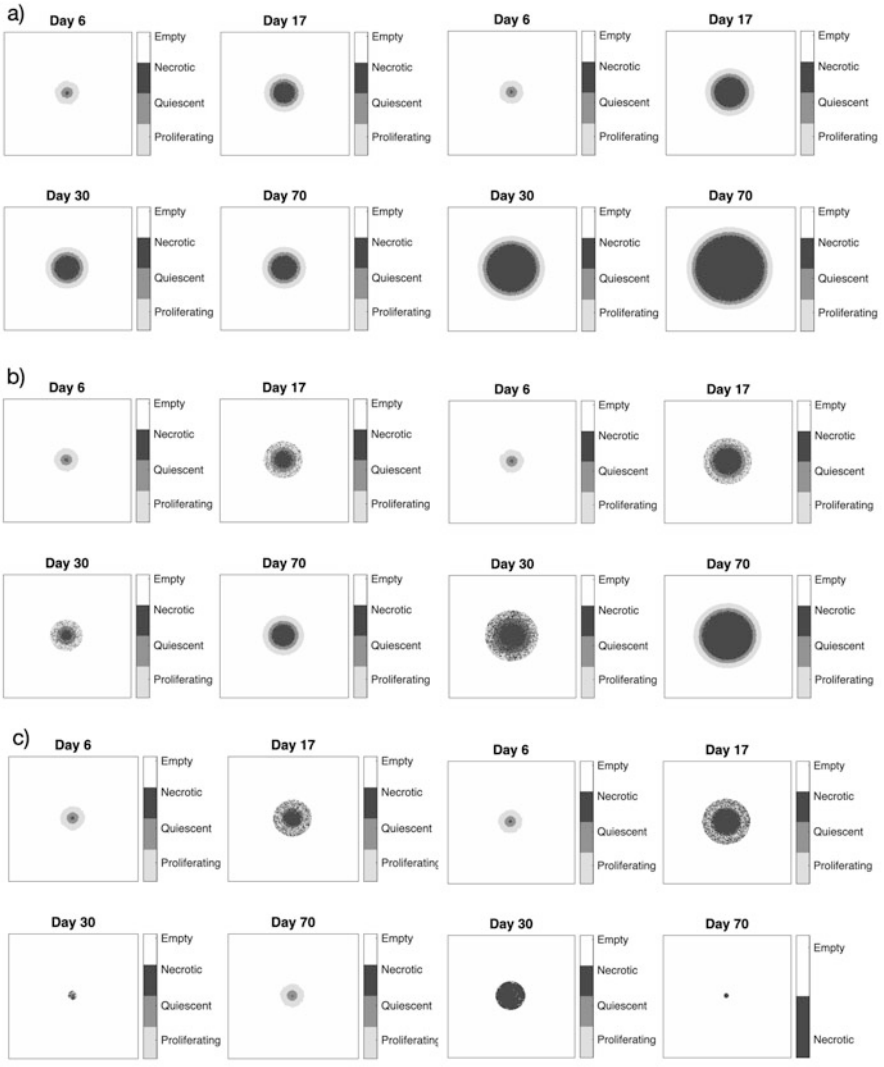

Fig. 4 (**a**) *No treatment.* Tumor composition on Days 6, 17, 30, and 60 when the rate of necrotic decay is high (left) and low (right), and no treatment is applied. (**b**) *Low radiosensitivity* ($\alpha/\beta = 9$). Tumor composition on Days 6, 17, 30, and 60 when the rate of necrotic decay is high (left) and low (right), and radiotherapy is applied. (**c**) *High radiosensitivity* ($\alpha/\beta = 1$). Tumor composition on Days 6, 17, 30, and 60 when the rate of necrotic decay is high (left) and low (right), and radiotherapy is applied

the model. It is important to note that many difficulties related to estimating parameters by fitting mathematical models to datasets may stem from lack of structural identifiability. In such cases, the system may not admit unique parameter

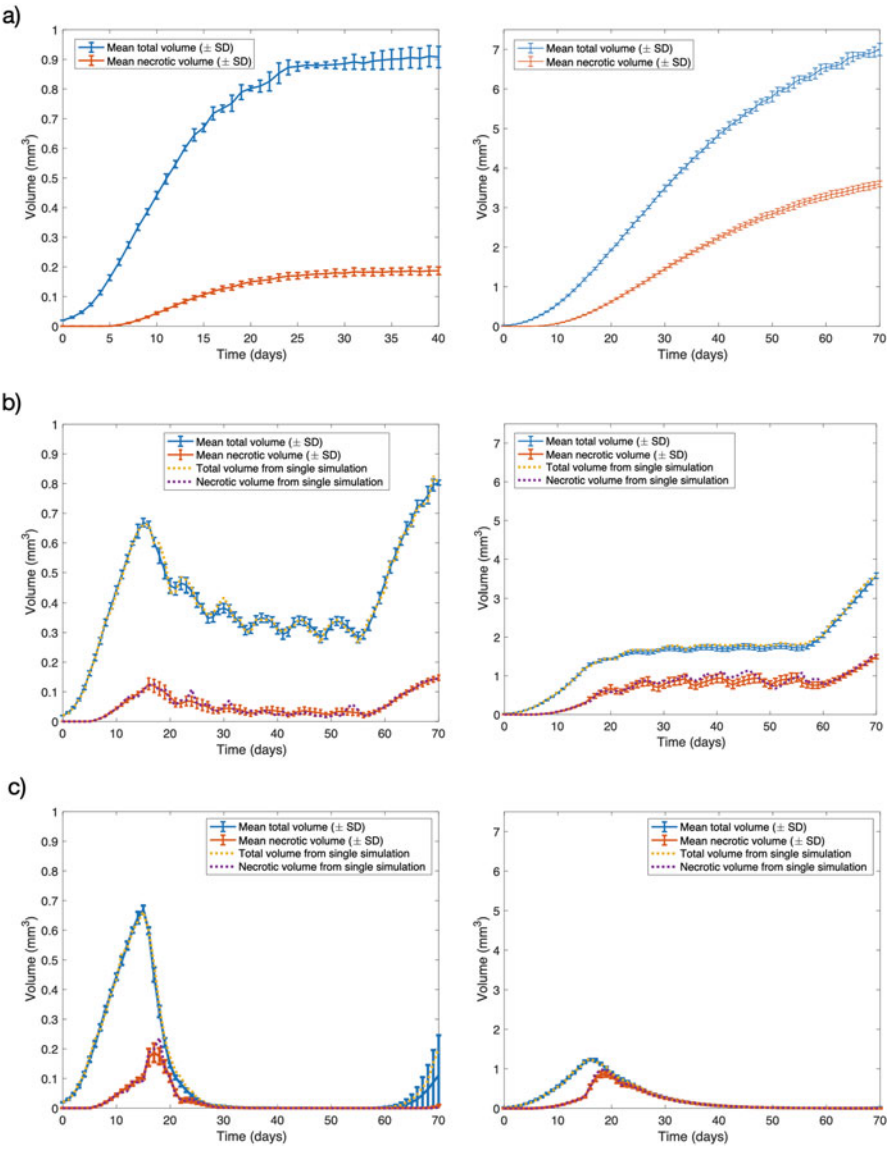

Fig. 5 (**a**) *Mean CA spheroid volume with no treatment.* The tumor grows in the absence of radiotherapy, when the rate of necrotic cell death is high (left), and low (right). (**b**) *Mean CA spheroid volume with low radiosensitivity.* The tumor exhibits radiosensitivity level $\alpha/\beta = 9$, when the rate of necrotic cell death is high (left), and low (right). (**c**) *Mean CA spheroid volume for tumor cells with high radiosensitivity.* The tumor exhibits radiosensitivity level $\alpha/\beta = 1$, when the rate of necrotic cell death is high (left), and low (right). The plots in Fig. 5 show the time evolution of the mean total tumor volume and mean necrotic volume, obtained by averaging over 100 realizations of the CA. The error bars indicate the standard deviation at each time point. The dashed lines show the total tumor volume and the necrotic volumes from the representative simulation used to fit to the ODE models. The figures on the left correspond to spheroids with high rates of necrotic decay ($p_{NR} = 0.015$) and small necrotic volumes; those on the right correspond to spheroids with low rates of necrotic decay ($p_{NR} = 0.004$) and large necrotic volumes

sets for a given observed model output. For further details about the techniques we use, we refer the interested reader to [5, 6]. In the remainder of this section we study the structural identifiability of the one- and two-compartment models presented in Sect. 2. Due to the form of the treatment terms, we employ the Taylor series approach [5, 23] to establish model structural identifiability and then compare our results with those obtained through a generating series approach implemented in Matlab through the GenSSI package [7]. To the best of our knowledge, assessing and characterising the structural identifiability of CAs remains an open problem. One approach would be to derive mean field descriptions of the CAs and to perform identifiability analysis of the resulting equations. Alternatively, several authors have shown how Approximate Bayesian Computation can be used for parameter inference of CAs [15, 16].

3.1.1 One-Compartment Model: No Radiation

In this case the model reduces to logistic growth, with

$$\begin{aligned} \frac{dV}{dt} &= AV\left(1 - \frac{B}{A}V\right), \\ V(0) &= V_0. \end{aligned} \tag{8}$$

We wish to establish the structural identifiability of the model with unknown parameters $p = \{A, B\}$, observable quantity $y(t; p) = V(t)$ and known initial conditions $V(t = 0) = V_0$.

Before we state results, we briefly outline the Taylor series approach, as delineated in [5, 23]. We assume that the observation function $y(t; p)$ is analytic in a neighborhood of some time. Then we can evaluate $y(t; p)$ and its successive time derivatives in terms of the model parameters and initial conditions at time $t = 0^+$ using

$$y(t, p) = y(0^+; p) + y^{(1)}(0^+; p)t + y^{(2)}(0^+; p)\frac{t^2}{2!} + \ldots + y^{(i)}(0^+; p)\frac{t^i}{i!} + \ldots, \tag{9}$$

where

$$y^{(i)}(0^+; p) = \frac{d^i y}{dt^i}(0^+; p). \tag{10}$$

Given Equations (9) and (10), the problem reduces to identifying a system of algebraic equations that relate the unknown model parameters (here A and B) to known values of the observable $y(t; p)$ and its derivatives at $t = 0^+$. If these equations admit unique solutions for each parameter, we consider the system to be structurally identifiable; otherwise, it is either locally identifiable or non-identifiable.

Since for our problem only A and B are unknown, we start by computing the first two Taylor series coefficients:

$$V'(0^+) = AV(0^+)\left(1 - \frac{B}{A}V(0^+)\right) \tag{11}$$

$$V''(0^+) = AV'(0^+) - 2BV(0^+)V'(0^+). \tag{12}$$

Denoting by $a_0 = V(0^+)$, $a_1 = V'(0^+)$ and $a_2 = V''(0^+)$, we obtain the following pair of simultaneous equations for the unknown parameters A and B:

$$a_1 = Aa_0 - Ba_0^2 \tag{13}$$

$$a_2 = a_1 A - 2a_1 Ba_0, \tag{14}$$

with solution

$$A = 2\frac{a_1}{a_0} - \frac{a_2}{a_1} \quad \text{and} \quad B = \frac{a_1}{a_0^2} - \frac{a_2}{a_1 a_0}. \tag{15}$$

Since a_0, a_1, and a_2 are known, the variables A and B are globally structurally identifiable except for at most a set of points of zero measure (i.e., points for which $a_0 = 0$, $a_1 = 0$). We remark that in this case we have exploited the reduced number of parameters due to model rescaling. For the original, dimensional model (1), the three parameters $p = \{\lambda, K, \eta\}$ are not uniquely identifiable: the Taylor series coefficients do not contain enough information to uniquely extract the three parameters from tumor volume observations.

Next, we repeat the above analysis using the generating series approach implemented in GenSSI [7]. Briefly, GenSSI implements a generating series approach coupled with identifiability tableaus [1] for linear and non-linear systems of ODEs. The underlying principle is to obtain equations for model parameters by computing successive Lie derivatives of the right hand side of the ODE system and model observable quantities ($y(t; p) = V(t)$ in our case). If the solution of the system of parameter equations is unique then the parameters are declared globally identifiable. We implemented our model in GenSSI and confirmed that $\{A, B\}$ in (2) are structurally identifiable, whereas $\{\lambda, K, \eta\}$ in (1) are not structurally identifiable. These results are consistent with results of the Taylor series approach outlined above.

3.1.2 One-Compartment Model: Point Radiation Treatment

The problem now reads

$$\frac{dV}{dt} = AV\left(1 - \frac{B}{A}V\right) \quad \text{for} \quad t_i^+ < t < t_{i+1}^-, \tag{16}$$

$$V(t_i^+) = \exp(-\alpha d - \beta d^2)\, V(t_i^-).$$

Since, in this case, the experimental observable $y(t; p) = V(t)$ corresponds to tumor volume, parameters A and B are identifiable from observations of the system without therapy (see Equation (15)). In addition, for a radiotherapy treatment the timing and dose t_i and d are also known quantities. Thus, the unknown parameters to identify from the tumor volume measurements are α and β. In order to identify α and β, we should compute, in principle at least, two Taylor series coefficients by employing a Taylor series expansion around the treatment time, t_i (note that we use one-sided limits and derivatives at $t_i^{\pm}$, as in the previous section).

Using Equation (16), it is straightforward to show:

$$A_1 = AA_0\Gamma - BA_0^2\Gamma^2, \tag{17}$$

$$A_2 = A_1(A - 2BA_0\Gamma), \tag{18}$$

where $A_0 = V(t_i^-)$, $A_1 = V'(t_i^+)$, and $A_2 = V''(t_i^+)$ are known, and $\Gamma = \exp(-\alpha d - \beta d^2)$.

Inspection of Equations (17) and (18) reveals that they do not admit unique solutions for α and β since α and β appear in both equations via the parameter grouping Γ. We declare α and β to be non-identifiable in this setting (the same results were obtained using GenSSI implementation). Therefore, in what follows, we fix α and vary β. This is reasonable since radiosensitivity of cancer is often characterized by the ratio α/β; we vary β to allow α/β to take on a range of values. In practice, estimates of the values of α and the ratio α/β could be obtained, for a particular biological tissue, by measuring the volume reduction caused by exposure to different radiotherapy doses and fitting these data to the linear quadratic model.

3.1.3 Two-Compartment Model

The two compartment model with treatment is given by the following model equations:

$$\frac{dY}{dt} = \lambda(1-\Phi)Y\left(1-(1-\Phi)\frac{Y}{K}\right) - \xi\Phi Y \tag{19}$$

$$\frac{d\Phi}{dt} = (1-\Phi)\left[\eta - \lambda\Phi\left(1-(1-\Phi)\frac{Y}{K}\right) - \xi\Phi\right], \quad \text{for} \quad t_i^+ < t < t_{i+1}^- \tag{20}$$

$$\Phi(t_i^+) = \Phi(t_i^-) + (1-\Phi(t_i^-))(1-\exp(-\alpha d - \beta d^2)). \tag{21}$$

Since the calculations are similar to, but more involved than those used for the one-compartment model, the details are presented in the Appendix. For completeness, we summarize our findings here. In the absence of treatment, with

$p = \{\lambda, K, \xi, \eta\}$, observable quantities $y(t; p) = \{Y(t), \Phi(t)\}$, and known initial conditions, we obtain unique solutions of the unknown parameters in terms of the observable quantities and their derivatives; we thus declare all four parameters structurally identifiable. We also repeat the analysis in the case in which only the tumor volume is observed, i.e., $y(t; p) = Y(t)$, but with known initial conditions for the tumor volume and necrotic fraction. In this case, we take higher order Taylor series coefficients (up to order 4) and find that $p = \{\lambda, K, \xi, \eta\}$ are structurally identifiable. GenSSI calculations confirm our findings.

When treatment is added, as detailed in the Appendix, we find that the radiation parameters α and β are not structurally identifiable, in agreement with the analysis for the single compartment model.

3.2 Practical Identifiability Analysis and Parameter Estimation

After establishing structural identifiability, it is natural to consider the practical identifiability of a model's parameters. In particular, given experimental data with measurement noise and a specific model, is it possible to uniquely determine a set of model parameter values that are most likely to produce the data?

One approach for determining the practical identifiability of the parameters while performing model calibration is through the use of a Metropolis algorithm, based on Markov Chain Monte Carlo (MCMC) techniques. Here, we construct Markov chains whose stationary distributions coincide with the posterior density of the parameters; thus, by sampling realizations of our parameters from this chain, we are effectively sampling from the parameter posterior density. The traditional Metropolis algorithm, as outlined in [25], constructs the posterior chains by drawing the next candidate, q^*, from a proposal function $J(q^*|q^{k-1})$, where q^{k-1} represents the previous parameter candidate. The goal of the Metropolis algorithm is to identify the set of parameter values that maximizes the likelihood function. This is equivalent to minimizing the sum-of-squares of the differences between the observed and predicted data. If the new candidate yields a smaller sum-of-squares error than the previous one, it is accepted as the next value in the posterior chain. Otherwise, we reject with some specified probability—see [25] for details—and the new state is taken to be the same as the old one, $q^k = q^{k-1}$. The traditional Metropolis algorithm assumes a symmetric proposal function J with respect to each of the individual parameters, though the Metropolis-Hastings algorithm [25] allows for asymmetric proposal functions. Here, we use the symmetric proposal function $J(q^*|q^{k-1}) = \mathcal{N}(q^{k-1}, C)$, where C is the covariance matrix for the parameter set.

In what follows, we perform model calibration using the Delayed Rejection Adaptive Metropolis (DRAM) algorithm for parameter estimation [11]. This extension of the traditional Metropolis algorithm includes two additional steps. The first, the delayed rejection step, allows for the proposal of an alternative parameter candidate from a narrower proposal distribution in place of outright rejection of the

original candidate. This results in greater mixing in our posterior MCMC chains by preventing the algorithm from stagnating on a single accepted candidate for long periods of time while multiple new candidates are rejected. During the adaptation step, a periodic adaptation of the parameter covariance matrix is performed to incorporate new information gained from accepted candidates. This covariance matrix is built into our proposal distribution; recall, we use $J(q^*|q^{k-1}) = \mathcal{N}(q^{k-1}, C)$, where C is the parameter covariance matrix. Thus, periodically updating C to reflect new information about the accepted parameter candidates will result in quicker convergence to the posterior densities. In this investigation, we use an adaptation interval of every 100 parameter candidates. For further information about the DRAM algorithm, we refer the interested reader to [11, 25].

After completing the parameter estimation process, we look for evidence of successful parameter recovery. First, we investigate the MCMC posterior chains for good mixing; we desire posterior chains that resemble white noise to suggest that the entire parameter space has been explored without extended stagnation on certain candidates. For visual examples of well-mixed posterior chains, we refer the reader to [25]. We also consider the pairwise parameter plots, as these can illustrate identifiability issues in several ways. Pairs of parameters whose chains are highly correlated in a strictly linear fashion are said to be unidentifiable in the sense that they cannot be uniquely identified by calibration with the available data; infinitely many pairs of parameter values would yield the same model response. Identifiability issues can also manifest as posterior densities that are unchanged from the specified prior distributions, indicating that the parameters are uninformed by the available data. By considering the above indicators, we can determine whether the quality and quantity of the data is sufficient to support the unique identification of all model parameters.

3.2.1 The Impact of Necrotic Fraction on Model Calibration

We now investigate the ability of the one- and two-compartment ODE models to fit synthetic data generated from the CA for different values of the CA parameter p_{NR}, the probability of removal of a necrotic cell. We begin by generating synthetic data in the form of a single representative realization of the CA with $\alpha/\beta = 1$ and all other parameters fixed at the nominal values provided in Table 1. For calibration, we use data from the first day of weeks 2–7 and day 70 (i.e. days 8, 15, 22, 29, 36, 43, 50, and 70), corresponding to the first day of treatment each week and then a post-treatment scan to check for tumor regrowth. We feed this data to the relevant ODE model and estimate all parameters (A, B, and β for the one-compartment ODE model; λ, k, η, ζ, and β for the two-compartment model). When investigating the performance of the two-compartment ODE model, we consider two cases: (1) providing tumor volume and necrotic fraction data, and (2) providing only tumor volume data, but still estimating necrotic fraction in the absence of that data.

The results presented in Fig. 6a show that the one-compartment ODE fit improves as p_{NR} increases, suggesting that the one-compartment ODE is better able to model

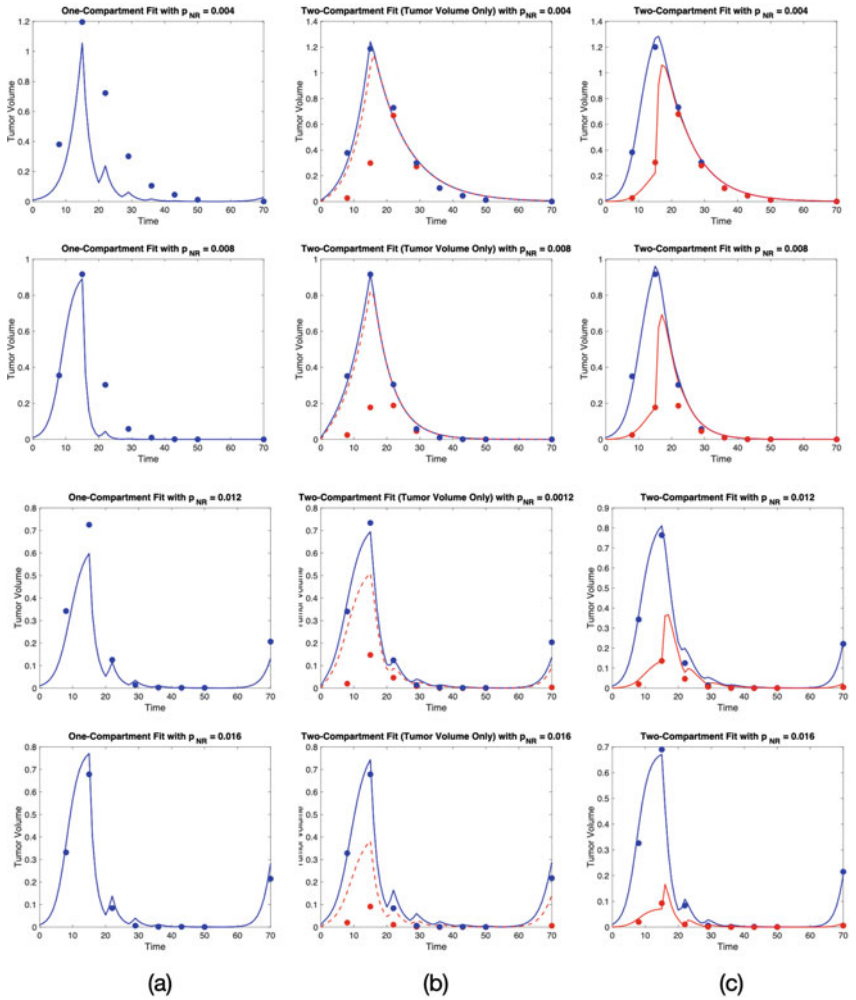

Fig. 6 Comparing the ability of the one- and two-compartment ODE models to fit synthetic data collected from the CA for varying values of p_{NR}, 0.004, 0.008, 0.012, and 0.016. (**a**) One-compartment model fit to tumor volume data from the CA model. (**b**) Two-compartment model fit to CA tumor volume data without necrotic core fraction data. (**c**) Two-compartment model fit to the CA data (for both tumor volume and necrotic fraction). Blue represents tumor volume; red represents necrotic fraction. Solid curves are fit to the given data; dashed curves show necrotic fraction estimate from the two-compartment ODE model generated in absence of necrotic data

scenarios in which the necrotic portion of the tumor is minimal. In contrast, the two-compartment ODE with both sets of data supplied fits the data with reasonable accuracy in all four cases, regardless of the p_{NR} value—see Fig. 6c. When the necrotic fraction data is not supplied, the two-compartment model generates good fits to total tumor volume data alone, for all values of p_{NR}. However, in all cases it

vastly overestimates the necrotic fraction towards the beginning of the observation period, as shown in Fig. 6b. Thus, using only tumor volume data, we can reliably use the two-compartment model to make predictions about tumor volume, but cannot rely on the calibration to produce an accurate portrayal of the tumor heterogeneity. Parameter values generated by fitting the models to the CA data for various values of p_{NR} are given in Table 2.

3.2.2 The Impact of Noise on Parameter Recovery

In this case study, we seek to understand how the addition of noise to the data affects our ability to recover model parameters; specifically, at what point does the noise level overcome our ability to determine what parameter values were used to generate the data? We generate noise uniformly on an interval centered over each data point, with a range depending on the level of noise desired; that is, at each time t,

$$y_{\text{noise}}(t) = y_{\text{exact}}(t)(1 + \varepsilon),$$

where $\varepsilon \sim \mathcal{U}(-x, x)$ if we desire $100x\%$ noise.

We generate noisy synthetic data for the one-compartment and two-compartment models in the presence and absence of radiotherapy, and then calibrate the data against the model used for its generation, varying the noise level from 1% to 20%. The results are presented in Fig. 7, separated into three cases for easier visualization: (a) one-compartment model, (b) two-compartment model with a low necrotic fraction (approximately 20% necrotic tissue over the long term without treatment), and (c) two-compartment model with a high necrotic fraction (approximately 50% necrotic tissue over the long term without treatment). For the one-compartment model, we estimate the parameter set $[A, B, \beta]$ and compare to the true parameter set $[0.5, 2, \beta]$, where the value of β used to generate the data depends on the radiosensitivity level specified. For the two-compartment model, we estimate $[\lambda, K, \eta, \zeta, \beta]$ and compare to the true parameter set $[1, 0.5, 0.5, 0.5, \beta]$ in the high-necrotic case, and $[1, 0.5, 0.5, 2, \beta]$ in the low-necrotic case. In all cases, we fix $\alpha = 0.14$ and estimate the radiosensitivity ratio α/β by varying β only, since we encounter structural identifiability issues when trying to fit both α and β simultaneously—see Sect. 3.1. Since we are interested in observing the behavior for a variety of radiosensitivity levels, we consider four cases for each model: no-treatment, high radiosensitivity ($\alpha/\beta = 1$, so $\beta = 0.14$), medium radiosensitivity ($\alpha/\beta = 3$, so $\beta = 0.0467$), and low radiosensitivity ($\alpha/\beta = 9$, so $\beta = 0.0156$). Data is supplied for calibration for two pre-treatment times (days 8 and 15), five treatment times (during days 22–50), and one post-treatment time (day 70). In each case, we measure the average relative error in parameter estimates (we compare the parameter values used to generate the data to those obtained via the parameter estimation procedure) over 10 runs, for four levels of noise: 1%, 5%, 10%, and 20%. As seen in Table 3 and Fig. 7, in all cases there is a positive correlation between the

Table 2 Parameter values for fits of one- and two-compartment models (with and without necrotic fraction data included) to the CA for varying p_{NR} values. These are the values of the parameters used to generate the graphs in Fig. 6

	One-comp. model				Two-comp. w/o Nec. data				Two-comp. w/o Nec. data			
p_{NR}	0.004	0.008	0.012	0.016	0.004	0.008	0.012	0.016	0.004	0.008	0.012	0.016
A	0.339	0.513	0.451	0.528								
B	0.111	0.552	0.703	0.668								
λ					1.696	2.724	2.959	2.749	0.565	0.517	0.570	0.691
k					2.975	2.639	1.358	2.572	1.138	1.106	0.673	0.758
η					1.459	2.529	2.542	2.340	0.074	0.073	0.133	0.136
ζ					0.098	0.171	0.898	2.177	0.124	0.202	0.639	1.144
β	0.113	0.243	0.140	0.170	0.470	0.426	0.123	0.116	0.235	0.245	0.132	0.185

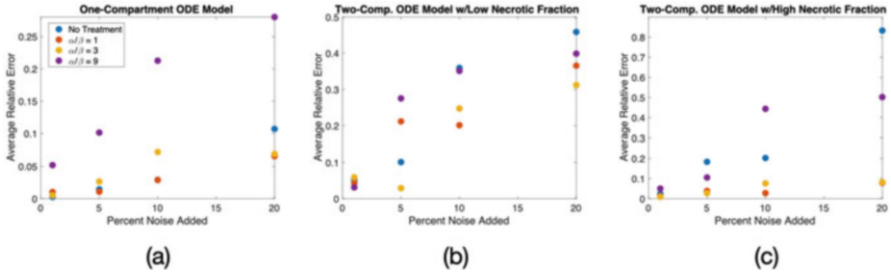

Fig. 7 Average relative error between the parameter values used to generate the synthetic data and those obtained by fitting to noisy data, averaged over 10 iterations in each trial for four different noise levels. (**a**) One-compartment ODE model with varying noise. (**b**) Two-compartment ODE model with low necrotic fraction ($\zeta = 2$) and varying noise. (**c**) Two-compartment ODE model with high necrotic fraction ($\zeta = 0.5$) and varying noise. In general, increased noise levels lead to higher relative errors

noise level and the average relative error in the parameter estimates. That is, as the level of noise in the data increases, the accuracy with which the true parameter values can be recovered decreases.

In Fig. 8, we focus on the one-compartment model with treatment and a low α/β ratio. For each noise level, we plot the posterior densities of our three parameters post-calibration. With 1% noise, the posterior distributions are centered at the values used to generate the data and are extremely well-informed, as illustrated by their narrow posterior densities. As the noise level increases, the posterior distributions widen, indicating less well-informed parameter estimates with greater variability. In addition, the distributions tend to drift rightwards as the noise level increases, suggesting that all parameter values are being overestimated. While we cannot be sure as to the cause of this "drifting" effect, we hypothesize that it may be due to our defining of the noise on a local scale, such that pre-treatment data will be "noisier" on average than data later on in the observation period. Further work is required to confirm this causation; meanwhile, we should remain cautious when dealing with noisy pre-treatment data, as it may play a significant role in the calibration of the parameters.

4 Model Selection

We now move to the question of determining how to choose the model most appropriate for the available data. When patient data is collected, we seek to identify a model that can be well calibrated to the data and make accurate predictions, while also retaining simplicity in terms of the underlying mechanisms and number of parameters. Minimizing the number of variables and parameters is desirable, due to the cost of collecting the data and conducting parameter estimation. For example,

Table 3 Mean relative error in parameter estimates, averaged over 10 runs in each category. High, Med and Low categories indicate radiosensitivity, corresponding to $\alpha/\beta = 1$, $\alpha/\beta = 3$, $\alpha/\beta = 9$, respectively. As the noise level increases, the average relative error tends to increase as well, indicating that parameter recovery is less successful

	Average relative error in parameter estimates											
	One-comp. model				Two-comp. w/o Low Nec.				Two-comp. w/o High Nec.			
Noise	None	High	Med	Low	None	High	Med	Low	None	High	Med	Low
1%	0.003	0.010	0.005	0.052	0.052	0.043	0.059	0.031	0.025	0.013	0.009	0.051
5%	0.015	0.011	0.027	0.010	0.101	0.212	0.029	0.275	0.184	0.039	0.028	0.106
10%	0.029	0.029	0.072	0.213	0.360	0.201	0.247	0.351	0.202	0.028	0.076	0.444
20%	0.107	0.066	0.069	0.280	0.459	0.366	0.312	0.399	0.833	0.079	0.082	0.502

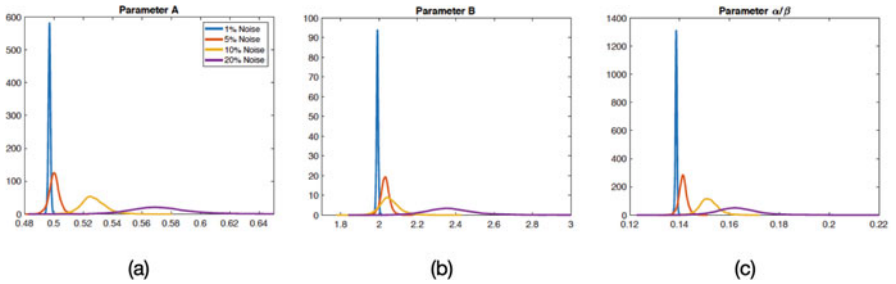

Fig. 8 Series of results showing how the posterior parameter distributions for the one-compartment model, with treatment and a low α/β ratio, depend on the level of noise in the data. As the noise level increases, the posteriors become less well-informed and tend to overestimate the parameter values compared to those used to generate the data

our one-compartment model requires only tumor volume data, whereas our two-compartment model also uses necrotic fraction data as an input, which is far more challenging and expensive to collect. Additionally, fitting the two-compartment model requires exploration of a five-dimensional parameter space, whereas fitting the one-compartment model requires investigation of a parameter space which is only three-dimensional. Therefore, we seek to understand when use of the one-compartment model is adequate, versus when the additional complexity of the two-compartment model (and cost for data collection) is necessary to accurately describe the tumor growth dynamics. In this section, we study the goodness of fit of the one-compartment model to synthetic data generated from the two-compartment model. In doing so, we address some of the questions raised here.

We generate synthetic data for total tumor volume and necrotic fraction from the two-compartment model, defined in Equations (4a) and (4b), by sweeping across the following region of five-dimensional parameter space: $\{(\lambda, K, \eta, \zeta, \beta) \in \mathbb{R}^5 \mid 0.2 \leq \lambda \leq 1,\ 0.1 \leq K \leq 1,\ 0 \leq \eta \leq \lambda,\ 0.5 \leq \zeta \leq 2,\ 0.014 \leq \beta \leq 0.14\}$. We generate a large number of samples ($O(10^3)$) using the Halton sequence, a quasi Monte Carlo method [13]. Then, we fit the synthetic tumor volume data using the one-compartment model defined in Equation (2).

We identify parameter regimes for which the one-compartment model provides a good approximation to the data from the two-compartment model and make note of other parameter regimes for which the one-compartment model does not approximate the two-compartment data well. The overall relation between the parameters are shown in Fig. 9; the parameter pairs (λ, A), (λ, B), are positively correlated since λ is the net tumor growth rate and, by definition, correlated to parameters representing the growth rate (A) and the inverse of the carrying capacity (B). Additionally, the values of β that represent radiosensitivity in the two models are positively correlated to each other, but not correlated to any other parameters, since the radiotherapy response is assumed to be independent of the tumor model parameters. We find negative correlation between the pair (η, A), since the natural death rate η is captured in the overall growth rate A, and between (K, B), also due

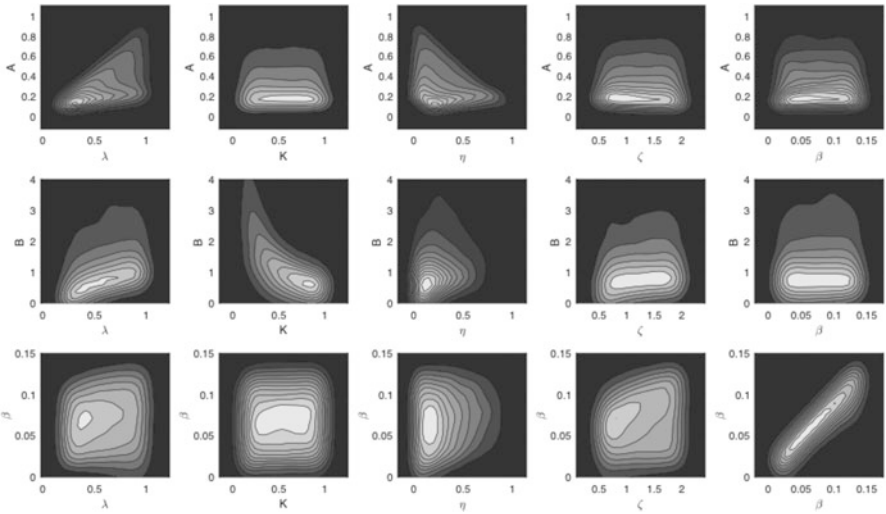

Fig. 9 Relationships between the two-compartment parameter samples $(\lambda, K, \eta, \zeta, \beta)$ and the fitted one-compartment parameters (A, B, β), depicted by contour plots showing the density of the parameter pairs. The parameter pairs (λ, A) and (λ, B) are positively correlated, while (η, A) and (K, B) are negatively correlated. The most apparent relation is between β from the two-compartment model and fitted β in the one-compartment model where the two values are linearly related

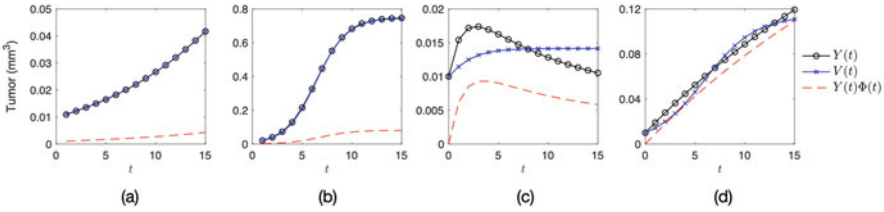

Fig. 10 Examples of tumor volume data $V(t)$ (∘) and necrotic volume (- -) generated from the two-compartment model defined by Equations (4a) and (4b) and corresponding tumor volume values $Y(t)$ (×) obtained by fitting the one-compartment model defined by Equation (2) to the data. In subplots (**a**) and (**b**), the approximation is accurate; in subplots (**c**) and (**d**), the one-compartment model fails to accurately capture the behavior of the data. Although in (**d**), the overall trend is captured by the one-compartment model, the slope difference toward the end time point may lead to inaccurate predictions in future time points

to their inverse relation. The necrosis clearing rate ζ is not strongly correlated with any of the parameters from the one-compartment model.

Figure 10 displays simulation data generated from the two-compartment model in the absence of treatment. We plot the total tumor volume $(Y(0), Y(1), \cdots, Y(t_f))$ and the corresponding necrosis fraction $(\Phi(0), \Phi(1), \cdots, \Phi(t_f))$, where $t_f = 15$. Also shown are the corresponding fits $(V(0), V(1), \cdots, V(t_f))$ of the one-compartment model in Equation (2) to the synthetic data. We find that the

Table 4 Parameter values used to generate the plots in Figs. 10 and 12. They are the parameter values of the two-compartment model in Equations (4a) and (4b) and the fitted parameter values in the one-compartment model Equation (7)

	Two-comp. model					One-comp. model		
	λ	K	η	ζ	β	A	B	β
Fig. 10a	0.33	0.60	0.26	1.97		0.1159	2.4326	
Fig. 10b	0.89	0.90	0.23	1.90		0.6738	0.9061	
Fig. 10c	0.92	0.11	0.92	0.77		0.0834	4.9750	
Fig. 10d	0.954	0.29	0.92	0.03		0.3906	3.3525	
Fig. 12a	0.811	0.934	0.0278	1.58	0.0276	0.779	0.851	0.0123
Fig. 12b	0.508	0.705	0.115	1.50	0.102	0.382	0.59	0.0984
Fig. 12c	0.819	0.170	0.806	0.723	0.0234	0.113	5.00	0.0419
Fig. 12d	0.583	0.972	0.0205	0.531	0.127	0.499	0.468	0.1006

one-compartment model accurately fits synthetic data which is either monotonically increasing (Fig. 10a), or increasing and saturating (Fig. 10b). It is unable to accurately fit data generated from the two-compartment model for which the growth dynamics are either non-monotonic (Fig. 10c) or for which the necrotic region is large (Fig. 10d). We note that the poor fit of the one-compartment model to non-monotonic growth (Fig. 10c) is expected, since solutions to the one-compartment model are monotonic. The parameter values used to generate the synthetic data presented in Fig. 10 are included in Table 4, together with the parameter values obtained by fitting the one-compartment model to the synthetic data.

To quantify how well the one-compartment model fits data generated from the two-compartment model (i.e., the goodness of fit of the one-compartment model), we compute the relative error $e \doteq \|\boldsymbol{y} - \boldsymbol{v}\|_2/\|\boldsymbol{y}\|_2$ between the data $\boldsymbol{y} = (Y(0), Y(1), \cdots, Y(t_f))$ and the fitted values $\boldsymbol{v} = (V(0), V(1), \cdots, V(t_f))$ for each parameter sample. Each point plotted in Fig. 11 represents the relative error when the one-compartment model is fitted to synthetic data generated from the two-compartment model for a sample parameter set. Figure 11a shows how the goodness of fit of the one-compartment model decreases as the tumor size reduction indicator $(Y(t_f)/\max_t Y(t))$ decreases. Smaller values of this quantity correspond to tumors which have more pronounced, non-monotonic growth dynamics that cannot be modeled using the one-compartment model that only produces logistic curves. Figure 11b shows that the goodness of fit of the one-compartment model decreases as the necrotic volume of the synthetic data increases. This suggests that measuring the necrotic volume at a given time could be used to decide whether a one-compartment model can accurately fit the data or whether a two-compartment model is needed. More specifically, we observed an increased level of relative error e in some of the fits of the two-compartment data when the necrotic proportion $\Phi(t_f)$ at the final time point is large; the relative error can be large when the necrotic proportion $\Phi(t_f)$ is large.

We use a similar workflow to study the goodness of fit of the one-compartment model to the two-compartment model with radiotherapy. As in the case without

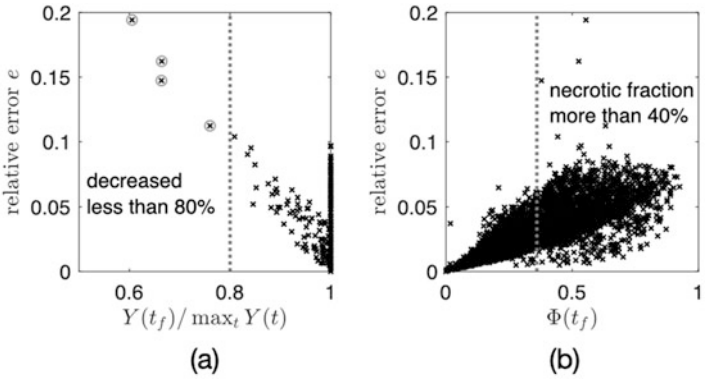

Fig. 11 Relative error e between the two-compartment tumor volume data generated from Equations (4a) and (4b) and the one-compartment fit using Equation (2) is plotted with respect to (**a**) tumor size reduction $Y(t_f)/\max_t Y(t)$ and (**b**) necrotic fraction $\Phi(t_f)$, where $t_f = 15$ is the final time of our simulation. Subplot (**a**) shows that data with non-monotonic tumor growth cannot be captured accurately by the one-compartment model. In particular, we highlight the data with $Y(t_f)/\max_t Y(t) < 0.8$ (○) that show large relative error values. Subplot (**b**) shows that the proportion of necrotic cells is also related to an increased error using the one-compartment model. We observe cases with increased error levels in data for which $\Phi(t_f) > 0.4$

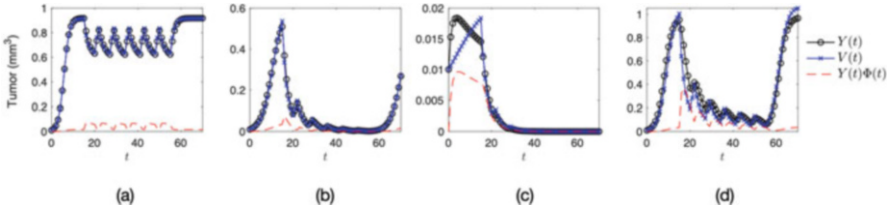

Fig. 12 Series of results showing the goodness of fit of the one-compartment model to data generated from the two-compartment model when treatment with radiotherapy is applied. We present synthetic data of tumor volume (○) and necrotic volume (- -) that were generated from the two-compartment model (see Equations (4a) and (4b)), together with tumor volume data (×) obtained by fitting to the one-compartment model (see Equation (7)). The fits are accurate in cases (**a**) and (**b**) but not in (**c**) or (**d**). In particular, the one-compartment model cannot reproduce the results of two-compartment model when the untreated growth dynamics are not well-fitted (**c**), and when the necrotic region is large (**d**)

treatment, Fig. 12 shows that there are situations for which the one-compartment model accurately captures the tumor dynamics (Fig. 12a, b) and others for which it does not (Fig. 12c, d). The two scenarios which typically yield inaccurate fits correspond to cases for which tumor growth before treatment can be not captured well (Fig. 12c) and/or the necrotic region is large (Fig. 12d). Although one can assume that the overall dynamics in Fig. 12d are captured reasonably well, the peak and trough during treatment could not be accurately fitted. This could potentially cause more inaccurate predictions when only a few, noisy data points are added. The parameter values used to generate the figure are included in Table 4.

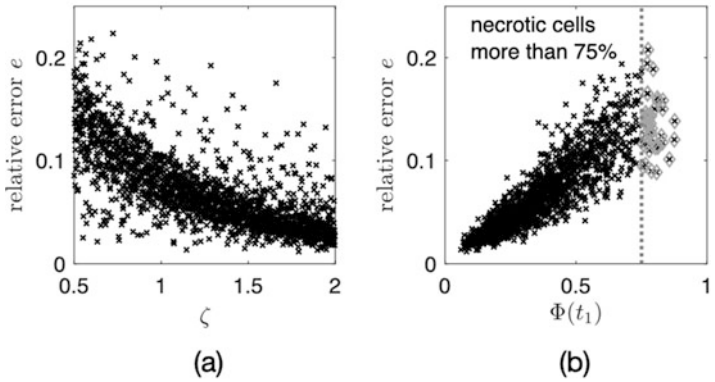

Fig. 13 Relative error e between the two-compartment data generated from Equations (4a) and (4b) with treatment and the one-compartment fit using Equation (7) with respect to (**a**) the parameter ζ and (**b**) necrotic fraction $\Phi(t_1)$, where $t_1 = 20$ is the time after the first week of treatment. The proportion of necrotic cells is positively correlated with the relative error. We highlight the data with $\Phi(t_f) > 0.75$ ($\Diamond$) that show large errors

As in the example without treatment, the relative error e is computed using the two-compartment tumor volume data $\boldsymbol{y}$ and the fitted tumor volume data from the one-compartment model $\boldsymbol{v}$. In this case, the data is collected daily until the final time $t_f = 70$. In Fig. 13b, the error is plotted with respect to the necrotic fraction after the first week of treatment $\Phi(t_1)$. The results show that the relative error increases as the ratio of necrotic core increases. We observe a correlation of the relative error with the necrotic core, that is more apparent compared to the study without treatment (Fig. 11b). In addition, in Fig. 13a, we observe correlated patterns with the necrotic core decay rate ζ, as this is the parameter that determines the size of the necrotic core.

Next, we verify that the ratio of necrotic core $\Phi(t)$ is a robust indicator of goodness of fit of the one-compartment model with treatment fitted to synthetic data with noise. Figure 14 shows the relative error, e, against the necrotic fraction $\Phi(t)$, while increasing the noise level of the data up to 20%. A positively correlated relationship between the relative error of the one-compartment fit and the necrotic fraction is apparent once treatment is applied, while the noise in the data reduces the effectiveness of the indicator in cases without treatment. In general, larger noise in the data impacts the ability to accurately fit the one-compartment model to two-compartment data, as is shown in Fig. 15, where the fitted β in the one-compartment model becomes inaccurate as the noise increases. In conclusion, we determine that the necrotic fraction is a good indicator of the quality of fit of the one-compartment model to synthetic data generated from the two-compartment model.

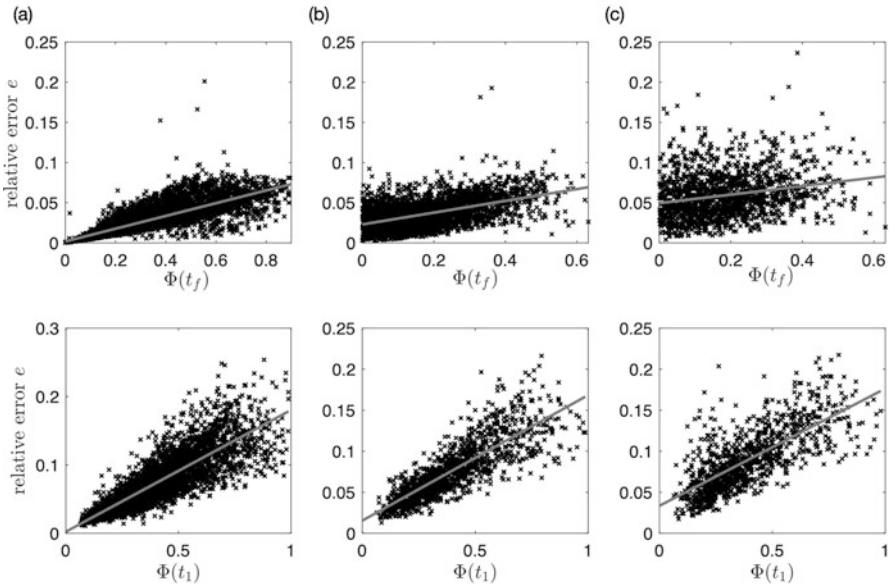

Fig. 14 The impact of adding noise to synthetic data from the two-compartment model on the ability to fit the one-compartment model without treatment (top) and with treatment (bottom). The relative error e is plotted with respect to the ratio of necrotic fraction $\Phi(t)$ while increasing noise in the data up to 20%. The least squares regression lines are shown. Despite the increased level of noise, the necrotic fraction remains a good indicator of the fitness of the one-compartment model when the treatment is given (bottom). (**a**) Noise 0%. (**b**) Noise 10%. (**c**) Noise 20%

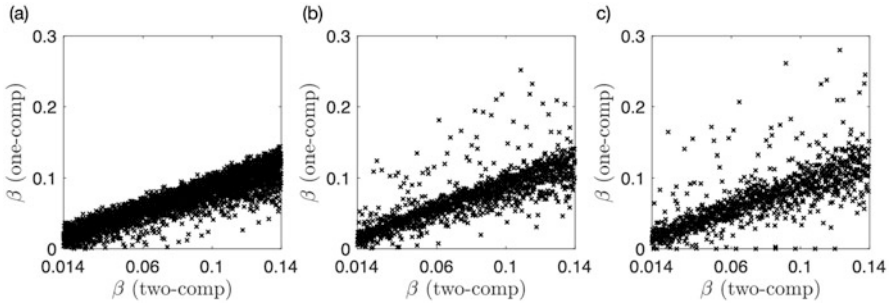

Fig. 15 The impact of adding noise to synthetic data from the two-compartment model reduces the ability to fit the one-compartment model. The parameter pairs of β that is used to generate the data from the two-compartment model and fitted using the one-compartment model show that larger noise in the data results in misfitted β values. (**a**) Noise 0%. (**b**) Noise 10%. (**c**) Noise 20%

5 Data Inclusion for Model Calibration

In an effort to predict how well simple mathematical models can be calibrated to clinical data of tumor volumes and to determine how much data is needed to accurately infer model parameters, we perform a sequential model calibration of: (i) the one-compartment model to data generated by the two-compartment model; (ii) the one-compartment model fit to CA data; and, (iii) the two-compartment model fit to synthetic data generated by the CA. In all cases, the data is generated from simulations in which tumors are treated with radiotherapy.

In each case, we fit the simpler model to a data set that includes tumor volumes, and necrotic fractions when relevant, measured once per week with 10% noise added. The calibration procedure is begun by fitting the lower-fidelity model to the first three data points (collected from the higher-fidelity model at days 8, 15, and 22); then the low-fidelity model is re-calibrated with the addition of each subsequent data point in an effort to determine a threshold at which we have "enough" data to accurately infer our model parameters with well-informed posterior distributions. When calibrating the one-compartment ODE model, the parameter set $[A, B, \beta]$ is estimated; for the two-compartment model, we estimate $[\lambda, K, \eta, \zeta, \beta]$. In all cases, α remains fixed at 0.14 to avoid identifiability issues.

At each calibration step, we calculate the relative error between the "fitted" and corresponding "true" parameter values. Since there is no explicit relationship between the parameters of the CA and those of the ODE models or mapping between parameters in the one- and two-compartment ODE models, we use a full set of in silico daily information to provide "true" parameter values. In each case, we initially fit the one- and two-compartment models to all 71 data points generated from the CA model and the one-compartment model to all 71 data points generated from the two-compartment model with no noise added and burn-in and subsequent MCMC chain lengths of 10,000. The parameter values generated from these fits are considered the "true" parameter values for each data set. We also assess the ability of the model to accurately predict the tumor regrowth (defined using the data point at day 70), by computing the absolute error in prediction when using the current calibrated model to predict forward in time. Figure 16 displays the calibration results when fitting to CA data in the "high" necrotic case (when the necrotic cells comprise 50% of total tumor volume in the absence of treatment). The final fits using all eight data points are generated using data with 10% noise added and burn-in and subsequent MCMC chain lengths of 10,000.

From Fig. 16, we see that when the necrotic fraction is high, the parameters converge well only when the two-compartment model is fit to both tumor volume and necrotic fraction data. In particular, the one-compartment model (Fig. 16a, b) does not accurately simulate a tumor with a high necrotic fraction; the fit is poor, and the estimated parameter values differ markedly from those used to generate the in silico data. Additionally, the entire fit changes when the final data point is added. This suggests that information about necrotic volume is needed to achieve a good fit to the data. When we fit the two-compartment model (Fig. 16c, d) to tumor

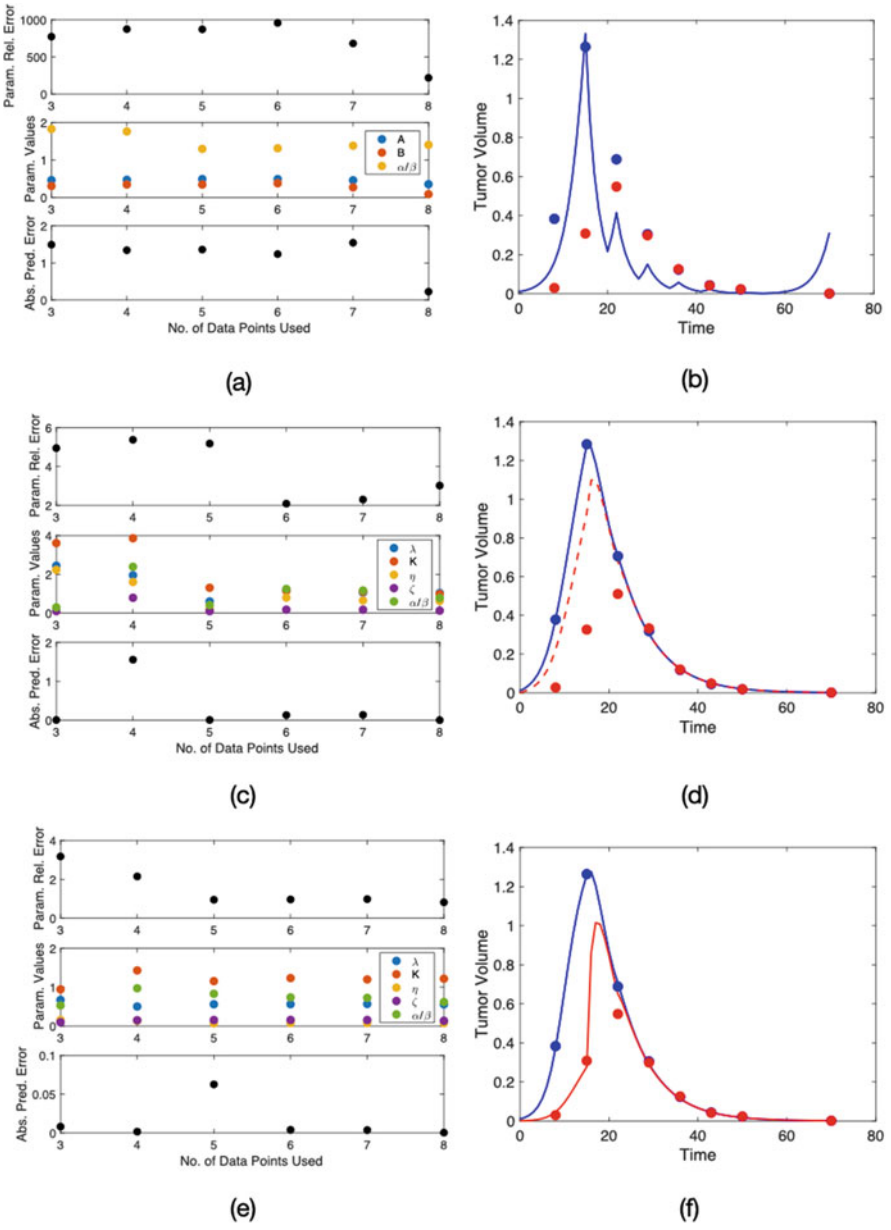

Fig. 16 Model fits with "high" necrotic fraction, "low" α/β ratio, and 10% noise. Blue represents tumor volume; red represents necrotic fraction. Solid curves are fit to the given data; dashed curves show necrotic fraction estimate from the two-compartment ODE model generated in absence of necrotic data. (**a**)–(**b**) One-compartment model fit, (**c**)–(**d**) Two-compartment model fit to tumor volume data only, (**e**)–(**f**) Two-compartment model fit to tumor volume and necrotic data generated from the CA

volume data only (no inclusion of necrotic information) generated from the CA, the overall fit is reasonable, but parameter values change markedly as the number of data points used for fitting varies. Additionally, the relative errors in the parameter values (as compared to those fit using the in silico data set) are high. Providing this supplemental information allows the slightly more complex two-compartment ODE model to accurately fit the high-necrotic tumor data; Fig. 16e, f shows a strong fit with greater consistency in the parameter estimates as data is added sequentially.

Analogous plots obtained by fitting CA data in the "low" necrotic case (when the necrotic cells comprise 20% of total tumor volume in the absence of treatment), are presented in Fig. 17. In this figure, all results relate to fits of the one-compartment model to the CA data as the tumor radiosensitivity varies. The radiosensitivity decreases from high ($\alpha/\beta = 1$) in (a)–(b), to medium ($\alpha/\beta = 3$) in (c)–(d), and low ($\alpha/\beta = 9$) in (e)–(f). Tumor regrowth data generated with medium and high α/β ratios is accurately predicted, even with only three data points. In the low α/β case, an additional data point is needed for accurate regrowth predictions. In all cases, the estimated parameter values converge rapidly to the true parameter values. We conclude that, despite its simplicity, for all three levels of radiosensitivity, the one-compartment model not only fits the data well, but can do so with just 3–4 few data points provided, for tumors with small necrotic regions.

6 Discussion

In this investigation, we have proposed a framework for choosing appropriate models, verifying the identifiability of their parameters, and calibrating those parameters with the available data. As a proof-of-concept, we investigated three distinct models of tumor growth: a one-compartment ODE model tracking tumor volume over time, a two-compartment model that includes an additional state variable representing the necrotic volume fraction, and a spatially explicit cellular automaton model that is more complex than the ODE models. We first showed that the one- and two-compartment model parameters are structurally identifiable without treatment, i.e. that one can uniquely recover all parameter values describing those models, given error-free model output data. However, we found that structural identifiability does not hold when treatment with radiotherapy is included in the models, since the radiosensitivity parameters are not uniquely identifiable. This led us to fix α and vary β only, when conducting model calibration.

Next we considered the practical identifiability of both ODE models. After adding varying levels of noise to synthetic tumor volume data, we found that the practical identifiability of parameter values becomes less well-informed for both the one-compartment and two-compartment models as the noise level increases. Further, the predicted parameter values tend to overestimate those values used to generate the data. By performing a sweep of the model parameters, we identified that the one-compartment model accurately fits synthetic data generated by the two-compartment model when the data is increasing monotonically or when the necrotic

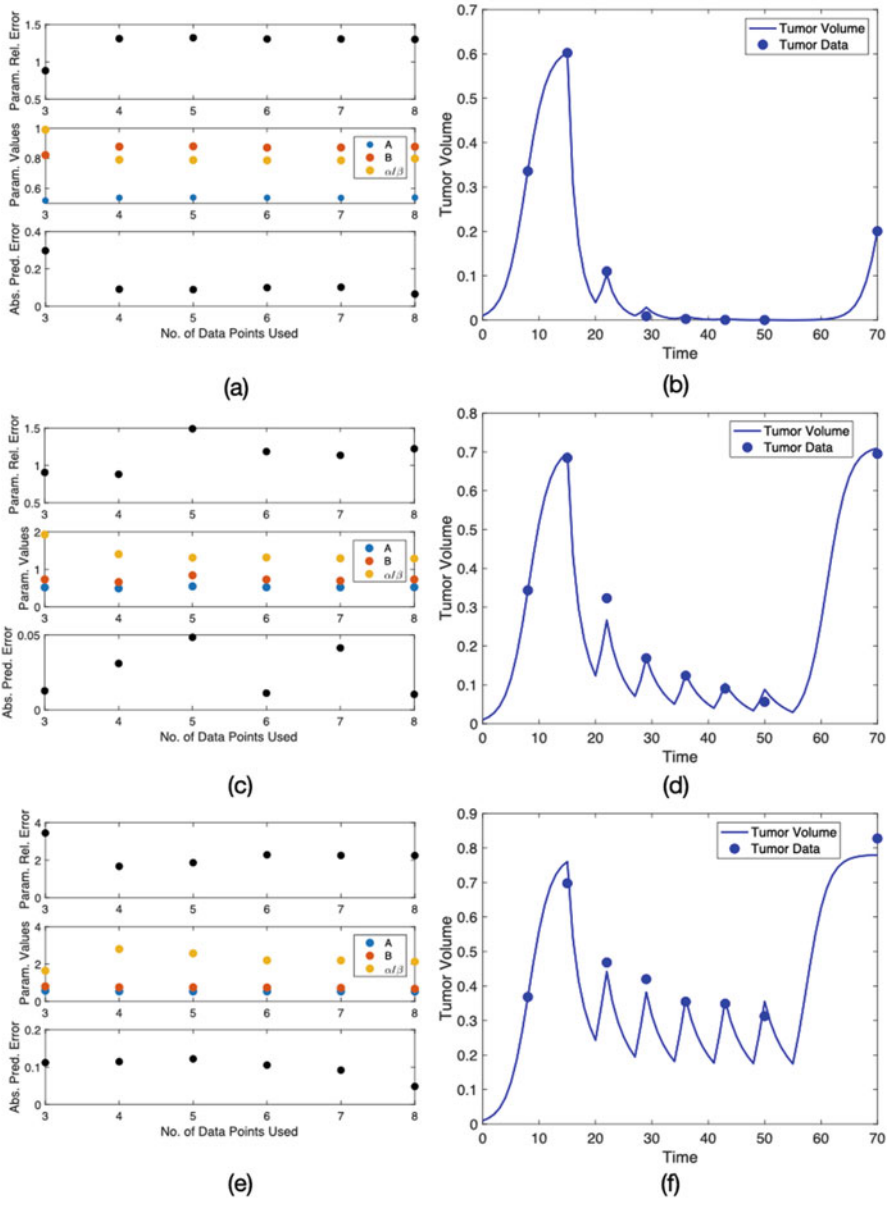

Fig. 17 The one-compartment model accurately describes tumors with a low necrotic fraction, for a range of α/β ratios. (**a**)–(**b**) $\alpha/\beta = 1$, (**c**)–(**d**) $\alpha/\beta = 3$, (**e**)–(**f**) $\alpha/\beta = 9$. In all three cases the relative errors in the estimated parameter values are small; the inclusion of the fourth data point results in absolute prediction error for tumor regrowth of less than 0.1 cm in all three cases

region is small. The error of the one-compartment fit to synthetic data increases as the size of the necrotic region increases, suggesting that in a clinical setting, the one-compartment ODE model becomes a less accurate predictor of tumor growth as the size of the necrotic region increases.

We also tested the ability of the simpler models to fit to data generated from the CA, using only a small number of data points. In agreement with the parameter sweep results, we found that for tumors with a large necrotic region, the estimated parameter values converge well only when fitting the two-compartment model to data on both total tumor and necrotic volumes. This implies that in cases with slow necrotic decay, information about tumor heterogeneity, rather than simply tumor volume, is necessary to fit these ODE models to the data. In such cases, i.e., when the necrotic region is large, the one-compartment ODE model will not accurately predict the response to treatment and tumor regrowth. However, when the necrotic fraction is small, it is possible to accurately identify parameters from the one-compartment model and to characterize tumor response to radiotherapy.

In the future, we aim to explore the sensitivity of the ODE models to intrinsic noise by adding noise to the parameter space before generating synthetic data and fitting the models to this data. We are also interested in incorporating multiple types of cells with different levels of radiosensitivity. We plan to examine how this additional complexity affects the overall predictive power of the ODE models and the amount of data needed to make accurate predictions. We also plan to conduct further experiments studying the impact of the quantity, type, and temporal location of available data on the ability to accurately calibrate various models.

Appendix: Structural Identifiability for the Two-Compartment Model

Below, we investigate the structural identifiability of the two-compartment model with radiotherapy, given in Equations (19) and (20), using the same techniques as presented in Sect. 3.1.

Case 1: No Radiation

In this case the model reads

$$\frac{dY}{dt} = \lambda(1-\Phi)Y(1-(1-\Phi)\frac{Y}{K}) - \xi\Phi Y$$
$$\frac{d\Phi}{dt} = (1-\Phi)\left[\eta - \lambda\Phi(1-(1-\Phi)\frac{Y}{K} - \xi\Phi)\right]$$

with unknown parameters $p = \{\lambda, K, \xi, \eta\}$, observable quantities $y(p; t) = \{Y, \Phi\}$, and known initial conditions. We repeat the analysis as before, using the Taylor coefficients. We define the following known quantities:

$$a_0 = Y(0^+)\ \ b_0 = \phi(0^+)\ \ a_1 = Y'(0^+)\ \ b_1 = \phi'(0^+) \quad a_2 = Y''(0^+)\ \ b_2 = \phi''(0^+).$$

We substitute these quantities into the model system to obtain:

$$\begin{aligned}
a_1 &= \lambda(1-b_0)a_0[1-(1-b_0)\frac{a_0}{K}] - \xi b_0 a_0 \\
&= [-(1-b_0)^2 a_0^2]\frac{\lambda}{K} + [(1-b_0)a_0]\lambda - [b_0 a_0]\xi \\
b_1 &= (1-b_0)\left[\eta - \lambda b_0(1-(1-b_0)\frac{a_0}{K} - \xi b_0)\right] \\
&\quad + [-(1-b_0)b_0]\lambda + [(1-b_0)^2 b_0 a_0]\frac{\lambda}{K} - [(1-b_0)b_0]\xi + [(1-b_0)]\eta.
\end{aligned}$$

We differentiate the model equations once more to obtain:

$$a_2 = \frac{(-2\lambda(-1+b_0)^2 a_0 - ((\xi+\lambda)b_0 - \lambda)K)a_1 - a_0 b_1(2(-1+b_0)\lambda a_0 + (\xi+\lambda)K)}{K}$$

$$b_2 = \frac{(3a_0 b_0^2\lambda + (-4\lambda a_0 + 2(\xi+\lambda)K)b_0 + \lambda a_0 - K(\xi+\eta+\lambda))b_1 + b_0 a_1 \lambda(-1+b0)^2)}{K}.$$

The above four equations can now be used to solve for each parameter as follows:

$$\begin{aligned}
K &= \frac{1}{[((b_0-1)b_2 - b_1^2)a_0^2 + a_2(b_0-1)^2 a_0 - a_1^2(b_0-1)^2]b_1} \\
&\quad \times [(b_1(b_0+1)(b_0 b_2 - b_1^2 - b_2)a_0^3 + ((a_1 b_2 + a_2 b_1)b_0^2 + ((-b_1^2 - b_2)a_1 + a_2 b_1)b_0 \\
&\quad - a_1 b_1^2 - a_2 b_1)(-1+b0)a_0^2 - (-a_2 b_0^3 + (a_1 b_1 + a_2)b_0^2 + 2a_1 b_1 b_0 - 2a_1 b_1) \\
&\quad \times a_1(-1+b_0)a_0 - a_1^3 b_0^2(-1+b_0)^2)(-1+b_0)], \\
\lambda &= \frac{-1}{(a_0 b_0 b_1 + a_1 b_0^2 - a_0 b_1 - 2a_1 b_0 + a_1)a_0^2 b_1} \\
&\quad \times [a_0^3 b_0^2 b_1 b_2 - a_0^3 b_0 b_1^3 + a_0^2 a_1 b_0^3 b_2 - a_0^2 a_1 b_0^2 b_1^2 \\
&\quad + a_0^2 a_2 b_0^3 b_1 - a_0 a_1^2 b_0^3 b_1 + a_0 a_1 a_2 b_0^4 - a_1^3 b_0^4 - a_0^3 b_1^3 \\
&\quad - 2a_0^2 a_1 b_0^2 b_2 - a_0 a_1^2 b_0^2 b_1 - 2a_0 a_1 a_2 b_0^3 \\
&\quad + 2a_1^3 b_0^3 - a_0^3 b_1 b_2 + a_0^2 a_1 b_0 b_2 + a_0^2 a_1 b_1^2 - 2a_0^2 a_2 b_0 b_1 + 4a_0 a_1^2 b_0 b_1 \\
&\quad + a_0 a_1 a_2 b_0^2 - a_1^3 b_0^2 + a_0^2 a_2 b_1 - 2a_0 a_1^2 b_1]
\end{aligned}$$

$$\eta = -\frac{a_0^2 b_0 b_2 - a_0^2 b_1^2 + a_0 a_2 b_0^2 - a_1^2 b_0^2}{b_1 a_0^2}$$

$$\xi = \frac{a_0^2 b_0 b_2 - a_0^2 b_1^2 + a_0 a_2 b_0^2 - a_1^2 b_0^2 - a_0^2 b_2 - a_0 a_1 b_1 - a_0 a_2 b_0 + a_1^2 b_0}{b_1 a_0^2}$$

Since we are able to obtain unique solutions for each of the four parameters, we declare them to be structurally identifiable. For GenSSI, only two Lie derivatives are needed which yield rank 4, and thus results show all four parameters are structurally identifiable, in agreement with our calculations above.

In addition to the above analysis, we also repeated the analysis in the case in which only the tumor volume could be observed, (i.e., $y(t; p) = Y(t)$), but with known initial conditions in tumor volume and necrotic fraction. In this case, we took higher order Taylor series coefficients (up to order 4) and obtained that $p = \{\lambda, K, \xi, \eta\}$ were structurally identifiable. Similarly, GenSSI took Lie derivatives up to order 4 and confirmed that all parameters were structurally identifiable.

Case 2: With Radiation Treatment

Similar to the single compartment model, here we examine the effect of a point treatment. The model equations read:

$$\begin{aligned}
\frac{dY}{dt} &= \lambda(1-\Phi)Y(1-(1-\Phi)\frac{Y}{K}) - \xi\Phi Y \\
\frac{d\Phi}{dt} &= (1-\Phi)\left[\eta - \lambda\Phi(1-(1-\Phi)\frac{Y}{K}) - \xi\Phi\right], \quad \text{for} \quad t_i^+ < t < t_{i+1}^- \\
\Phi(t_i^+) &= \Phi(t_i^-) + (1-\Phi(t_i^-))(1-\Gamma),
\end{aligned}$$

where $\Gamma = \exp(-\alpha d - \beta d^2)$. Since the other parameters are known and measured prior to treatment, as in the previous section, we want to solve for $p = \{\alpha, \beta\}$ assuming $y(p; t) = \{Y, \Phi\}$ as observable quantities. We let

$$A_0 = Y(t_i^+) \quad B_0 = \Phi(t_i^-) \quad A_1 = Y'(t_i^+) \quad B_1 = \Phi'(t_i^+),$$

and substitute these quantities into the model equations:

$$A_1 = \frac{-A_0^2\lambda(B_0-1)^2}{K}\Gamma^2 - A_0(\xi+\lambda)(B_0-1)\Gamma - A_0\xi$$

$$B_1 = \frac{[A_0\lambda(B_0-1)^2\Gamma^2 + (B_0-1)((K+A_0)\lambda+\xi K)\Gamma + K(\xi-\eta+\lambda)]\Gamma(B_0-1)}{K}$$

As with the one-compartment model, we find that the equations are not informative for α and β simultaneously, thus, we again declare the pair (α, β) to be non-identifiable in this setting. As before, we choose to fix α for all subsequent model calibrations and measure the ratio α/β to use as a measure of radiosensitivity.

References

1. E. Balsa-Canto, A. A. Alonso, J. R. Banga. An iterative identification procedure for dynamic modeling of biochemical networks. BMS Systems Biology. **4(11)**, (2010). https://doi.org/10.1186/1752-0509-4-11.
2. R. Bellman, K. J. Astrom. On structural Identifiability. Mathematical Biosciences, **7**, 329–339 (1970).
3. M.A. Boemo, H.M. Byrne. Mathematical modelling of a hypoxia-regulated oncolytic virus delivered by tumour-associated macrophages. J Theor Biol **461**, 102–116 (2019).
4. H. Byrne, L. Preziosi (2003). Modelling solid tumour growth using the theory of mixtures, Math Med Biol **20**, 341–366.
5. M. J. Chappell, K. R. Godfrey, S. Vajda. Global identifiability of the parameters of nonlinear systems with specified inputs: a comparison of methods. Mathematical Biosciences, **102**, 41–73 (1990).
6. O.T. Chis, J.R. Banga, E. Balsa-Canto. Structural identifiability of systems biology models: a critical comparison of methods. PLOS One. **6(11)**, 1–16 (2011). https://doi.org/10.1371/journal.pone.0027755.
7. O.T. Chis, J.R. Banga, E. Balsa-Canto. GenSSI: a software toolbox for structural identifiability analysis of biological models. Bioinformatics. **27(18)**, 2610–2611 (2011). https://doi.org/10.1093/bioinformatics/btr431.
8. J. Collis, A.J. Connor, M. Paczkowski, P. Kannan, J. Pitt-Francis, H.M. Byrne, M.E. Hubbard. Bayesian calibration, validation and uncertainty quantification for predictive modeling of tumor growth: a tutorial. Bull. Math. Biol. **79**(4), 939–974. (2017).
9. J.M.J. da Costa, H.R.B. Orlande, W.B. da Silaa. Model selection and parameter estimation in tumor growth models using approximate Bayesian computation – ABC. Comp. Appl. Math. **37**(3), 2795–2815. (2018).
10. H. Enderling, M.A.J. Chaplain, P. Hahnfeldt. Quantitative modeling of tumor dynamics and radiotherapy. Acta Biotheoretica. **58**(4), 341–353. (2010).
11. H. Haario, M. Laine, A. Mira, et al.: Efficient adaptive MCMC. Stat. Comput. **26**, 339–354 (2006).
12. E.J. Hall. Radiobiology for the radiologist. J.B. Lippincott, Philadelphia, 478–480 (1994).
13. N. Harald. Random number generation and quasi-Monte Carlo method. SIAM (1992).
14. P. Kannan, M. Paczkowski, A. Miar, et al.: Radiation resistant cancer cells enhance the survival and resistance of sensitive cells in prostate spheroids. bioRxiv (2019). https://doi.org/10.1101/564724.
15. J. Kursawe, R.E. Baker, A.G. Fletcher. Approximate Bayesian computation reveals the importance of repeated measurements for parameterising cell-based models of growing tissues. J. Theor. Biol. **443**, 66–81. (2018)
16. B. Lambert, A.L. MacLean, A.G. Fletcher, A.N. Combes, M.H. Little, H.M. Byrne. Bayesian inference of agent-based models: a tool for studying kidney branching morphogenesis. J. Math. Biol. **76**(7), 1673–1697. (2018).
17. D.E. Lea, D.G. Catcheside. The mechanism of the induction by radiation of chromosome aberrations in tradescantia. Journal of Genetics. **44**, 216–245 (1942).
18. T.D. Lewin. Modelling the impact of heterogeneity in tumor composition on the response to fractionated radiotherapy. D. Phil. Thesis, University of Oxford, 2018.

19. T.D. Lewin, H.M. Byrne, P.K. Maini, J.J. Caudell, E.G. Moros, H. Enderling. The importance of dead material within a tumour on the dynamics in response to radiotherapy. Physics in Medicine and Biology. https://doi.org/10.1088/1361-6560/ab4c27 (2019).
20. T.D. Lewin, P.K. Maini, E.G. Moros, H. Enderling, H.M. Byrne. A three-phase model to investigate the effects of dead material on the growth of avascular tumours. Mathematical Modelling of Natural Phenomena (in press) (2019).
21. E. Lima, J.T. Oden, D.A. Hormuth 2nd, T.E. Yankeelov, R.C. Almeida. Selection, calibration, and validation of models of tumor growth. Mathematical Models and Methods in Applied Sciences **26**(12), 2341–2368. (2016).
22. J.T. Oden, A. Hawkins, S. Prudhomme. General diffuse-interface theories and an approach to predictive tumour growth modelling. Math. Models Meth. Appl. Sci. **20** (3), 477–517 (2010).
23. H. Pohjanpalo. System identifiability based on the power series expansion of the solution. Mathematical Biosciences. **41**, 21–33 (1978).
24. B. Ribba, N.H. Holford, P. Magni, I Troconiz, I Gueorguieva, P. Girard, C. Sarr, M. Elishmereni, C. Kloft, L.E. Friberg. A review of mixed-effects models of tumor growth and effects of anti-cancer treatment used in population analysis. CPT Pharmacometrics Syst. Pharmacol. **3**, e113. (2014).
25. R.C. Smith. Uncertainty Quantification: Theory, Implementation, and Applications. SIAM Computational Science and Engineering Series (CS12). (2014).