

# ICES REPORT 17-14

---

June 2017

## Selection and Validation of Predictive Models of Radiation Effects on Tumor Growth Based on Noninvasive Imaging Data

by

E. A. B. F. Lima, J. T. Oden, B. Wohlmuth, A. Shahmoradi, D. A. Hormuth II, T. E. Yankeelov, L. Scarabosio, and T. Horger



**The Institute for Computational Engineering and Sciences**  
The University of Texas at Austin  
Austin, Texas 78712

*Reference: E. A. B. F. Lima, J. T. Oden, B. Wohlmuth, A. Shahmoradi, D. A. Hormuth II, T. E. Yankeelov, L. Scarabosio, and T. Horger, "Selection and Validation of Predictive Models of Radiation Effects on Tumor Growth Based on Noninvasive Imaging Data," ICES REPORT 17-14, The Institute for Computational Engineering and Sciences, The University of Texas at Austin, June 2017.*

# Selection and Validation of Predictive Models of Radiation Effects on Tumor Growth Based on Noninvasive Imaging Data

E. A. B. F. Lima<sup>a</sup>, J. T. Oden<sup>a,1,\*</sup>, B. Wohlmuth<sup>b</sup>, A. Shahmoradi<sup>a</sup>, D. A. Hormuth II<sup>a</sup>, T. E. Yankeelov<sup>a,c,d</sup>, L. Scarabosio<sup>b</sup>, T. Horger<sup>b</sup>

<sup>a</sup>*Institute for Computational Engineering and Sciences (ICES), The Center of Computational Oncology (CCO), The University of Texas at Austin*

<sup>b</sup>*Technical University of Munich, Germany, Department of Mathematics, Chair of Numerical Mathematics (M2)*

<sup>c</sup>*Department of Biomedical Engineering, The University of Texas at Austin*

<sup>d</sup>*Department of Internal Medicine, Livestrong Cancer Institutes, Dell Medical School, The University of Texas at Austin*

---

## Abstract

The use of mathematical and computational models for reliable predictions of tumor growth and decline in living organisms is one of the foremost challenges in modern predictive science, as it must cope with uncertainties in observational data, model selection, model parameters, and model inadequacy, all for very complex physical and biological systems.

In this paper, large classes of parametric models of tumor growth in vascular tissue are discussed including models for radiation therapy. Observational data is obtained from MRI of a murine model of glioma and observed over a period of about three weeks, with X-ray radiation administered 14.5 days into the experimental program. Parametric models of tumor proliferation and decline are presented based on the balance laws of continuum mixture theory, particularly mass balance, and from accepted biological hypotheses on tumor growth. Among these are new model classes that include characterizations of effects of radiation and simple models of mechanical deformation of tumors.

The Occam Plausibility Algorithm (OPAL) is implemented to provide a Bayesian statistical calibration of the model classes, 39 models in all, as well as the determination of the most plausible models in these classes relative to the observational data, and to assess model inadequacy through statistical validation processes. Discussions of the numerical analysis of finite element approximations of the system of stochastic, nonlinear partial differential equations characterizing the model classes, as well as the sampling algorithms for Monte Carlo and Markov chain Monte Carlo (MCMC) methods employed in solving the forward stochastic problem, and in computing posterior distributions of parameters and model plausibilities are provided. The results of the analyses described suggest that the general framework developed can provide a useful approach for predicting tumor growth and the effects of radiation.

**Keywords:** MRI imaging, radiotherapy, phenomenological models, calibration, validation, model plausibilities, Bayesian inference, Monte Carlo methods, statistical sampling.

---

## 1. Introduction

The development of computational models that reliably predict the growth or decline of tumors in living tissue is one of the great challenges in modern predictive science. It is a mission made difficult not only because of the considerable physical and biological complexity of heterogeneous interacting media in which tumors exist and evolve, but also because of the presence of uncertainties. Uncertainty in the observational data, in the selection of reliable models, in model parameters, in combination with errors in the computed approximations result in uncertainties in the predicted quantities of interest.

---

\*Corresponding author

*Email addresses:* lima@ices.utexas.edu (E. A. B. F. Lima), oden@ices.utexas.edu (J. T. Oden), wohlmuth@ma.tum.de (B. Wohlmuth), amir@ices.utexas.edu (A. Shahmoradi), david.hormuth@utexas.edu (D. A. Hormuth II), thomas.yankeelov@utexas.edu (T. E. Yankeelov), scarabos@ma.tum.de (L. Scarabosio), horger@ma.tum.de (T. Horger)

A comprehensive account of the mathematical foundations of cancer modeling and relevant literature is found in the articles of Bellomo et al. [4, 3]. More recently, the survey article of Oden et al. [35] discussed the emergence of predictive medicine, and, particularly, the statistical, mathematical, computational, and biological issues relevant in predictive multiscale modeling of tumor growth. Further accounts and additional references to contemporary literature can be found in the books of Cristini and Lowengrub [9] and Deisboeck and Stamatakos [11].

The terminology “predictive modeling” and “predictive science” refer to the appeal to all of the philosophical, logical, and scientific principles that underlie actual scientific predictions of realities and events in the physical universe. True predictive science must take into account all of the uncertainties mentioned above.

In Lima et al. [29], a general framework and process for selecting and validating predictive phenomenological models of tumor growth in living tissue was put forth in the form of an adaptive algorithm referred to as OPAL: the Occam Plausibility Algorithm, described in [16, 33]. The algorithm attempts to address and quantify uncertainties in model selection, observational data, and model parameters and to address model inadequacy and parameter sensitivity, while also quantifying uncertainties in outputs - the quantities of interest. Underlying OPAL are Bayesian methodologies: Bayesian methods of model selection based on posterior model plausibilities, and Bayesian statistical inverse methods for model calibration and validation.

The present investigation extends the work in [29] to the prediction of effects of X-ray radiation applied to a murine model of brain cancer. New models of the decline in tumor cell proliferation due to radiation are presented and calibrated against dynamic contrast-enhanced MRI (DCE-MRI) data. As the MRI data are provided in 2D tomographic slices, most of the simulations and model analyses is described for 2D cases, although 3D simulations of tumor growth and decline are also presented. In addition, summary discussions of key numerical methods are presented, including sampling algorithms generating posterior distribution, algorithms for Bayesian plausibility calculations and Monte Carlo solvers for solving stochastic forward problems.

The critical starting point for developing predictive computational models of tumor growth is the selection of a broad class  $\mathcal{M}$  of possible models. Following arguments in [21, 28, 32, 34], a natural starting point for developing plausible models for tumor growth involves, first, an appeal to fundamental balance laws of physics, represented by the phenomenological models of continuum mechanics, particularly the continuum theory of mixtures [7], and then proceeds to the consideration of mathematical characterizations of well-known biological events that promote or control the growth and decline of tumors (e.g., [18, 19]), all subject to constraints imposed by the second law of thermodynamics.

In this study, the physical region in which events of interest are to be modeled and from which observational (MRI) data are to be collected is the domain  $\Omega \subset \mathbb{R}^3$  representing the envelope of the brain of a murine subject in which a tumor mass, occupying a region  $\Omega_T(t) \subset \Omega$  at time  $t$ , evolves over a time period  $t \in [0, T]$ . The conservation of mass of the heterogeneous microenvironment of  $N$  interacting constituents (cell phenotypes, nutrients, extracellular matrix, etc.) is modeled as a continuum mixture by the system

$$\frac{\partial}{\partial t}(\rho_\alpha \phi_\alpha) = \rho_\alpha(S_\alpha(\boldsymbol{\phi}, \mathbf{u}) - \nabla \cdot \mathbf{J}_\alpha(\boldsymbol{\phi}, \mathbf{u})), \quad (1)$$

$1 \leq \alpha \leq N$ , and the mechanical behavior of the tumor is assumed to be captured by the quasi-static linear and angular momentum balances

$$\nabla \cdot \mathbf{T}(\boldsymbol{\phi}, \mathbf{u}) = \mathbf{0}, \quad \mathbf{T} = \mathbf{T}^t; \quad (2)$$

these balance laws holding at points  $(\mathbf{x}, t) \in \Omega_T(t) \times [0, T]$ . Here,

- $\rho_\alpha$  = mass per unit volume,
- $\phi_\alpha$  = volume fraction of constituent  $\alpha$ ,  $\boldsymbol{\phi} = \{\phi_1, \phi_2, \dots, \phi_N\}$ ,
- $S_\alpha$  = source term describing mass supplied to  $\alpha^{th}$  constituent by other constituents,
- $\mathbf{J}_\alpha$  = mass flux through the boundary of constituent  $\alpha$ ,
- $\nabla = \mathbf{e}_i \frac{\partial}{\partial x_i}$  = the spatial gradient (with Einstein’s convention on indices),

- $\mathbf{T}$  = the Cauchy stress tensor,
- $\mathbf{u}$  = the displacement field at  $\mathbf{x} \in \Omega_T(t)$  at time  $t$ .

In (1) and (2), convection and inertia terms do not appear, as they are dropped due to the slow (many days) evolution of tumor in living organisms.

Among dependent variables of interest are, for example,  $\phi_1 = \phi_T = \phi_T(\mathbf{x}, t)$ ,  $\mathbf{x} \in \bar{\Omega}$ ,  $t \in [0, \infty)$ , the volume fraction of tumor cells;  $\phi_P$ ,  $\phi_H$  and  $\phi_N$ , the volume fraction of proliferative, hypoxic and necrotic cells, respectively ( $\phi_T = \phi_P + \phi_H + \phi_N$ );  $\phi_\sigma$ , the nutrient concentration, etc. Key biological properties known to be relevant in the emergence of tumors and cancer are represented empirically in the mathematical structures of the source terms  $S_\alpha$  and the mass fluxes  $\mathbf{J}_\alpha$ , and are based on the highly-cited ‘‘Hallmarks of Cancer’’ of Hanahan and Weinberg [18, 19]. To close the models, constitutive equations must be added to characterize  $\mathbf{J}_\alpha$  and  $\mathbf{T}$ . These are discussed in Section 4, as well as boundary and initial conditions. Mechanical deformations and stress are captured by (2) to provide models of deformation-dependent mobilities and growth rates in mass balances of key constituents.

Proposals of mathematical models of the effects of X-ray radiation on living cells date back to 1924 with the creation of target theory by Crowther [10], and various generalizations of this theory appeared through the 1960’s (e.g., [13]). The popular so-called linear-quadratic (LQ) model [26] appeared in 1972 and is, to date, the most frequently used model in radiotherapy. A comprehensive historical and critical review of such models appeared in the recent paper of Bodgi et al. [5]. Traditionally, the effectiveness of cancer radiation therapy is depicted by a relation describing the survival fraction (the number of cells that are alive and can proliferate) versus the radiation dose, generally measured in units of Grays. In the present study, these survival-versus-dose models are used to develop mass sources or sinks in mass balance equations for tumor cell volume fraction (the  $S_\alpha$  functions in (1)). We consider, among the parametric model classes, those accounting for radiation modeled using the models described in Bodgi et al. [5], as well as new models designed to mimic observed effects in murine subjects used in this investigation.

The paper is organized as follows: after this introduction, a brief review of our Bayesian framework for model calibration, validation, and selection is given. This includes a discussion of nested sampling (or ‘‘Monte Carlo Lebesgue Integration’’) and Parallel Tempering algorithms for computing Bayesian model plausibilities, our principal tool for model selection for given observational data. Section 3 focuses on the acquisition and analysis of observational data - here MRI data on murine subjects. In particular, we review key aspects of DCE-MRI technology with the aim to characterize observational noise.

The development of several classes of parametric models of tumor growth is taken up in Section 4. These classes include families of reaction-diffusion models and phase-field models. An important aspect of this study is that appropriate models for predicting the evolution of tumors in living tissue are selected on the basis of Bayesian model plausibilities computed for fixed observational data (here furnished by MRI) and by well-defined model-validation processes.

In Section 5, various mathematical and numerical analysis issues are addressed, while Section 6 focuses on specific applications of the methodology to the prediction of glioma in murine subjects. Here the OPAL algorithm is briefly reviewed and used to calibrate, validate, and select models that best correspond to the observational (MRI) data. The effects of X-ray radiation on tumor growth are modeled using a characterization of the decline in tumor cell mass (volume fraction) due to radiation. The model selection, calibration, and validation processes are first implemented for data provided in 2D tomographic slices, but we go on to describe full 3D simulations. Results of several 3D simulations show good agreement with experimental data.

Major conclusions of the study are summarized in Section 7.

## 2. Bayesian Setting for Model Calibration, Validation, and Selection

To lay down notation and preliminaries to the analysis and developments described in later sections, we provide here a brief account of a framework for predictive modeling based on Bayesian inference and statistical inverse analysis that is drawn from earlier work (e.g., [35, 29, 16, 33, 21, 32, 34]). As noted in the introduction, the goal is to develop tools and algorithms for predictive modeling in the presence of uncertainties, including uncertainties in the observational data, in the selection of plausible models from a set  $\mathcal{M}$  of proposed models (large families of various reaction-diffusion and phase-field models with different characterizations of the effects of radiotherapy), model

parameters, which are generally represented by probability distributions, and target quantities of interest (QoIs). In the present work, the principal QoIs considered are the tumor volume (or area in 2D models), or equivalently the tumor mass, with model predictions of the volume / area at different sample times defining the parameter-to-observation correspondences.

### 2.1. Model Calibration and Validation

Each mathematical model  $\mathcal{M}_i$ , among a set  $\mathcal{M}$  of parametric model classes, can be best characterized as an abstract forward problem of the form,

$$\mathcal{M}_i : \mathcal{A}_i(\boldsymbol{\theta}_i, S; u_i(\boldsymbol{\theta}_i, S)) = 0, \quad 1 \leq i \leq m, \quad (3)$$

where the  $\mathcal{A}_i$  are operators (here generally partial differential operators),  $\boldsymbol{\theta}_i \in \boldsymbol{\Theta}_i$  is the vector of parameters for model  $\mathcal{M}_i$  with  $\boldsymbol{\Theta}_i$  being the corresponding parameter space of the model,  $u_i(\boldsymbol{\theta}_i, S)$  is the solution of the forward problem, and  $S$  is the scenario in which the model is applied (typically the domain of the solution  $u_i(\boldsymbol{\theta}_i, S)$  or the experimental set up designed to collect observational data).

The scenarios are separated into three categories: the calibration scenario  $S_c$  in which calibration data  $\mathbf{y}_c \in \mathcal{Y}_c$  is collected,  $\mathcal{Y}_c$  being the space of calibration observational data, the validation scenario  $S_v$  with validation data  $\mathbf{y}_v \in \mathcal{Y}_v$ , and the prediction scenario  $S_p$  in which the QoIs are evaluated. In our applications, the calibration process follows a variant of statistical cross validation in which training samples are taken over a specified time interval and model validation occurs at further times outside the training set.

The predictions of the calibration data  $\mathbf{y}_c$  by model  $\mathcal{M}_i$  of (3) are denoted  $\mathbf{d}^c(\boldsymbol{\theta}_i, S_c)$ . Based on the assumptions of a Gaussian noise model and independent and identically distributed (i.i.d.) samples of experimental data, we then construct the likelihood function  $\pi(\mathbf{y}_c | \boldsymbol{\theta}_i, \mathcal{M}_i)$ :

$$\ln(\pi(\mathbf{y}_c | \boldsymbol{\theta}_i, \mathcal{M}_i)) = \frac{N}{2} \ln(2\pi) + \sum_{j=1}^N \left[ -\ln(\sigma_j) - \frac{1}{2} \left( \frac{y_{cj} - d_j^c(\boldsymbol{\theta}_i)}{\sigma_j} \right)^2 \right], \quad (4)$$

where  $N$  is the number of data points (days) used for sampling calibration data,  $d_j^c(\boldsymbol{\theta}_i)$  is the tumor area (2D)/volume (3D) predicted by the model at day  $j$ , and  $\sigma_j^2$  is the estimated noise variance. Here by noise we mean both the measurement error in the data and the modeling error. Denoting the calibration prior by  $\pi(\boldsymbol{\theta}_i, \mathcal{M}_i)$ , the posterior distribution of the parameters for the calibration scenario is,

$$\pi(\boldsymbol{\theta}_i | \mathbf{y}_c, \mathcal{M}_i) = \frac{\pi(\mathbf{y}_c | \boldsymbol{\theta}_i, \mathcal{M}_i) \pi(\boldsymbol{\theta}_i, \mathcal{M}_i)}{\pi(\mathbf{y}_c, \mathcal{M}_i)}, \quad 1 \leq i \leq m. \quad (5)$$

These posterior densities are used as priors in the validation process, yielding new updated validation posterior densities,

$$\pi(\boldsymbol{\theta}_i | \mathbf{y}_v, \mathbf{y}_c, \mathcal{M}_i) = \frac{\pi(\mathbf{y}_v | \boldsymbol{\theta}_i, \mathcal{M}_i) \pi(\boldsymbol{\theta}_i | \mathbf{y}_c, \mathcal{M}_i)}{\pi(\mathcal{M}_i, \mathbf{y}_v, \mathbf{y}_c)}. \quad (6)$$

The validation predictions of the model, with parameters defining the posterior (6), of the data  $\mathbf{y}_v$ , are denoted  $\mathbf{d}^v(\boldsymbol{\theta}_i, S_v)$ . The validity of model  $\mathcal{M}_i$  is then determined by the accuracy of these predictions, computed using a pre-specified metric  $\mathfrak{d}(\cdot, \cdot)$  and an error tolerance  $\gamma_{tol}$ :

$$\mathfrak{d}(\mathbf{y}_v, \mathbf{d}^v(\boldsymbol{\theta}_i, S_v)) \begin{cases} \leq \gamma_{tol} & \text{model is not-invalid,} \\ > \gamma_{tol} & \text{model is invalid.} \end{cases} \quad (7)$$

We discuss such metrics and tolerance selection in Section 5. Once a model has been declared ‘‘valid’’ (not-invalid, according to (7)), the forward problem (3) is solved in the prediction scenario  $S_p$ , and the QoIs are evaluated. In the present investigation, Monte Carlo methods are employed to solve the stochastic forward problem. As mentioned earlier, a typical QoI is the tumor area or volume.

## 2.2. Model Selection: Numerical Methods for Calculating Bayesian Model Plausibilities

A major challenge in modeling the evolution of tumor cells in living tissue is the selection of an appropriate mathematical model that best complies with the observational data in hand, independently of the model parameters (which are fit to data in the calibration and validation processes). Here we employ the theory of Bayesian model plausibility, as a powerful approach to model selection, and we describe the algorithm for calculating such plausibility measures.

For a given model  $\mathcal{M}$  and data  $\mathbf{y}$ , the posterior distribution of the parameters can be written as:

$$\pi(\boldsymbol{\theta}|\mathbf{y}, \mathcal{M}) = \frac{\pi(\mathbf{y}|\boldsymbol{\theta}, \mathcal{M}) \pi(\boldsymbol{\theta}, \mathcal{M})}{\pi(\mathbf{y}, \mathcal{M})}. \quad (8)$$

Let us suppose that we have access to a set of  $m$  competing parametric model classes  $\mathcal{M} = \{\mathcal{M}_1, \dots, \mathcal{M}_m\}$ , each of which is capable of describing data  $\mathbf{y}$ . For each model  $\mathcal{M}_i$  in the set of models, one can integrate the Bayes rule (8) over the entire parameter space  $\Theta_i$  of the model and rearrange the equation to obtain

$$\int_{\Theta_i} \pi(\boldsymbol{\theta}|\mathbf{y}, \mathcal{M}_i) d\boldsymbol{\theta} = \int_{\Theta_i} \frac{\pi(\mathbf{y}|\boldsymbol{\theta}, \mathcal{M}_i) \pi(\boldsymbol{\theta}, \mathcal{M}_i)}{\pi(\mathbf{y}, \mathcal{M}_i)} d\boldsymbol{\theta} = 1, \quad (9)$$

so that

$$\pi(\mathbf{y}, \mathcal{M}_i) = \int_{\Theta_i} \pi(\mathbf{y}|\boldsymbol{\theta}, \mathcal{M}_i) \pi(\boldsymbol{\theta}, \mathcal{M}_i) d\boldsymbol{\theta}. \quad (10)$$

Equation (10) provides a method of calculating the denominator of the Bayes rule in (8), which can be interpreted as the probability of observing data  $\mathbf{y}$  and that these data come from the model  $\mathcal{M}_i$ . It is called *marginal likelihood* since it is calculated by marginalizing the likelihood function times prior over the entire parameter space of the model  $\mathcal{M}_i$ ; it is also called the *Bayesian evidence* or *model evidence* or simply the *evidence* [16].

As a marginalization over parameters  $\boldsymbol{\theta}_i$ , parameter dependence is averaged out and does not appear in the arguments of the model evidence. The utility of evidence goes beyond just serving as a normalizing constant in the Bayes rule, as it can be used for the calculation of *Bayesian plausibilities*. The latter follow from a second application of the Bayes theorem, this time for the set of models  $\mathcal{M}$ :

$$\pi(\mathcal{M}_i|\mathbf{y}, \mathcal{M}) = \frac{\pi(\mathbf{y}|\mathcal{M}_i, \mathcal{M}) \pi(\mathcal{M}_i, \mathcal{M})}{\pi(\mathbf{y}, \mathcal{M})}. \quad (11)$$

Equation (11) gives the posterior probability density of the  $i^{\text{th}}$  model  $\mathcal{M}_i$  in the set of all rival models  $\mathcal{M}$ . It is called a Bayesian plausibility, since it provides a measure of the plausibility of model assumptions in the light of available data and prior knowledge about the models in  $\mathcal{M}$ . In the case of complete prior ignorance about all competing models, the Jaynes' principle of maximum entropy [25] dictates the assignment of uniform equal prior probabilities to each of the models. Clearly, denoting by  $\omega_i$  the model plausibility,

$$\omega_i = \pi(\mathcal{M}_i|\mathbf{y}, \mathcal{M}), \quad \sum_{i=1}^m \omega_i = 1. \quad (12)$$

The model(s) with the largest value of  $\omega_i$  are deemed the most plausible (the "best") among the set of models  $\mathcal{M}$ .

## 3. Observational Data: Magnetic Resonance Imaging in a Murine Model of Glioma

Biomedical imaging, in general, and magnetic resonance imaging (MRI), in particular, has matured to the point where it offers a wealth of information on tumor status [44, 36]. Examples of the use of imaging data to ascertain the predictivity of tumor growth models can be found in [2, 1, 29, 45, 46, 43, 42, 41]. In this study, we make use of contrast-enhanced MRI to identify the boundaries of the tumor. We note that some portions of this data have already been reported in the literature [22, 23, 24]. These data are then used to initialize, constrain, and predict the

spatio-temporal evolution of tumor growth and response to treatment. More specifically, the time evolution is divided into three sets of data. The first several (three or four, depending on the experimental group) time points are used to calibrate the model and update the prior. With the updated prior, we use the subsequent time points to validate the model and further update the prior. The computed posterior is used to compute the model prediction and is compared with the last point drawn from the experimental data.

### 3.1. Murine Glioma Data

Rats were anesthetized with 2% isoflurane in 98% oxygen for all imaging, surgical, and irradiation procedures. Three female Wistar (245 to 300g) rats were anesthetized and inoculated intracranially with C6 glioma cells ( $1 \times 10^5$ ) via stereotaxic injection on day 0. Permanent jugular catheters were placed on day 8 for the injection of an MRI contrast agent. The control rat was imaged on days 10, 12, 14, 15, 16, 18 and 20, while the treated rats were imaged on days 10, 12, 14, 16.5, 18.5, 20.5, and 22.5. The treated rats each received either 20 or 40 Gy with a Therapax DXT 300 X-ray machine (Pantak Inc., East Haven, CT, USA) on day 14.5. During the irradiation protocol, lead blocks were placed around the rat head to minimize radiation exposure outside of the brain.

### 3.2. Contrast Enhanced MRI

MRI acquisitions employed a 9.4T 21 cm bore spectrometer (Agilent, Santa Clara, CA, USA) and a 38 mm volume coil (Doty Scientific, Columbia, SC, USA) at the Vanderbilt University Institute of Imaging Science. MR images were acquired over a  $32 \times 32 \times 16 \text{ mm}^3$  field of view sampled with a  $96 \times 96 \times 16$  matrix (zero padded to a  $128 \times 128 \times 16$  matrix). At the beginning of the second through the final imaging session, a mutual information based rigid registration algorithm [22, 8, 30] was used at the scanner to register the current imaging time point to the initial imaging session.

To obtain the contrast enhanced MRI data, a spoiled gradient echo sequence was employed, with  $TR/TE = 47/1.4$  ms, two excitations, and a flip angle  $= 20^\circ$  was used for the dynamic portion of the experiment yielding a temporal resolution of 9s. These dynamic images were collected before and after the injection of a  $200 \mu\text{L}$  bolus ( $0.05 \text{ mmol kg}^{-1}$ ) of gadolinium-diethylenetriamine pentaacetic acid (Gado-DTPA<sup>TM</sup>, BioPhysics Assay Laboratory, Worcester, MA) through an indwelling jugular catheter. These DCE-MRI data are then used to identify the boundaries of the tumor.

The contrast agent spreads throughout the entire network of blood vessels, and starts extravasating through vessel walls into the extravascular-extracellular environment in the body. This extravasation is much more enhanced in vessels with leaky walls. Since vascularization of tumors is often very leaky and immature, the amount of contrast-agent extravasation is much greater in the areas surrounding the tumor. Therefore, the signal intensity within and adjacent to the tumor area is significantly enhanced in DCE-MRI images. This enhancement can be used to define the boundary of tumor versus healthy tissue surrounding it, as illustrated in Figure 1.

## 4. Parametric Classes of Models of Tumor Growth

A fundamental aspect of this study is that our approach is “model agnostic” - in the sense that we ultimately choose models among large classes of parametric models which best fit observational data according to their Bayesian plausibilities. In this section, we present several large classes of tissue-scale models of tumor growth. These classes can be generated from the general mass balance (1), once models of the mass source terms  $S_\alpha$  and constitutive equations for the mass flux  $\mathbf{J}_\alpha$  are specified.

### 4.1. Mass Flux $\mathbf{J}_\alpha$

It is shown in [32] that mass flux appears in the Clausius-Duhem inequality, derived from the second law of thermodynamics, as conjugate to the gradient of the chemical potential,  $\mu_\alpha$ , leading to the scalar product  $\mathbf{J}_\alpha \cdot \nabla \mu_\alpha$ . For models with  $N$  constituents  $\phi_\alpha$ ,  $1 \leq \alpha \leq N$ , this term is consistent with the entropy (dissipation) inequality if we choose a mass flux linear in the gradient of the chemical potential:

$$\mathbf{J}_\alpha = -\mathbf{M}_\alpha(\boldsymbol{\phi})\nabla\mu_\alpha, \quad 1 \leq \alpha \leq N, \quad (13)$$

where  $\mathbf{M}_\alpha(\boldsymbol{\phi})$  is the positive semi-definite mobility matrix, which may depend upon the volume fraction vector,  $\boldsymbol{\phi} = (\phi_1, \phi_2, \dots, \phi_N)$ .

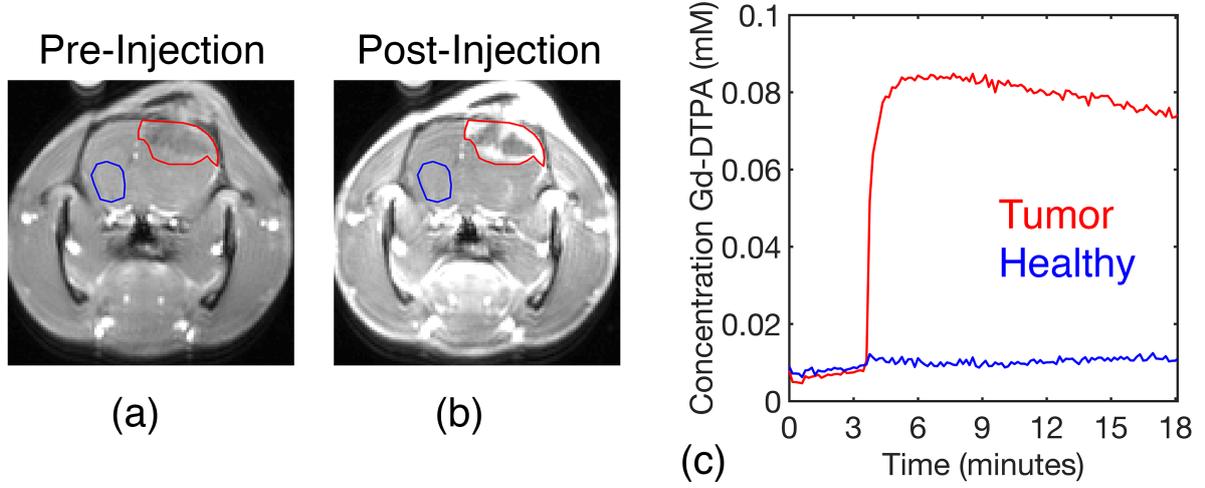


Figure 1: An illustration of the  $T_1$ -weighted signal enhancement in DCE-MRI. Plot (a) illustrates the luminosity of a healthy brain region (blue) and tumor region (red) before the injection of contrast agent. Plot (b) is the same image, but taken with 5 minutes delay after the contrast agent injection. The signal enhancement is clearly visible in and around the tumor area. Plot (c) shows the temporal evolution Gd-DTPA concentration after the injection of the contrast agent. This is an example of a contrast agent that enhances the signal in tumor region by changing the  $T_1$  relaxation time of the tumor due to high concentration of the agent in and around the tumor.

#### 4.2. Mass Sources $S_\alpha$

In developing mathematical representations of the mass source terms  $S_\alpha$  in (1), one begins by identifying the constituent volume fraction,  $\phi_\alpha$ , with key components of the tissue morphology and by developing representations of what is perceived to be of their interactions based on accepted explanations of the biological factors affecting the growth and decline of cancer in living tissue. Examples of key constituents of interest are the following volume fraction constituents:

- $\phi_C \sim$  healthy cells,
- $\phi_T \sim$  tumor cells.

The mixture is said to be saturated if  $\sum_\alpha \phi_\alpha(\mathbf{x}, t) = 1$ .

To express the mass sources  $S_\alpha$  in the terms of the volume fractions  $\phi_\alpha$ , we call upon the Hanahan and Weinberg [18, 19] hallmarks of cancer and design mathematical switches and amplification terms to represent these cancer-promoting factors. For example, ignoring convection due to the slow (days or months) progression of tumor growth and setting the mass densities  $\rho_\alpha$  to be constant (since the mass densities of cellular constituents are generally all close to that of water), we have, for the change in the mass of tumor cells, the equation,

$$\frac{\partial \phi_T}{\partial t} = -\nabla \cdot \mathbf{J}_T + S_T, \quad (14)$$

and one can take

$$S_T = \lambda_T^g \phi_T (1 - \phi_T) - \mathcal{R}(T_{treat}, \tilde{\mathcal{D}}_t) \phi_T, \quad (15)$$

where  $\lambda_T^g$ ,  $T_{treat}$  and  $\tilde{\mathcal{D}}_t$  are positive parameters (constants for the moment, but ultimately probability densities) and  $\mathcal{R}(T_{treat}, \tilde{\mathcal{D}}_t)$  the radiotherapy function. In (15), the first term on the right-hand side describes a logistic increase in tumor cell concentration at a growth rate  $\lambda_T^g$ ; the second term is designed to mimic tumor cell decline within the tumor volume due to the effects of the radiation therapy at a rate of  $\mathcal{R}(T_{treat}, \tilde{\mathcal{D}}_t)$ , where  $T_{treat}$  is the time that the radiation dose  $\tilde{\mathcal{D}}_t$  is applied. The radiotherapy models used are described in subsections 4.2.1, 4.2.2 and 4.2.3 below.

#### 4.2.1. Memory Model (T01)

A new model is proposed that is designed to capture memory effects, in which the death of the tumor cells propagates for a period after the treatment. By this is meant that the effects of the treatment do not diminish immediately after the treatment. Upon the application of X-ray radiation, the tumor mass decreases over a period of time and then begins to increase as growth renews. Assuming that  $\lambda_k(\tilde{\mathcal{D}}_t)$  represents the death effect by the dose of radiation applied and that  $\lambda_r$  is the rate of recovery from the treatment, the following model is proposed,

$$\mathcal{R}(T_{treat}, \tilde{\mathcal{D}}_t) = \begin{cases} \frac{\lambda_k(\tilde{\mathcal{D}}_t)}{1 + \lambda_r(t - T_{treat})}, & \text{for } t \geq T_{treat}, \\ 0, & \text{for } t < T_{treat}, \end{cases} \quad (16)$$

where  $\lambda_r(t - T_{treat})$  indicates that the time to kill tumor cells increases as the time after treatment increases. The model described by (16) reflects that the likelihood to kill tumor cells decreases with the time elapsed after the day of the treatment and that the influence of the treatment vanishes in time if the tumor is not killed completely.

#### 4.2.2. Partial Memory Model (T02)

A family of radiotherapy models is described in [24] which is based on two assumptions: 1) the effects of radiotherapy are instantaneous, without a memory effect on the death of the tumor by radiation; 2) there is a reduction in the net proliferation rate after the treatment. Under these hypotheses, some tumor cells are killed the moment that the treatment is applied; however, a memory effect holds in the surviving tumor with a decrease of the tumor growth rate. The death of the tumor cells by the radiotherapy is modeled as

$$\mathcal{R}(T_{treat}, \tilde{\mathcal{D}}_t) = \begin{cases} \bar{\alpha}_{\mathcal{D}}(\tilde{\mathcal{D}}_t), & \text{for } t = T_{treat}, \\ 0, & \text{for } t \neq T_{treat}, \end{cases} \quad (17)$$

where  $\bar{\alpha}_{\mathcal{D}}(\tilde{\mathcal{D}}_t)$  is the rate of tumor cell death as a function of radiation dose. In this work it is assumed that  $\bar{\alpha}_{\mathcal{D}}(\tilde{\mathcal{D}}_t) = \alpha_{\mathcal{D}}$ , with  $\alpha_{\mathcal{D}}$  being a positive constant to be calibrated. The reduction in the growth rate is given by

$$\lambda_T^g(t) = \begin{cases} \lambda_T^g R^g, & \text{for } t \geq T_{treat}, \\ \lambda_T^g, & \text{for } t < T_{treat}, \end{cases} \quad (18)$$

where  $0 \leq R^g \leq 1$  is the radiotherapy effect in growth rate, and  $\lambda_T^g$  on the right-hand side is the same parameter as in (15).

#### 4.2.3. No Memory Model (T03)

One class of models is based on the assumption that the effects of the treatment are instantaneous, killing some cells at the moment that the treatment is applied, without any lasting effects on the tumor growth rate. This approach can be modeled by the following function:

$$\mathcal{R}(T_{treat}, \tilde{\mathcal{D}}_t) = \begin{cases} \kappa(1 - \mathbb{S}(\tilde{\mathcal{D}}_t)), & \text{for } t = T_{treat}, \\ 0, & \text{for } t \neq T_{treat}, \end{cases} \quad (19)$$

where  $\kappa$  is a positive parameter inserted to comply dimensionally. In (19),  $\mathbb{S}(\tilde{\mathcal{D}}_t)$  is the survival fraction. Several options for representing the survival fraction  $\mathbb{S}(\tilde{\mathcal{D}}_t)$  are compared in Bodgi et al. [5]. However, as our experimental data is obtained for a single treatment application and single dose fraction (20Gy or 40Gy), calibration of the different models would involve only a single dose to calibrate, suggesting that all of the models are equivalent.

In the present work, the classical linear-quadratic (LQ) survival-versus-dose model [12, 26] is selected, as it is regarded as the ‘‘gold standard’’ in the field [39, 6]. The LQ model is given by

$$\mathbb{S}(\tilde{\mathcal{D}}_t) = \exp(-\alpha\tilde{\mathcal{D}}_t - \beta\tilde{\mathcal{D}}_t^2), \quad (20)$$

where  $\alpha$  and  $\beta$  are positive parameters to be calibrated.

Table 1 shows the three radiotherapy models presented and their parameters.

Table 1: Number of parameters for the three different models described in subsections 4.2.1–4.2.3.

Model	Parameters ( $\theta$ )							# $\theta$
	$\lambda_r$	$\lambda_k$	$\alpha$	$\beta$	$\kappa$	$\alpha_{\mathcal{D}}$	$R^g$	
T01	✓	✓						2
T02						✓	✓	
T03			✓	✓	✓			3

### 4.3. Mechanical Effects

To account for mechanical effects, such as compression of the tumor mass, which can strongly influence mobility and diffusion rates, we introduce the linear and angular momentum balance (2). Rather than endow each constituent with independent mechanical fields (stress, strain, displacement), we consider models of the total tumor domain  $\Omega_T(t)$ . Following Gurtin [17], we consider quasi-static, infinitesimal, elastic deformations of tissues characterized by free energies of the form

$$\tilde{\Psi} = \int_{\Omega_T} [\Psi(\phi_T, \nabla\phi_T) + W(\phi_T, \boldsymbol{\varepsilon}(\mathbf{u}))] dx, \quad (21)$$

where the tumor volume fraction  $\phi_T$  acts as an order parameter,  $\phi_T = \phi_T(\mathbf{x}, t)$ ,  $\Psi(\phi_T, \nabla\phi_T)$  is the Ginzburg-Landau free energy, and  $W(\phi_T, \boldsymbol{\varepsilon}(\mathbf{u}))$  is the free energy of elastic deformation,  $\boldsymbol{\varepsilon}(\mathbf{u})$  being the strain tensor associated with the displacement field  $\mathbf{u}$ . The Ginzburg-Landau free energy is given by

$$\Psi(\phi_T, \nabla\phi_T) = \hat{\Psi}(\phi_T) + \frac{\epsilon_T^2}{2} |\nabla\phi_T|^2. \quad (22)$$

Here  $\hat{\Psi}(\phi_T)$  describes a double-well potential with wells that define phases in the media, and  $\epsilon_T \geq 0$  is a surface-energy parameter weighing gradients in the constituent  $\phi_T$ . In (21),  $W(\phi_T, \boldsymbol{\varepsilon}(\mathbf{u}))$  is defined as

$$W(\phi_T, \boldsymbol{\varepsilon}(\mathbf{u})) = \frac{1}{2} \boldsymbol{\varepsilon}(\mathbf{u}) : \mathbf{C}(\phi_T) \boldsymbol{\varepsilon}(\mathbf{u}) + \boldsymbol{\varepsilon}(\mathbf{u}) : \bar{\mathbf{T}}(\phi_T), \quad (23)$$

where the stored energy is characterized as a quadratic function of strains,

$$\boldsymbol{\varepsilon}(\mathbf{u}) = \frac{1}{2} (\nabla\mathbf{u} + \nabla\mathbf{u}^t), \quad (24)$$

$\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$ ,  $\mathbf{x} \in \Omega_T(t)$ , being the displacement field,  $\mathbf{C}(\phi_T)$  is the symmetric fourth-order elasticity tensor, and  $\bar{\mathbf{T}}(\phi_T)$  is the symmetric compositional stress tensor [17]. This form of the free energy leads to the following constitutive equations for stress  $\mathbf{T}$  and chemical potential  $\mu$ :

$$\left. \begin{aligned} \mathbf{T} &= \frac{\partial W(\phi_T, \boldsymbol{\varepsilon}(\mathbf{u}))}{\partial \boldsymbol{\varepsilon}} = \mathbf{C}(\phi_T) \boldsymbol{\varepsilon}(\mathbf{u}) + \bar{\mathbf{T}}(\phi_T), \\ \mu &= \frac{\partial \tilde{\Psi}(\phi_T, \nabla\phi_T, \boldsymbol{\varepsilon}(\mathbf{u}))}{\partial \phi_T} = \frac{\partial \hat{\Psi}(\phi_T)}{\partial \phi_T} + \frac{1}{2} \boldsymbol{\varepsilon}(\mathbf{u}) : \mathbf{C}'(\phi_T) \boldsymbol{\varepsilon}(\mathbf{u}) + \boldsymbol{\varepsilon}(\mathbf{u}) : \bar{\mathbf{T}}'(\phi_T) - \epsilon_T^2 \Delta\phi_T, \end{aligned} \right\} \quad (25)$$

with  $\mathbf{C}'(\phi_T) = \partial\mathbf{C}(\phi_T)/\partial\phi_T$ ,  $\bar{\mathbf{T}}'(\phi_T) = \partial\bar{\mathbf{T}}(\phi_T)/\partial\phi_T$  and  $\Delta$  denotes the spatial Laplacian.

The quasi-static approximation of the moment balance is thus given by

$$\nabla \cdot (\mathbf{C}(\phi_T) \boldsymbol{\varepsilon}(\mathbf{u}) + \bar{\mathbf{T}}(\phi_T)) = 0. \quad (26)$$

For the avascular tumor evolution considered here, a linear isotropic compositional stress is assumed of the form  $\bar{\mathbf{T}}(\phi_T) = \lambda\phi_T\mathbf{I}$ , where  $\mathbf{I}$  is the identity tensor and  $\lambda$  depends on the tumor growth rate. Thus, rewriting (26) in terms of the displacement vector  $\mathbf{u}$ , we obtain the mechanical equilibrium equation:

$$\frac{1}{2} \nabla \cdot [\mathbf{C}(\phi_T) (\nabla\mathbf{u} + \nabla\mathbf{u}^t)] = -\lambda \nabla\phi_T, \quad (27)$$

describing quasi-static deformations of the tumor.

A major simplification occurs if we assume that the material is isotropic and homogeneous, in which case Hooke's law holds with the addition of a source term proportional to  $\nabla\phi_T$ . For the plane-strain case,

$$0 = \nabla \cdot G\nabla\mathbf{u} + \nabla \frac{G}{1-2\nu} (\nabla \cdot \mathbf{u}) + \lambda\nabla\phi_T. \quad (28)$$

In (28),  $\nu$  is Poisson's ratio and  $G = \frac{E}{2(1+\nu)}$  is the shear modulus, with  $E$  Young's modulus. The various forms of the free energy density  $\Psi(\phi_T, \nabla\phi_T)$  yield different tumor growth models considered here. As a further simplification, we assume constant mechanical properties throughout the whole domain  $\Omega_T(t)$  [29].

A model of the effect of mechanical deformation on tumor mobility and growth is proposed in [22], and is given by

$$M_T := \exp(-\gamma\Sigma)\overline{M_T}, \quad \lambda_T^g := \exp(-\gamma^g\Sigma)\overline{\lambda_T^g}, \quad (29)$$

where  $\Sigma$  is the Von Mises stress and  $\overline{M_T}$  and  $\overline{\lambda_T^g}$  are the latent values of mobility and growth rate. In order to simplify the notation, we drop hereafter the over bar from the constants  $\overline{M_T}$  and  $\overline{\lambda_T^g}$ .

#### 4.4. Model Classes

The choice of the free energy  $\tilde{\Psi}$  leads to the three different base model classes described more fully in following subsections: Phase-Field (PF) models, Reaction-Diffusion models with the mechanical term (23) included in (21) (MD) and without the mechanical term included in (21) (RD).

##### 4.4.1. Phase-Field Model Classes

For phase-field models, the free energy is given by (21), with the the Ginzburg-Landau free energy (22). In (22) the double-well potential is given by

$$\tilde{\Psi}(\phi_T) = \bar{E}_T\phi_T^2(1 - \phi_T)^2, \quad (30)$$

where  $\bar{E}_T > 0$  is the energy scale associated with constituent  $\phi_T$ . The chemical potential in (25) is

$$\mu = -\epsilon_T^2\Delta\phi_T + \frac{\partial\tilde{\Psi}(\phi_T)}{\partial\phi_T} + \lambda\nabla \cdot \mathbf{u}. \quad (31)$$

We add to this system the following boundary conditions and initial conditions:

$$\nabla\phi_T \cdot \mathbf{n} = 0, \quad \nabla\mu \cdot \mathbf{n} = 0, \quad \mathbf{u} = \mathbf{0}, \quad \text{on } \partial\Omega \times (0, T), \quad (32)$$

$$\phi_T(\mathbf{x}, 0) = \phi_{T0}(\mathbf{x}), \quad \text{on } \Omega, \quad (33)$$

$\forall \mathbf{x} \in \Omega$ ,  $\mathbf{n}$  being normal to  $\partial\Omega$ . By definition of volume fraction,  $\phi_{T0}(\mathbf{x}) \geq 0$ ,  $\forall \mathbf{x} \in \Omega$ . From (13), (14), (21), (30) and (31), the phase-field model is given by

$$\left. \begin{aligned} \frac{\partial\phi_T}{\partial t} &= \nabla \cdot M_T(\phi_T)\nabla\mu + S_T, \\ \mu &= 2\bar{E}_T\phi_T(1 - 3\phi_T + 2\phi_T^2) - \epsilon_T^2\Delta\phi_T + \lambda\nabla \cdot \mathbf{u}, \end{aligned} \right\} \text{in } \Omega \times (0, T), \quad (34)$$

to be completed with equation (28) for  $\mathbf{u}$ .

##### 4.4.2. Reaction-Diffusion Model Classes

The volume fraction mass transfer equations (e.g., (1)) reduce to nonlinear reaction diffusion equations when the elastic energy is disregarded and the Ginzburg-Landau free energy is quadratic in the field  $\phi_T$ . For this simple model,

$$\tilde{\Psi} = \int_{\Omega} \left[ \frac{c}{2}\phi_T^2 \right] dx, \quad (35)$$

$$\mu = c\phi_T, \quad (36)$$

where  $c$  is a positive constant. Otherwise, with the mechanical term (23) included in (21), and (35) in force,

$$\mu = c\phi_T + \lambda \nabla \cdot \mathbf{u}. \quad (37)$$

From (13) and (14), the reaction-diffusion model is given by

$$\frac{\partial \phi_T}{\partial t} = \nabla \cdot M_T(\phi_T) \nabla \mu + S_T, \quad \text{in } \Omega \times (0, T), \quad (38)$$

where the chemical potential is given by (36) for the RD class (for which we define  $M_T^* = cM_T$ ) or (37) for the MD class. Table 2 presents the 13 models that can be derived from the three different base model classes obtained so far [29].

Table 2: Models derived from the reaction-diffusion (RD) (equations (38) and (36)), reaction-diffusion with the mechanical term included in the free-energy (MD) (equations (38) and (37)) and phase-field (PF) (equation (34)).

Model	Variables			Parameters ( $\theta$ )										# $\theta$	
	$\phi_T$	$\mu$	$\mathbf{u}$	$M_T$	$M_T^*$	$c$	$\lambda_T^g$	$\bar{E}_T$	$\epsilon_T$	$G$	$\nu$	$\lambda$	$\gamma$		$\gamma^g$
RD01	✓				✓		✓								2
PF01	✓	✓		✓			✓	✓	✓						4
RD02	✓		✓		✓		✓			✓	✓	✓	✓		6
RD03	✓		✓		✓		✓			✓	✓	✓		✓	
MD01	✓	✓	✓	✓		✓	✓			✓	✓	✓			
RD04	✓		✓		✓		✓			✓	✓	✓	✓	✓	7
MD02	✓	✓	✓	✓		✓	✓			✓	✓	✓	✓		
MD03	✓	✓	✓	✓		✓	✓			✓	✓	✓		✓	
PF02	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓			8
MD04	✓	✓	✓	✓		✓	✓			✓	✓	✓	✓	✓	
PF03	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓		
PF04	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓		✓	9
PF05	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓	

Combining the three treatment models presented in Table 1, with the thirteen tumor growth models presented in Table 2, we obtain 39 tumor growth models with radiotherapy treatment. These models are divided into 9 Occam categories according to the number of parameters in each model. Table 3 presents the 39 models, where the names of the models are a combination of the names presented in Table 2 (e.g., RD01) with the treatment model used (e.g., T01 = memory model, T02 = partial memory model and T03 = no memory model).

## 5. Algorithms and Numerical Methods

In this section, we provide a brief account of the principal computational algorithms, numerical methods, and software frameworks used in the tumor growth predictions taken up in the next section.

### 5.1. Mixed Finite Element Approximations

The general purpose finite-element library `libMesh` [27], developed at ICES (the Institute for Computational Engineering and Sciences) at UT Austin, is used to construct numerical approximations of all of the model classes described up to this point. A typical example of a phase-field model (model PF01 from Table 2) is provided by the tumor volume fraction mass balance,

$$\left. \begin{aligned} \frac{\partial \phi_T}{\partial t} &= \nabla \cdot M_T(\phi_T) \nabla \mu + \lambda_T^g \phi_T (1 - \phi_T), \\ \mu &= 2\bar{E}\phi_T (1 - 3\phi_T + 2\phi_T^2) - \epsilon_T^2 \Delta \phi_T, \end{aligned} \right\} \text{in } \Omega \times (0, T), \quad (39)$$



$L^2(0, T; \mathcal{V}) \times L^2(0, T; \mathcal{V})$ , with  $\frac{\partial \phi_T}{\partial t} \in \mathcal{V}'$  (the dual space of  $\mathcal{V}$ ), such that for all  $t \in J$ ,

$$\left. \begin{aligned} \left( \frac{\partial \phi_T}{\partial t}, \widehat{\phi}_T \right) &= \left( \lambda_T^g \phi_T (1 - \phi_T), \widehat{\phi}_T \right) - \left( M_T(\phi_T) \nabla \mu, \nabla \widehat{\phi}_T \right), \quad \forall \widehat{\phi}_T \in \mathcal{V}, \\ (\mu, \widehat{\mu}) &= \left( 2\bar{E} \phi_T (1 - 3\phi_T + 2\phi_T^2), \widehat{\mu} \right) + \left( \epsilon_T^2 \nabla \phi_T, \nabla \widehat{\mu} \right), \quad \forall \widehat{\mu} \in \mathcal{V}. \end{aligned} \right\} \quad (40)$$

To define the finite element approximation to (40), we introduce the family of subspaces

$$\mathcal{V}^h = \{v \in \mathcal{V} : v|_\tau = \hat{v} \circ F_\tau, \hat{v} \in \mathbb{Q}_1, \tau \in \mathcal{T}^h\}, \quad (41)$$

where  $\mathcal{T}^h$  is a quasi-uniform family of triangulations of  $\Omega$ ,  $F_\tau$  is an affine map from the master element  $\hat{\tau}$  into  $\tau$ , and  $\mathbb{Q}_1$  is the tensor product of polynomials of degree 1.

We employ a discrete-time scheme based on the energy convex-nonconvex splitting proposed by Eyre [15], where the energy functional can be split into contractive and expansive terms. The contractive term is treated implicitly and the expansive term is treated explicitly. Such splittings always exist for each form of the energy functional, although there is not a unique way to split the energy [15].

The time domain is divided into  $n$  time steps,  $\Delta t_n = t_{n+1} - t_n$ , assumed constant for simplicity ( $\Delta t_n = \Delta t$ ,  $n = 0, 1, \dots$ ). The nonlinear term can be approximated by using the Picard (fixed-point) iteration method, where  $k$  is the iteration index. The fully discrete solution for the tumor growth model is obtained by using finite difference approximations of the time derivatives. Then, the fully discrete variational form is given by: Find  $\phi_{T_{n+1}}^{k+1}$  and  $\mu_{n+1}^{k+1} \in \mathcal{V}^h$ , such that:

$$\left. \begin{aligned} \left( \phi_{T_{n+1}}^{k+1} - \phi_{T_n}, \widehat{\phi}_T \right) - \Delta t \left( \lambda_T^g \phi_{T_{n+1}}^{k+1} (1 - \phi_{T_{n+1}}^k), \widehat{\phi}_T \right) + \Delta t \left( M_T(\phi_{T_{n+1}}^k) \nabla \mu_{n+1}^{k+1}, \nabla \widehat{\phi}_T \right) &= 0, \quad \forall \widehat{\phi}_T \in \mathcal{V}^h, \\ \left( \mu_{n+1}^{k+1}, \widehat{\mu} \right) - \left( \epsilon_T^2 \nabla \phi_{T_{n+1}}^{k+1}, \nabla \widehat{\mu} \right) - \left( 2\bar{E} \phi_{T_{n+1}}^{k+1}, \widehat{\mu} \right) &= \left( 2\bar{E} \phi_{T_n} (-3\phi_{T_n} + 2(\phi_{T_n})^2), \widehat{\mu} \right), \quad \forall \widehat{\mu} \in \mathcal{V}^h. \end{aligned} \right\} \quad (42)$$

## 5.2. The OPAL framework

The Occam Plausibility Algorithm (OPAL) (see Figure 2) [16, 33] is used to calibrate, validate, and select models that best correspond to the observed MRI data. The sensitivity analysis step was not activated in these calculations in anticipation of exploring multiple QoIs. Different QoIs have different sensitivities to the parameters, making the OPAL process biased to one QoI. In the following subsections, each step in the process is described using, as an example, the model  $\mathcal{M}_1 = \text{RD01T01} \in \mathcal{M}$ , given by

$$\frac{\partial \phi_T}{\partial t} = \nabla \cdot M_T^* \nabla \phi_T + \lambda_T^g \phi_T (1 - \phi_T) - \mathcal{R}(T_{\text{treat}}, \tilde{\mathcal{D}}_t) \phi_T, \quad \text{in } \Omega \times (0, T), \quad (43)$$

where  $\mathcal{R}(T_{\text{treat}}, \tilde{\mathcal{D}}_t)$  is defined by (16). The data used in this example are drawn from the images given in Figure 1.

### 5.2.1. The Calibration Step

For the particular example (43), the calibration scenario  $S_c$ , in which observational data,  $y_c$ , are collected, is the murine data with radiation treatment of 20Gy and 40Gy. Given a prior probability density,  $\pi(\theta_1, \mathcal{M}_1)$  (e.g., uniform prior), on the vector of model parameters  $\theta_1 \in \Theta_1$ , compute the posterior probability density according to (5). Figure 3 presents the computed calibration posterior  $\pi(\theta_1 | y_c, \mathcal{M}_1)$ , and also the validation posterior, discussed in the next subsection. The calibration priors are assumed to be uniform distributions:  $M_T^* = \mathcal{U}(0.1, 1.0)$ ,  $\lambda_T^g = \mathcal{U}(0.2, 1.6)$ ,  $\lambda_k = \mathcal{U}(0.5, 1.5)$  and  $\lambda_r = \mathcal{U}(0.5, 1.5)$ . Figure 4 presents scatter plots of the calibration posterior of the RD01T01 model with  $T_{\text{treat}} = 14.5$ . In Figure 4, a strong correlation between  $M_T^*$  and  $\lambda_T^g$ , and between  $\lambda_r$  and  $\lambda_k$  can be observed.

The model evidence (10) is computed, and, subsequently, the plausibility (11) is computed for all of the models. The most plausible model, within the same Occam category, is selected for the validation step. The Tempered Monte Carlo algorithm (see [38]) is used for the stochastic approximation of Bayesian evidence (10). This algorithm is available in the library QUESO (Quantification of Uncertainty for Estimation, Simulation, and Optimization) [37]. The details regarding the available options in QUESO and sample codes can be found in [14].

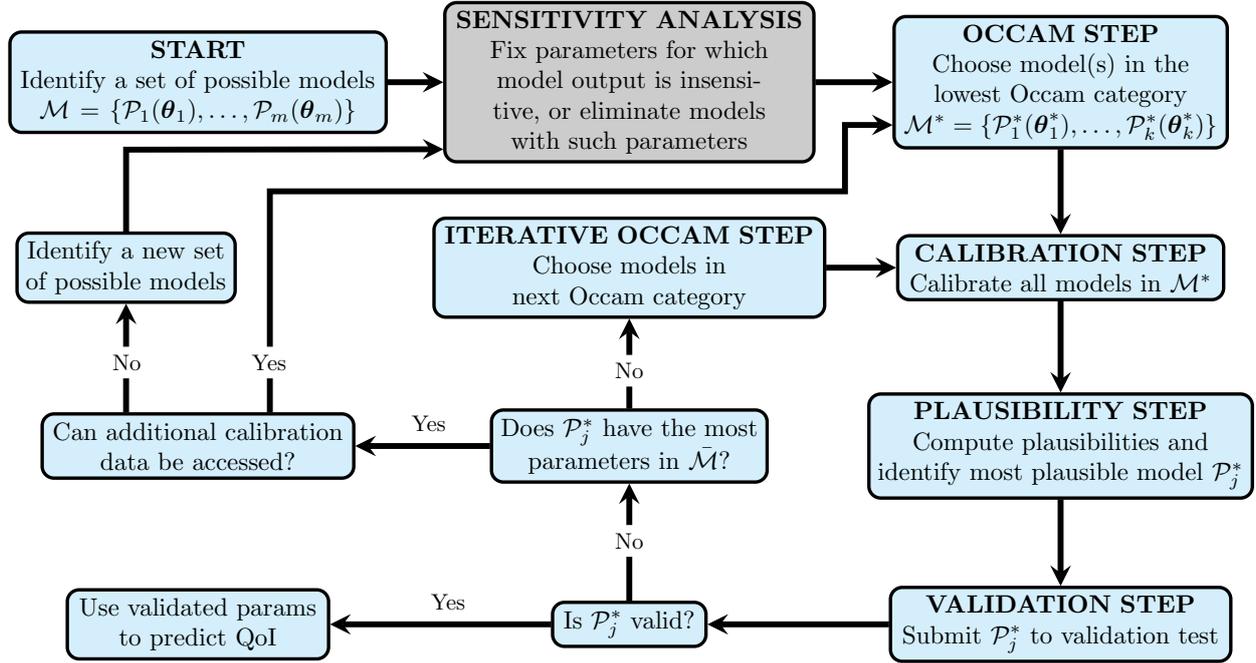


Figure 2: The Occam Plausibility Algorithm [16, 33]. Starting with a set of  $\mathcal{M}$  of parametric model classes, one eliminates models with parameters that do not appreciably effect quantities of interest and then partitions the surviving models into categories, category 1 been the class with the fewest parameters. Then a Bayesian calibration is done in scenario  $S_c$ , model plausibilities are computed, model validation is performed in  $S_v$ , and models which pass the validation test are used to predict quantities of interest in the prediction scenario  $S_p$ .

### 5.2.2. The Validation Step: The Choice of Validation Metrics

Using  $\pi(\theta_1|y_c, \mathcal{M}_1)$  as a validation prior, the parameters in the validation scenario  $S_v$  are updated using (6). The prior used at the validation step (the calibration posterior) and the validation posterior  $\pi(\theta_1|y_v, y_c, \mathcal{M}_1)$ , which are used to compute the stochastic forward problem if the model is not-invalid, are shown in Figure 3. The fact that the validation posterior is quite different from the calibration posterior can be explained by the fact that the calibration scenario is run only up to one day after the treatment. The fact that little information about the treatment can be captured by the calibration scenario is clearly observed in Figure 3c, where the calibration posterior is similar to uniform prior. This also explains why, as seen in Figures 4b and 4c, the samples for  $\lambda_r$  cover the whole support of the prior. However, from the validation posterior (Figure 3c), it is clear that the limits  $[0.5, 1.5]$  prescribed for  $\lambda_r$  adequately covered the parameter domain. We present a more detailed discussion of this point in Section 5.3.

Figure 5 presents the scatter plots of the validation posterior of the RD01T01 model. At the validation scenario, it is possible to notice that the correlation between  $M_T^* \times \lambda_T^g$  persists.

In order to determine if the model is not-invalid, a validation metric must be selected. As one metric to measure the accuracy of the predictions, the cumulative probability distribution functions (cdf) in  $L^1(\mathbb{R})$  is proposed: Let  $F_t(x)$  and  $S_t(x)$  be the cumulative distribution functions for the model and the data at day  $t$ , respectively. The  $L^1(\mathbb{R})$  metric for the data  $x = \text{area} / \text{volume}$  of the tumor,  $x \geq 0$ , is then

$$d_t(F_t, S_t) = \int_0^\infty |F_t(x) - S_t(x)| dx. \quad (44)$$

In order to be not-invalid, we demand that  $d_t < \gamma_{tol}$ , with  $\gamma_{tol}$  being 6% of the tumor area at the last day of the validation scenario. As is shown in Section 6, such a model proves to be not-invalid (“valid”). Figure 6 presents the cdf of the experimental data (solid blue line) and the model results for the validation scenario using the updated posterior (dotted red line) and also the uniform prior (dashed green line). The results with the uniform prior are presented in order to compare the possible variance of the tumor area before the correct calibration of the parameters.

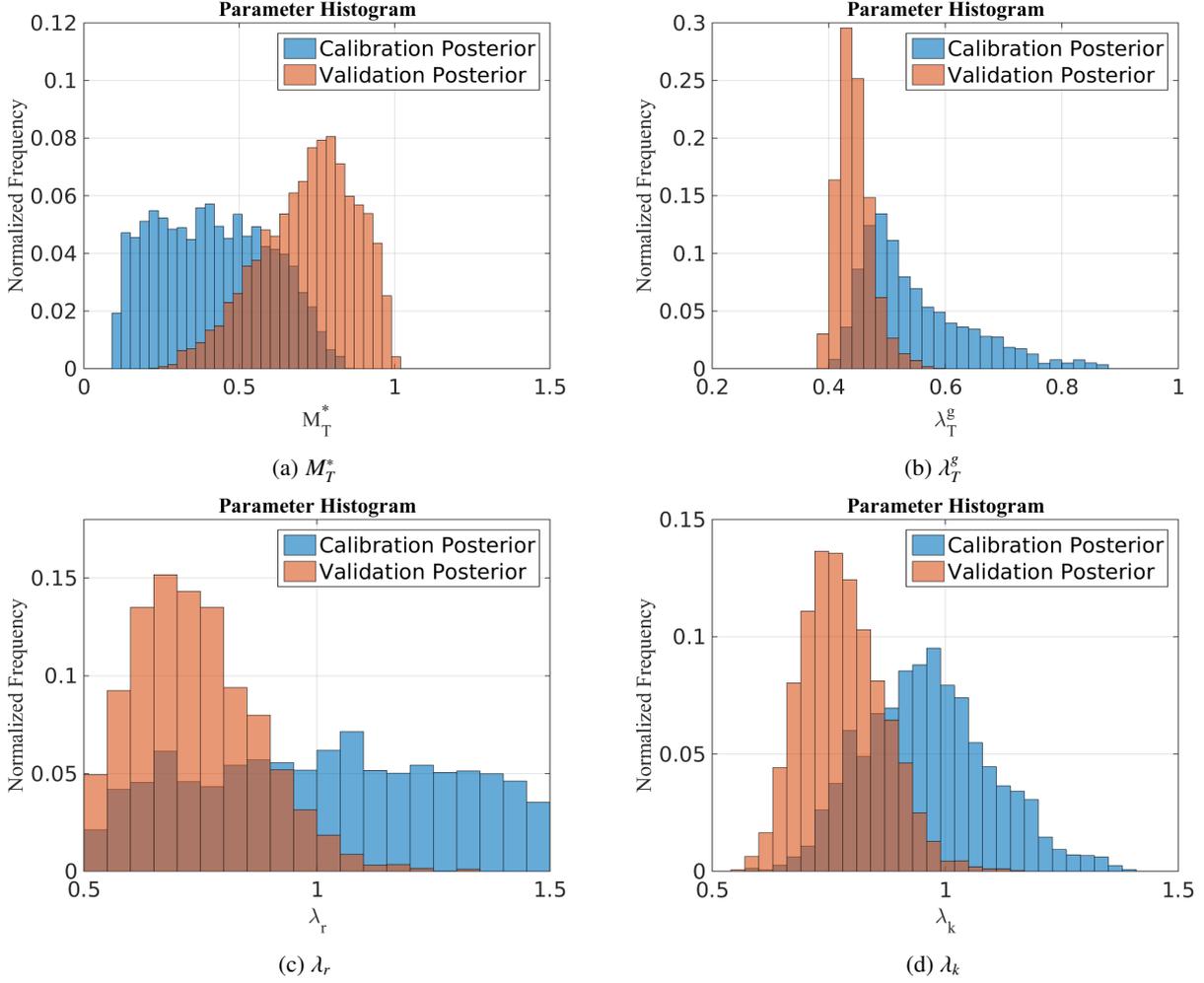


Figure 3: Histogram of the calibration posterior (blue) and validation posterior (red) parameters, normalized to sum 1, of the RD01T01 model ( $M_T^*$ ,  $\lambda_T^g$ ,  $\lambda_k$  and  $\lambda_r$ ).

### 5.2.3. The Prediction Step: The Forward Problem

Using the updated parameters  $\theta_1$ , the forward problem is now solved in the prediction scenario  $S_p$ . Traditionally, the forward problem is solved using only the Maximum A Posteriori (MAP) estimate of the parameters of the models. Here, however, in order to also compute the relevant uncertainties in the predictions of the models, we adopt a Monte Carlo approach to advance the forward problem.

Instead of relying on the MAP estimates of the parameters, we draw repeatedly, randomly and uniformly from the posterior sample  $\pi(\theta_1|y_v, y_c, \mathcal{M}_1)$  of the parameters of the model obtained in validation step. Then, corresponding to each set of parameters drawn from the validation posterior distribution, the forward problem is solved to obtain the corresponding values of QoI. By repeating this process, we then obtain a distribution for the QoI. This approach provides us with a systematic method of estimating the uncertainties in the QoI, for example, by measuring the mean and variance of the distribution of QoI. It is notable that this sampling procedure is valid, only because the sampled parameters from a validation posterior distribution are assumed to be independent and identically distributed (i.i.d), as they are in our case. Therefore, it suffices for QUESO, our Monte Carlo library [37], to randomly and uniformly draw sets of parameter values from the validation posterior distribution of parameters, and compute the QoI corresponding to each parameter set.

Using  $\mathcal{N}$  samples of  $\theta_1 = \{M_T^*, \lambda_T^g, \lambda_k, \lambda_r\}$  from the full posterior, the forward problem is solved with a conventional

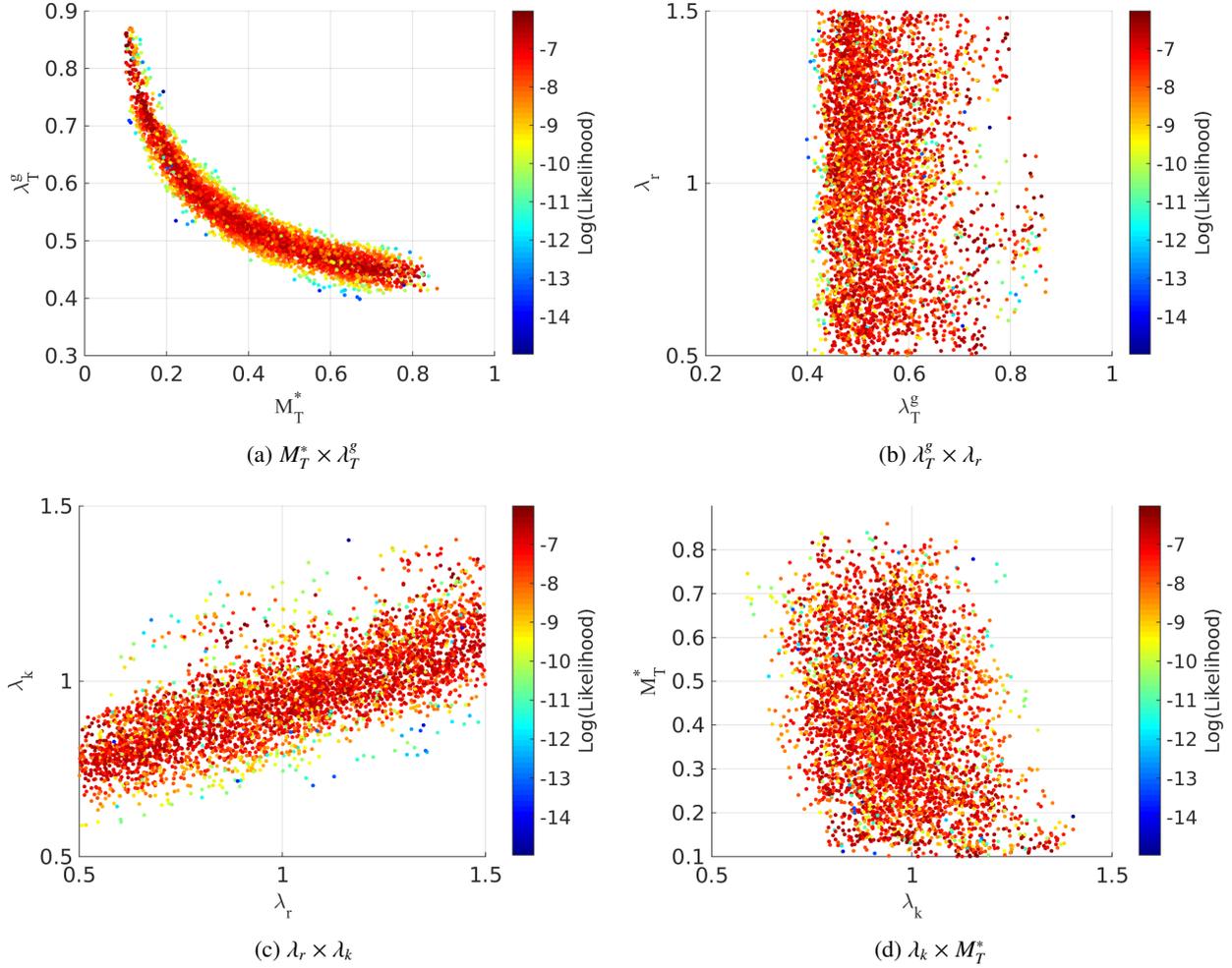


Figure 4: Scatter plot of the calibration posterior of the RD01T01 model. The color scheme represents the log-likelihood, with red being the highest and blue the lowest values.

Monte Carlo method, and the distribution of the QoI (here the tumor area/volume) is computed. In Figure 7, a comparison of the experimental data cumulative distribution function with the model prediction is presented. The variance of the experimental data is assumed to be 10%, while the predicted variance is much smaller than the data. Once more, the uniform prior is seen to allow a wide range of tumor area (dashed green line); however, due to the calibration and validation step, the correct distribution of the parameters reduce the variance of the prediction.

### 5.3. Dealing with Parameter Degeneracy in Models

In scenarios in which the number of model parameters is larger than the number of independent observations, there is always a possibility of overfitting and degeneration of parameter values. This likelihood increases with increase in the complexity of the model. This is an unfortunate result of the scarcity of data, which leads to degenerated model parameter values, even for our best model as illustrated in Figure 4. The Bayesian framework, however, naturally takes into account the effects of overfitting by penalizing the complex models for better fits at the cost of overfitting and degenerate parameter values. Overfitting and parameter degeneracy can be rectified by reducing the number of parameters in the model. To illustrate this, we fix the parameter values of  $M_T^*$  and  $\lambda_r$  in the most plausible model recognized by the OPAL algorithm to 0.2 and 1.0 respectively, and rerun the entire simulation for this reduced model.

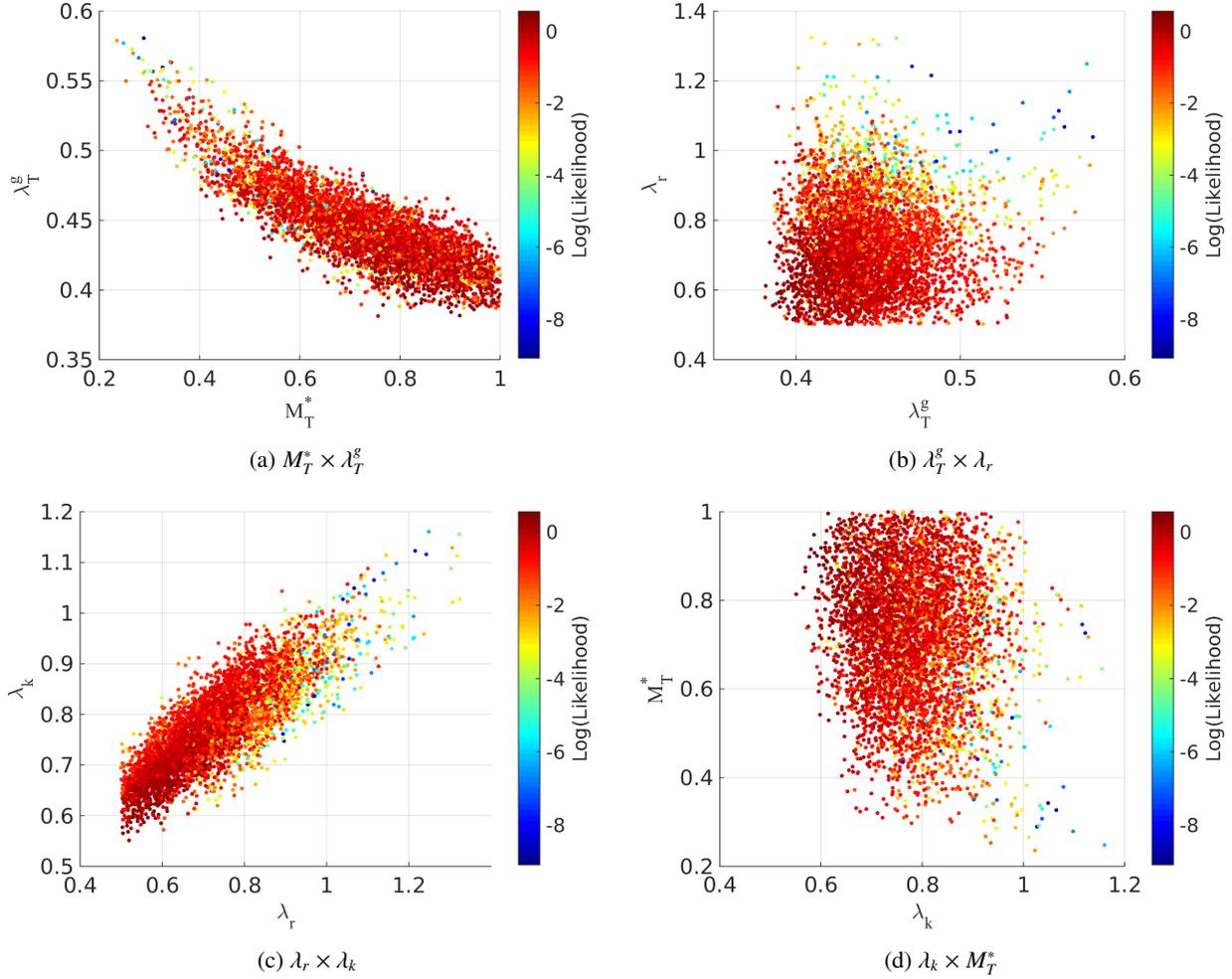


Figure 5: Scatter plot of the validation posterior of the RD01T01 model. The color scheme represents the log-likelihood, with red being the highest and blue the lowest values.

Despite the fact that the original 4-parameter model RD01T01 has been already recognized as the best model, we find that reducing the number of parameters of this model to only 2, results in an even higher Bayesian plausibility ( $2.2 \times 10^{-4}$ ) for this model. The resulting calibration and validation posterior densities are depicted in the marginal distribution plots of Figure 8 and 9. As evidenced by the plots, the reduction in the number of model parameters dramatically changes the MCMC sampled values in the calibration and validation scenarios, reflecting, in this case, the much smaller variance in the posteriors. The validation parameter estimates are also more consistent with the calibration parameter estimates, compared to the original 4-parameter model.

#### 5.4. Ensuring Sample Independence in MCMC

QUESO's sampling algorithm, which is used in this work, is very similar to other methods employing the classical Metropolis-Hasting Markov Chain Monte Carlo (MH-MCMC) algorithm for sampling [31, 20]; in particular, the Parallel-Tempering or Replica Exchange algorithm [40] is employed. A major challenge with MH-MCMC algorithm is that the chain of samples has a tendency to become locked in local modes of the target density being explored. One solution to this locality problem is to use *Parallel Tempering* which is a basis for the QUESO library [37, 38] used in this work.

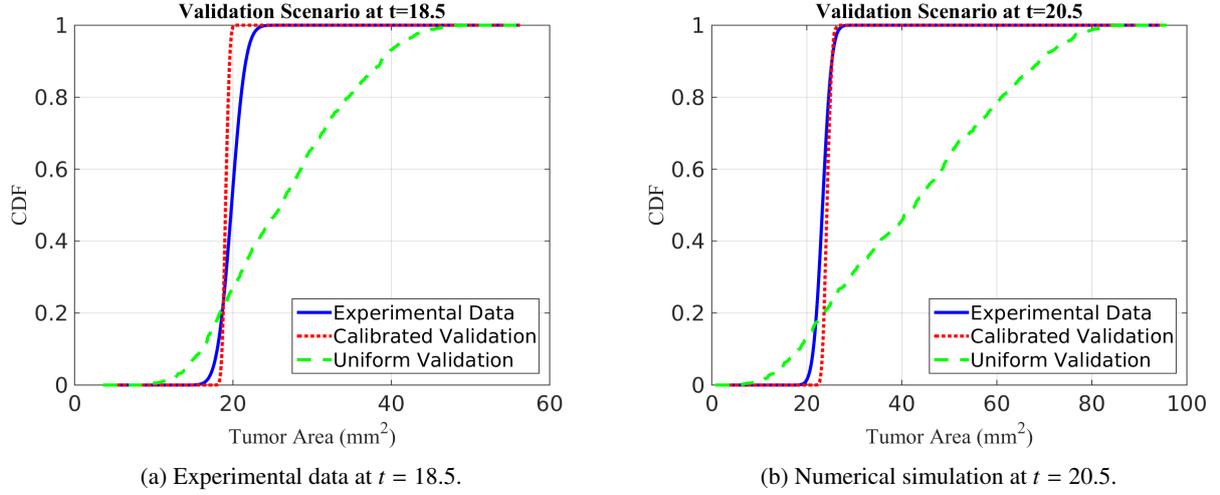


Figure 6: Comparison of the cdf of the measured data (solid blue line), computed simulation using updated posterior (dotted red line) and computed simulation using uniform prior (dashed green line) at the validation scenario. According to the metric defined by (44), model RD01T01 is not-invalid.

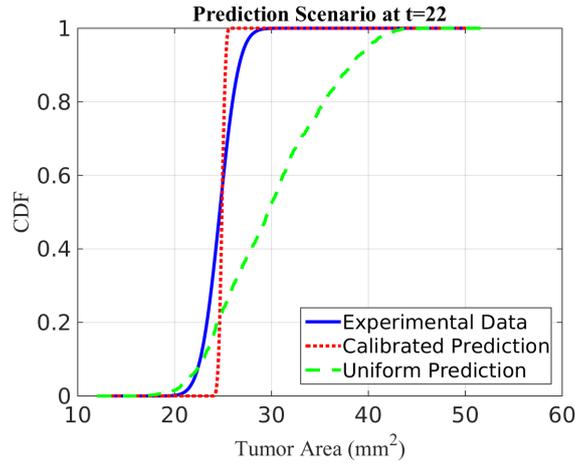


Figure 7: Comparison of the cdf of the measured data (solid blue line), computed simulation using updated posterior (dotted red line) and computed simulation using uniform prior (dashed green line) at the prediction scenario. As mentioned earlier, model RD01T01 is not-invalid.

In Parallel Tempering, we create a discrete set of progressively flatter versions of the target distribution using a temperature parameter  $\mathcal{T}$ . For  $\mathcal{T} = 1$ , the distribution is the desired target distribution. For  $\mathcal{T} > 1$ , the distribution is much flatter, depending on the magnitude of  $\mathcal{T}$ . The basic idea is that by repeatedly flattening the distribution, the sampler can escape from local modes and increase its chance of reaching all regions of the target distribution that contain significant probability. The inference is based on samples drawn from the target density with  $\mathcal{T} = 1$  and *the samples from other higher-temperature chains are discarded*.

In a similar fashion, QUESO's algorithm for sampling starts by sampling from the prior distribution and gradually introduces the likelihood function into the target density to be sampled, through a sequence of temperature levels, one after the other. The target density at the final level represents the full likelihood multiplied by the prior. This method of sampling ensures a full exploration of all posterior modes by gradually moving from prior towards the regions of high probability in the likelihood function. It also provides an approximate value for the integral of the posterior distribution over the parameter space, the so-called Bayesian evidence or marginal likelihood (recall (10)).

The final posterior sample obtained from any algorithm that relies on MCMC methods, including QUESO, has to

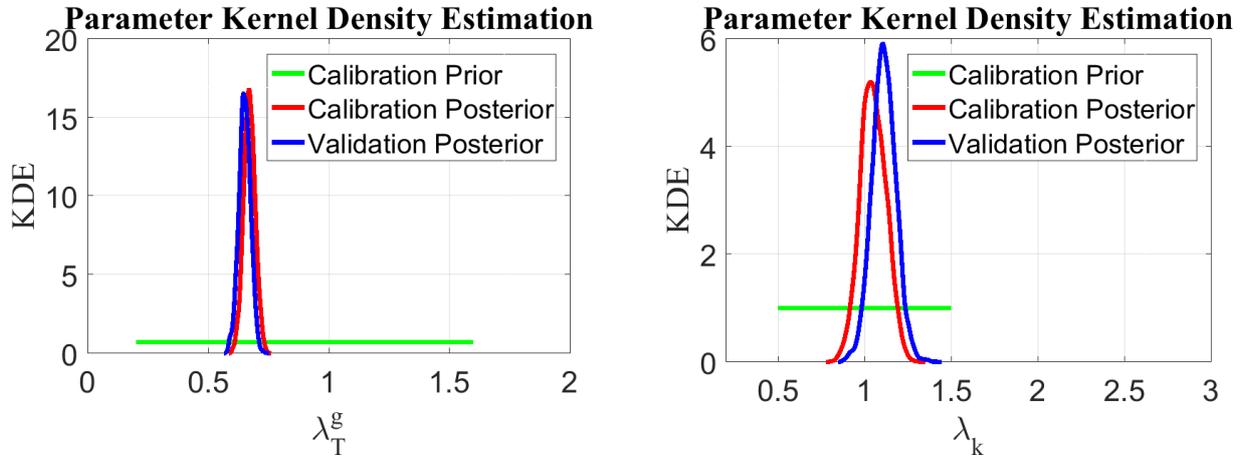


Figure 8: Marginal distribution of the calibration prior (green), calibration posterior (red) and validation posterior (blue) probability densities of the reduced-parameter model discussed in Section 5.3.

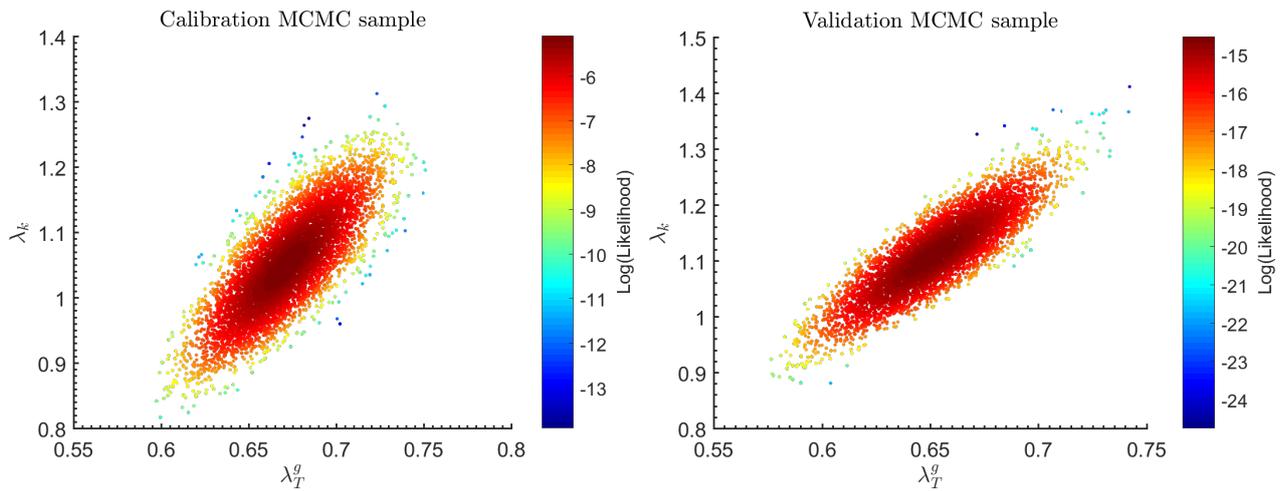


Figure 9: Scatter plots of the calibration and validation posterior probability densities of the two parameters of the reduced model discussed in Section 5.3. The color scheme represents the log-likelihood, the color code from blue to red, represents the likelihood value of each MCMC sampled point from minimum to maximum value.

be also checked for independence. This is most commonly done by plotting the auto-correlation function (ACF) of the MCMC sequence with itself. If the sampling algorithm were able to explore the entire parameter space sufficiently, the plot of ACF should show no significant correlation of each sampled point with any point in the past or future in the sequence. In other words, the occurrence of each individual sample has to have no effect on the occurrences of any of the future points to be sampled.

Sampling results of QUESO are tested for independence using ACF as illustrated by the example plots in Figure 10. In cases where there is significant correlation in the sample, the sample is “thinned” by picking only a fraction of the QUESO’s sampling result. This fraction is determined by the total sample size divided by the observed length of correlation in the sample. Alternatively, one can also increase the sample size, and keep all the sampled points. This approach is also valid, since in the limit of very large sample size, the existence of correlations in the sample becomes irrelevant.

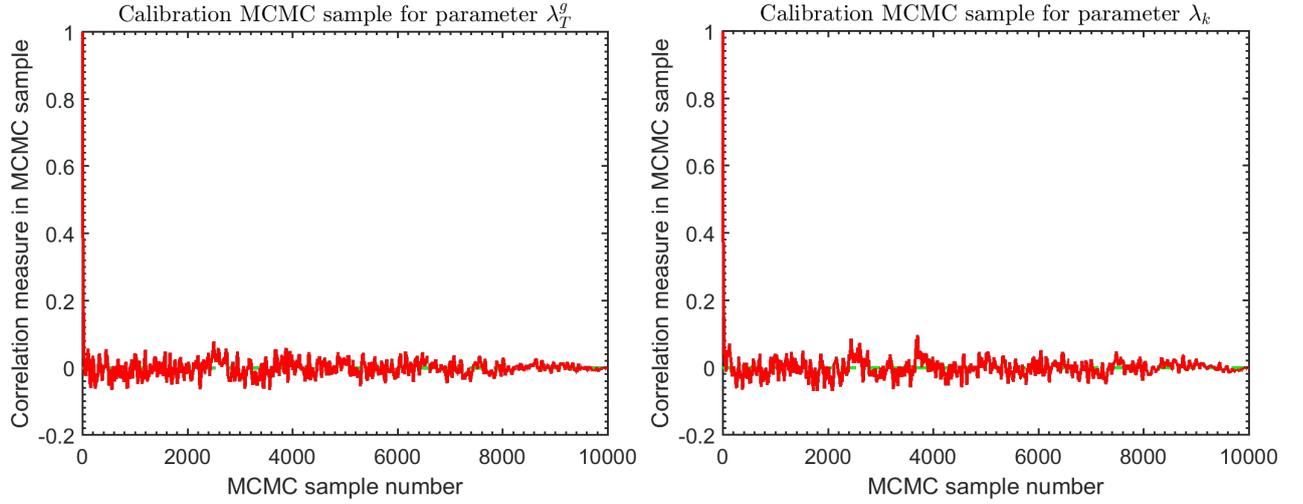


Figure 10: An illustration of the Auto-Correlation Function (ACF) for the calibration sample of the 2-parameter model discussed in Section 5.3. For all models, the initial few-tens of points, exhibiting significant auto-correlation, are excluded from the final sample as they represent the initial *burn-in* points. The corresponding ACF plots for validation scenario also exhibit a similar behavior to that observed in the calibration scenario.

## 6. Numerical Experiments

Numerical experiments are performed using the 39 models exhibited in Table 3, with different reaction terms to capture the radiotherapy effects. The total simulation time, including calibration, validation and prediction of each model, requires from 4 (RD01T01 model) to 36 (PF05T03 model) hours, for the 2D, on serial and 100 hours (RD01T01 model), for the 3D, on 24 processors on the Lonestar 5 Cluster, at the Texas Advanced Computing Center (TACC) of The University of Texas at Austin (URL: <http://www.tacc.utexas.edu>). The experiments are computed assuming that  $\sigma_i^2 = 10\%$  of the area at day  $i$  in (4). Figure 11 presents the experimental data used for model calibration, validation, and prediction. The results for the central MRI slice of the tumor are shown in Figure 11.

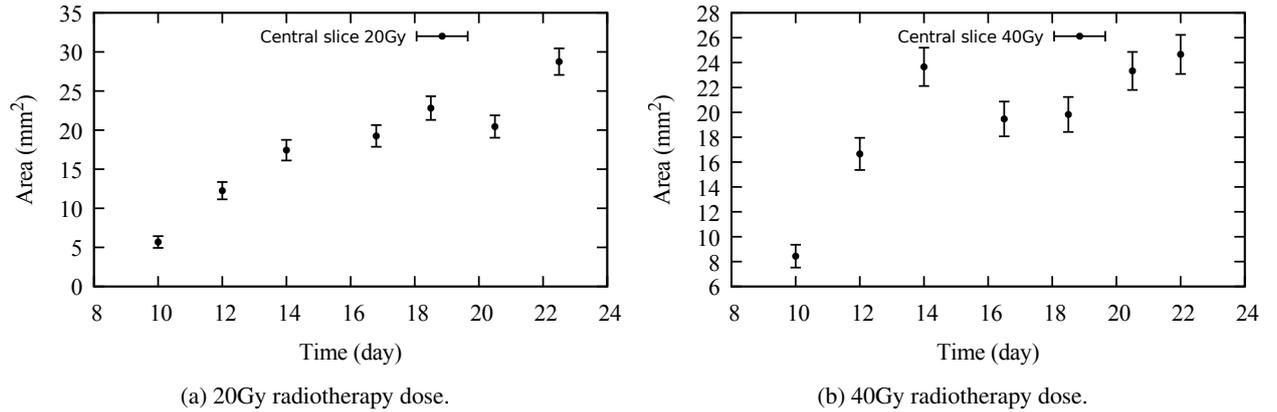


Figure 11: Data used by the OPAL framework. The mean tumor area is obtained by the MRI data and the variance is assumed to be 10% of the tumor area. Central slice of the MRI data, on the left a rat subjected to 20 Gy radiotherapy treatment and on the right a 40Gy radiotherapy applied at  $t = 14.5$  days.

### 6.1. Results with 20Gy radiotherapy

For the data presented in Figure 11a, days 10, 12, 14 and 16.8 are selected for the calibration step, days 18.5 and 20.5 for the validation step, and the prediction is made for day 22.5. The tolerances selected for the validation metric

are:

$$\mathbb{d}_{18.5} \leq \gamma_{tol}^{18.5} = 2.28 \text{ mm}^2, \quad \mathbb{d}_{20.5} \leq \gamma_{tol}^{20.5} = 2.05 \text{ mm}^2, \quad (45)$$

which correspond to an average error of 10% in the model compared to the data. One data point is set out to compare the prediction, the prediction is also desired to be within 10% of the experimental data, that is  $\mathbb{d}_{22.5} \leq \gamma_{tol}^{22.5} = 2.89 \text{ mm}^2$ . The results obtained using the OPAL framework are presented in Table 4 for the murine data with 20 Gy radiotherapy. For this tolerance, none of the models are “valid”. This is due to the fact that the models are unable to capture the decrease in tumor area at day  $t = 20.5$ . If the tolerances are increased to 20% average error, only the models with memory (T01) are considered “not-invalid”.

Generally, it is assumed to be infeasible to compute the plausibility for all the models at the validation step, mostly due to the large computational time required. However, as shown in Table 4, the plausibilities for all the models at the validation step are also computed for this experiment. From the models in Occam category 5, it is seen that the most plausible model from the calibration step is not necessarily the most plausible model at the validation step. This behavior is due to the acquisition of a new data set at the validation step. For the case in which the computation of the plausibilities of all the models is feasible, it is not necessary to discard models at the calibration step, but instead, one can select the most plausible and valid model at the validation step in order to move to the prediction step.

## 6.2. Results with 40Gy radiotherapy

For the data presented in Figure 11b, days 10, 12, 14 and 16.5 are selected for the calibration step, days 18.5 and 20.5 for the validation step, and the prediction is made for day 22. The tolerances selected for the validation metric are:

$$\mathbb{d}_{18.5} \leq \gamma_{tol}^{18.5} = 1.98 \text{ mm}^2, \quad \mathbb{d}_{20.5} \leq \gamma_{tol}^{20.5} = 2.33 \text{ mm}^2, \quad (46)$$

and  $\mathbb{d}_{22} \leq \gamma_{tol}^{22} = 2.47 \text{ mm}^2$  for the prediction. These values correspond to an average error of 10% in the model compared to the experimental data with a 40Gy radiotherapy dose applied at  $t = 14.5$  days. The results obtained using the OPAL framework are presented in Table 5. It is observed that the newly developed model with the new treatment term (model RD01T01) is able to meet the desirable tolerance for the validation criteria used. As the model RD01T01 was selected as the best model, the OPAL framework stops at Occam category 1, without the necessity of performing simulations with models with higher number of parameters. The prediction of the model RD01T01 exhibits an error within 10% (on average a 1.11  $\text{mm}^2$  difference from the experimental data). It is also seen that all the models with memory (T01) and some of the partial memory models (T02) are “not-invalid”. However, all the models without memory are invalid. This means that for the data used, and on the scenario modeled, this type of treatment was unable to correctly represent the decline in tumor area due to the radiotherapy. To model this scenario, the effects of the radiotherapy should continue during some time after the treatment, instead of having an instantaneous killing effect [39, 6].

If we require that the average error in the model compared to the data be less than 6% rather than 10%, then the new tolerance are given by

$$\mathbb{d}_{18.5} \leq \gamma_{tol}^{18.5} = 1.19 \text{ mm}^2, \quad \mathbb{d}_{20.5} \leq \gamma_{tol}^{20.5} = 1.40 \text{ mm}^2. \quad (47)$$

With a 10% tolerance, RD01T01, RD02T01 and RD03T01 are determined to be “not-invalid”.

In a strict application of OPAL, the model selection and validation processes would be terminated after Occam category 1 as model RD01T01 is found to be valid and to have a high plausibility for the given data in both the calibration and the validation scenarios. However, results for the full set of models are provided to display the full spectrum of model plausibilities and model inadequacies. It is observed that the most plausible models in an Occam category may not be valid. Surprisingly, for high-category models with many parameters, the opposite behavior is observed: models with low plausibility may, in fact, be valid. This can be attributed to several possible effects: first, the available data is sparse, so a type of “over-fitting” of parameters can result when fitted to validation data, a phenomenon not expected in Bayesian approaches, and one that could conceivably be avoided with more informative priors. Second, full convergence of the MCMC algorithm may have not been achieved, and, finally, the validity of such models could be refuted with the acquisition of additional data, an acquisition not possible in the present investigation.

Table 4: Results using the variance as  $\sigma_i^2 = 10\%$  of the area at day  $i$ . The tolerances prescribed captures an average error of 20% in the model compared to the data, setting the new tolerances to  $\gamma_{tol}^{18.5} = 4.56$  and  $\gamma_{tol}^{20.5} = 4.09$ . The not-invalid models are highlighted in orange.

Model	Occam Category	Calibration	Validation			Prediction
		Plausibility	Plausibility	$d_{18.5}$ (mm <sup>2</sup> )	$d_{20.5}$ (mm <sup>2</sup> )	$d_{22.5}$ (mm <sup>2</sup> )
RD01T01	1	0.99	1.00	3.96	2.71	5.17
RD01T02	1	0.01	0.00	1.72	4.77	5.39
RD01T03	2	n/a	n/a	1.21	11.11	2.31
PF01T01	3	1.00	1.00	3.88	2.91	5.35
PF01T02	3	0.00	0.00	1.65	4.50	5.32
PF01T03	4	n/a	n/a	1.11	8.16	3.35
RD02T01	5	0.33	0.23	3.72	3.48	5.31
RD02T02	5	0.00	0.00	1.66	5.16	5.16
RD03T01	5	0.23	0.50	3.81	3.26	4.99
RD03T02	5	0.00	0.00	1.48	6.31	4.69
MD01T01	5	0.42	0.27	4.25	2.84	5.63
MD01T02	5	0.01	0.00	1.80	4.47	5.61
RD02T03	6	0.00	0.00	1.17	10.69	2.55
RD03T03	6	0.00	0.00	1.15	9.55	3.16
MD01T03	6	0.01	0.00	1.14	7.89	3.93
RD04T01	6	0.03	0.03	3.67	3.63	4.50
RD04T02	6	0.00	0.00	1.54	5.54	5.00
MD02T01	6	0.05	0.37	4.11	3.22	5.62
MD02T02	6	0.00	0.00	1.81	4.48	5.55
MD03T01	6	0.07	0.52	4.06	2.86	5.56
MD03T02	6	0.00	0.00	1.62	5.24	5.18
PF02T01	6	0.80	0.08	3.61	3.37	5.09
PF02T02	6	0.03	0.00	1.75	4.23	5.50
RD04T03	7	0.00	0.00	1.26	11.59	2.24
MD02T03	7	0.02	0.00	1.13	8.48	3.64
MD03T03	7	0.01	0.00	1.11	9.28	3.25
PF02T03	7	0.00	0.00	1.13	8.63	3.06
MD04T01	7	0.03	0.12	4.02	2.99	5.72
MD04T02	7	0.00	0.00	1.28	7.71	3.97
PF03T01	7	0.30	0.11	3.76	3.15	5.22
PF03T02	7	0.00	0.00	1.50	5.42	4.92
PF04T01	7	0.63	0.77	3.87	2.89	5.29
PF04T02	7	0.00	0.00	1.24	6.56	4.27
MD04T03	8	0.01	0.00	1.14	8.39	3.66
PF03T03	8	0.39	0.00	1.13	9.28	2.99
PF04T03	8	0.10	0.00	1.22	9.73	2.58
PF05T01	8	0.50	1.00	3.93	2.90	5.32
PF05T02	8	0.00	0.00	1.61	4.48	5.27
PF05T03	9	n/a	n/a	1.21	9.47	2.73

Table 5: Results using the variance as  $\sigma_i^2 = 10\%$  of the area at day  $i$ . The tolerances prescribed captures an average error of 10% in the model compared to the data, setting the new tolerances to  $\gamma_{tol}^{18.5} = 1.98$  and  $\gamma_{tol}^{20.5} = 2.33$ . The not-invalid models are highlighted in orange.

Model	Occam Category	Calibration Plausibility	Validation			Prediction $\mathbb{d}_{22}$ (mm <sup>2</sup> )
			Plausibility	$\mathbb{d}_{18.5}$ (mm <sup>2</sup> )	$\mathbb{d}_{20.5}$ (mm <sup>2</sup> )	
RD01T01	1	0.99	1.00	1.13	0.98	1.11
RD01T02	1	0.01	0.00	3.83	6.81	1.70
RD01T03	2	n/a	n/a	3.74	6.58	1.64
PF01T01	3	1.00	1.00	1.31	1.72	1.06
PF01T02	3	0.00	0.00	2.55	2.81	1.23
PF01T03	4	n/a	n/a	6.67	12.40	3.45
RD02T01	5	0.15	0.71	1.08	1.07	1.12
RD02T02	5	0.00	0.00	1.64	1.08	1.13
RD03T01	5	0.20	0.23	1.16	1.14	0.99
RD03T02	5	0.01	0.00	4.11	7.26	1.94
MD01T01	5	0.63	0.06	1.43	1.49	1.07
MD01T02	5	0.01	0.00	1.97	1.84	1.14
RD02T03	6	0.01	0.00	3.69	6.46	1.62
RD03T03	6	0.01	0.00	4.45	8.31	2.03
MD01T03	6	0.05	0.00	4.18	7.69	1.85
RD04T01	6	0.06	0.15	0.92	2.22	0.91
RD04T02	6	0.00	0.00	2.16	2.36	1.15
MD02T01	6	0.05	0.02	1.69	2.08	1.20
MD02T02	6	0.00	0.00	2.14	2.43	1.16
MD03T01	6	0.04	0.53	1.29	1.45	1.07
MD03T02	6	0.00	0.00	2.62	3.73	1.21
PF02T01	6	0.74	0.30	1.15	1.72	1.07
PF02T02	6	0.03	0.00	1.82	1.22	1.11
RD04T03	7	0.01	0.00	4.03	7.12	1.87
MD02T03	7	0.02	0.00	7.99	15.30	4.26
MD03T03	7	0.02	0.00	4.58	8.88	2.05
PF02T03	7	0.05	0.00	7.64	14.73	4.08
MD04T01	7	0.04	0.41	1.51	1.75	1.12
MD04T02	7	0.01	0.00	1.98	2.03	1.16
PF03T01	7	0.41	0.26	1.30	1.71	1.09
PF03T02	7	0.00	0.00	1.78	1.17	1.13
PF04T01	7	0.43	0.33	1.26	1.75	1.09
PF04T02	7	0.00	0.00	1.85	1.28	1.11
MD04T03	8	0.05	0.00	6.96	15.16	3.47
PF03T03	8	0.05	0.00	5.60	11.76	2.59
PF04T03	8	0.03	0.00	7.12	13.53	3.74
PF05T01	8	0.83	1.00	1.25	1.88	1.07
PF05T02	8	0.03	0.00	1.72	1.14	1.14
PF05T03	9	n/a	n/a	7.39	14.12	3.92

### 6.3. 3D experiments

The results presented in Sections 6.1 and 6.2 are obtained from experiments in 2D tomographic slices of the tumor, where the QoI is the tumor area at the central slice. However, for the 3D simulation, the tumor volume is computed as the QoI. In Figure 12, we compare the experimental data for the time evolution of the tumor area at the central slice and for the tumor volume, for the rat treated with 40Gy radiotherapy. The overall tumor growth rate is clearly higher for the 3D scenario.

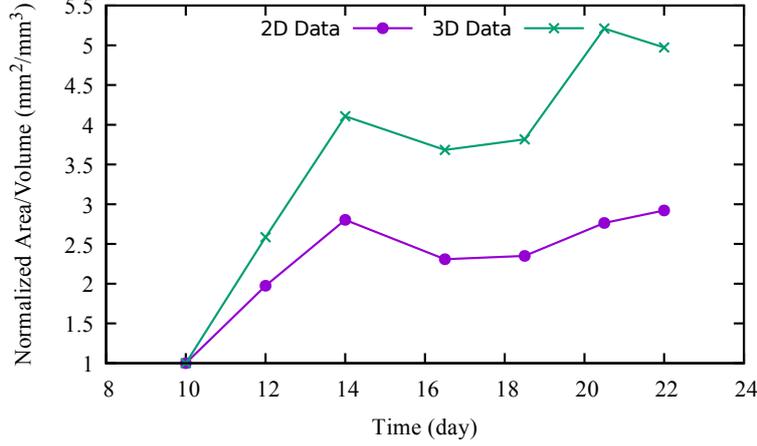


Figure 12: Time evolution of the mean area (purple) and mean volume (green) of experimental data for the 40Gy data with treatment at  $t = 14.5$  days. The area and volume exhibited have been normalized by their value at  $t = 10$  days.

In order to compare the results between the 2D and 3D experiments, we select the RD01T01 model, which is the most plausible model found by the 2D OPAL experiments (Table 5). For the data presented in Figure 12, days 10, 12, 14 and 16.5 are selected for the calibration step with treatment at day  $T_{treat} = 14.5$ , days 18.5 and 20.5 for the validation step, and the prediction is made for day 22. The tolerances selected for the validation metric are:

$$\mathfrak{d}_{18.5}^{2D} \leq \gamma_{tol}^{18.5} = 1.98 \text{ mm}^2, \quad \mathfrak{d}_{20.5}^{2D} \leq \gamma_{tol}^{20.5} = 2.33 \text{ mm}^2, \quad (48)$$

$$\mathfrak{d}_{18.5}^{3D} \leq \gamma_{tol}^{18.5} = 14.20 \text{ mm}^3, \quad \mathfrak{d}_{20.5}^{3D} \leq \gamma_{tol}^{20.5} = 19.39 \text{ mm}^3, \quad (49)$$

and  $\mathfrak{d}_{22}^{2D} \leq \gamma_{tol}^{22} = 2.47 \text{ mm}^2$  and  $\mathfrak{d}_{22}^{3D} \leq \gamma_{tol}^{22} = 18.50 \text{ mm}^3$  for the prediction. In Table 6, the results for the 2D and 3D validation and prediction are presented. The 3D simulations are computed for two different scenarios: one scenario uses the same uniform calibration priors as the 2D experiments, and the second 3D scenario uses the 2D validation posterior as the calibration prior. As presented in Table 6, for the 3D scenarios, the RD01T01 model is not able to correctly predict the tumor volume decrease at day 22 within the tolerance desired. This decrease in the measured value can be due to radiotherapy delayed effects that the model is not able to capture or to uncertainties in observational data and the MRI acquisition process.

Table 6: Results using the variance as  $\sigma_i^2 = 10\%$  of the area at day  $i$ . The tolerances prescribed capture an average error of 10% in the model compared to the data. Comparison between the most plausible model (RD01T01) at 2D and 3D scenarios.

Model RD01T01	Validation		Prediction
	$\mathfrak{d}_{18.5} (\text{mm}^2)$	$\mathfrak{d}_{20.5} (\text{mm}^2)$	$\mathfrak{d}_{22} (\text{mm}^2)$
2D	1.13	0.98	1.11
3D (uniform prior)	3.84	5.04	31.98
3D (2D prior)	3.85	2.90	32.45

Figure 13 compares the parameter distribution between the 2D and the 3D scenarios, with both using the uniform calibration prior. The tumor growth rate  $\lambda_T^g$  is higher in the 3D scenario. The parameters  $M_T^*$  and  $\lambda_r$  have a similar mean between the two scenarios, while the  $\lambda_k$  is higher in the 3D scenario. In order to use all the information available,

the 2D validation posterior (blue histogram in Figure 13) is used as a prior for the 3D experiments. Figure 14 compares the histogram of the validation parameters using the 2D validation posterior and the uniform priors. The mean of  $\lambda_T^g$ ,  $\lambda_r$  and  $\lambda_k$  parameters for this new scenario is closer to the 2D mean.

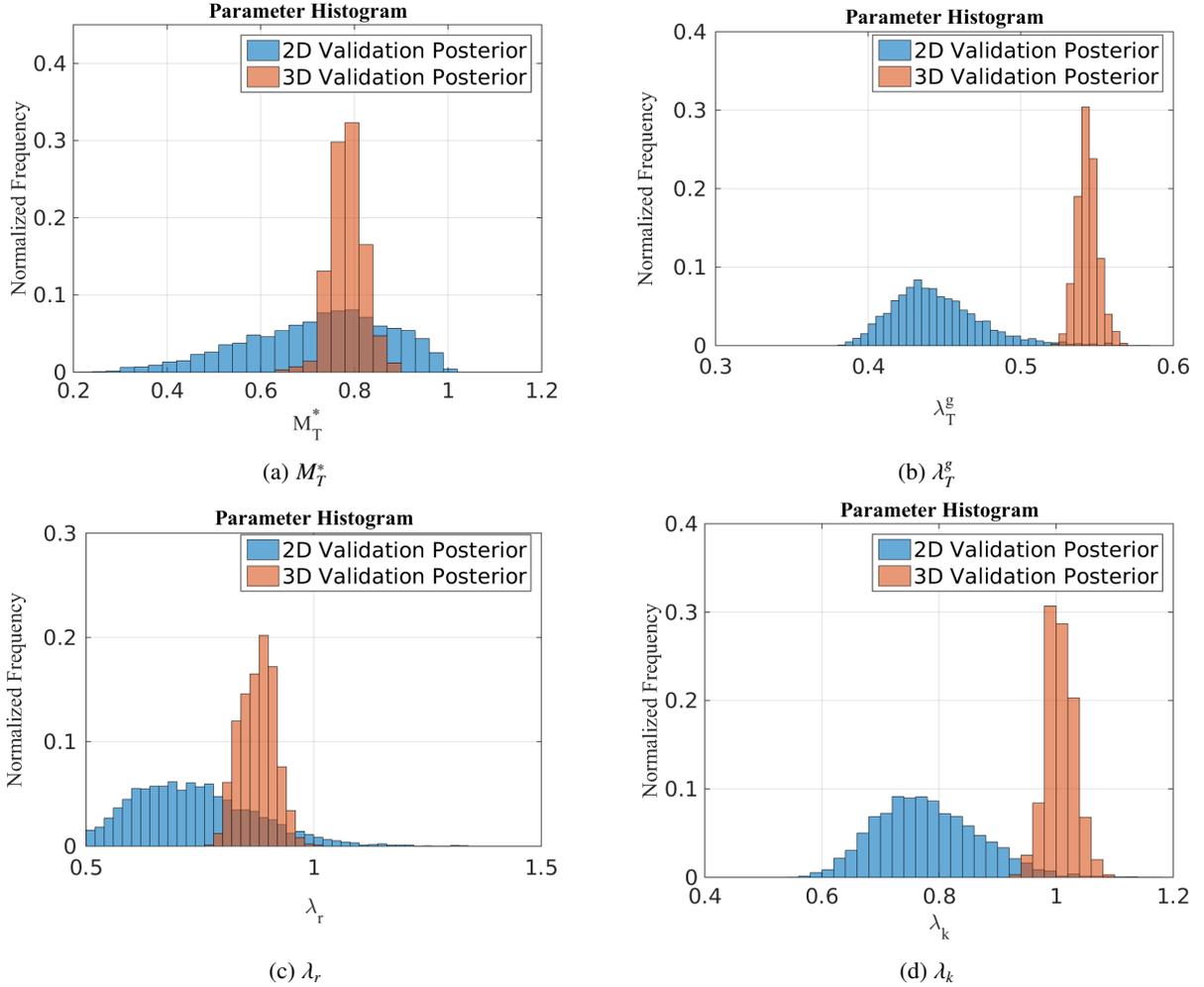


Figure 13: Histogram of the 2D validation posterior (blue) and the 3D validation posterior (red) parameters, normalized to sum 1, of the RD01T01 model ( $M_T^*$ ,  $\lambda_T^g$ ,  $\lambda_k$  and  $\lambda_r$ ).

In Figures 15 and 16, 3D deterministic results obtained using the RD01T01 model with radiotherapy are shown. The parameter values are obtained through the 3D calibration and validation of the tumor volume. The values selected are the ones corresponding to maximum likelihoods. In Figure 15, the calibration posterior is used in order to obtain the initial characterization of the parameters, while in Figure 16, the model is restarted at day  $t = 16.5$  using the updated values for the parameters from the validation posterior. Once more, the increase of tumor volume is observed at time  $t = 22$  days.

## 7. Conclusions

This investigation has demonstrated that a number of emerging and existing technologies, theories, numerical and statistical tools can be brought to bear in the problem of reliable prediction of the evolution of glioma tumors in laboratory animals that include models of the effects of X-ray radiation. Effective models of tissue-scale phenomena can be developed from continuum models of balance laws and accepted biological principles.

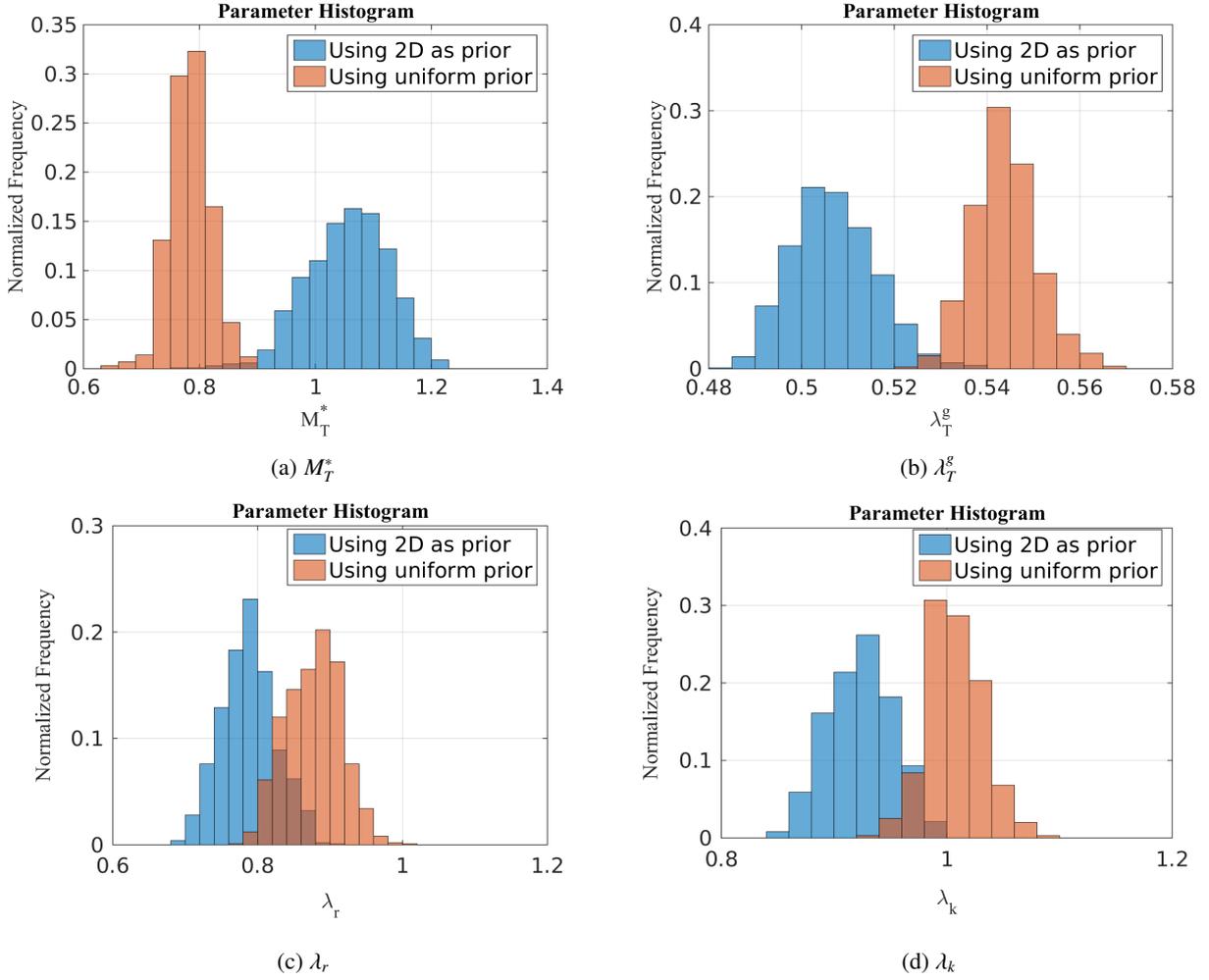


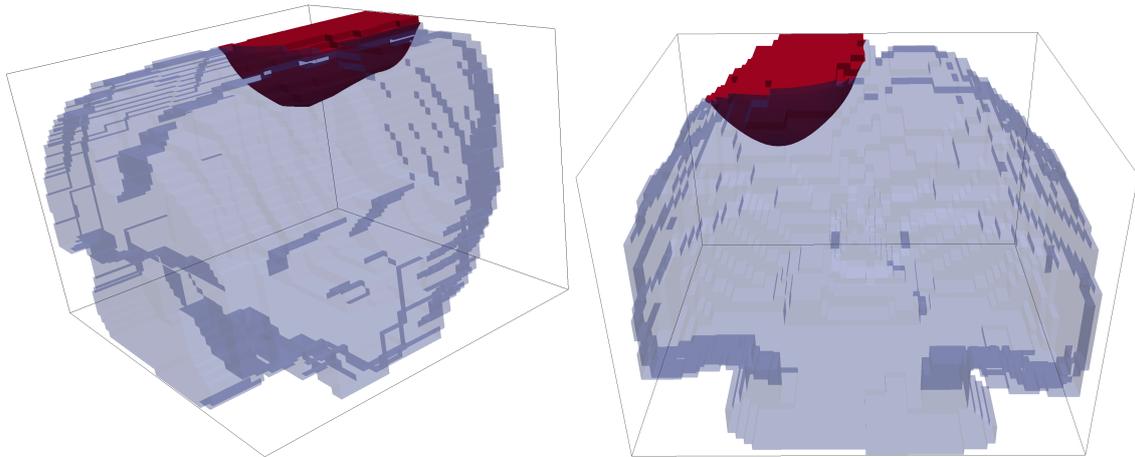
Figure 14: Histogram of the 3D validation posterior using the 2D as a prior (blue) and the 3D validation posterior using the uniform prior (red) parameters, normalized to sum 1, of the RD01T01 model ( $M_T^*$ ,  $\lambda_T^g$ ,  $\lambda_k$  and  $\lambda_r$ ).

MRI can be used non-invasively to deliver data essential in model calibration and validation processes. The OPAL algorithm provides a general framework for model selection, model validation, and prediction when implemented with the use of effective Monte Carlo stochastic PDE-solvers and MCMC sampling methods. The approach is regarded as “model agnostic” as it does not presume that a valid model is known in advance of a prediction.

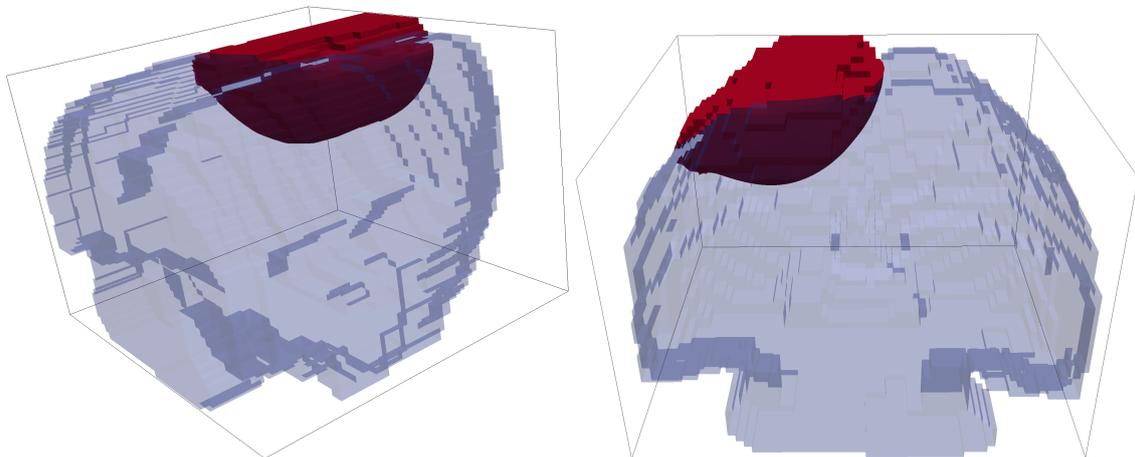
The results suggest that, for the data used, and on the scenario modeled, the classic linear-quadratic model (T03) is not able to capture the effects of the radiotherapy. This work has demonstrated that in order to obtain a better agreement with the data, the effects of the radiotherapy should be prolonged after the treatment (T01 and T02 models), instead of having an instantaneous killing effect (T03 model).

## Acknowledgements

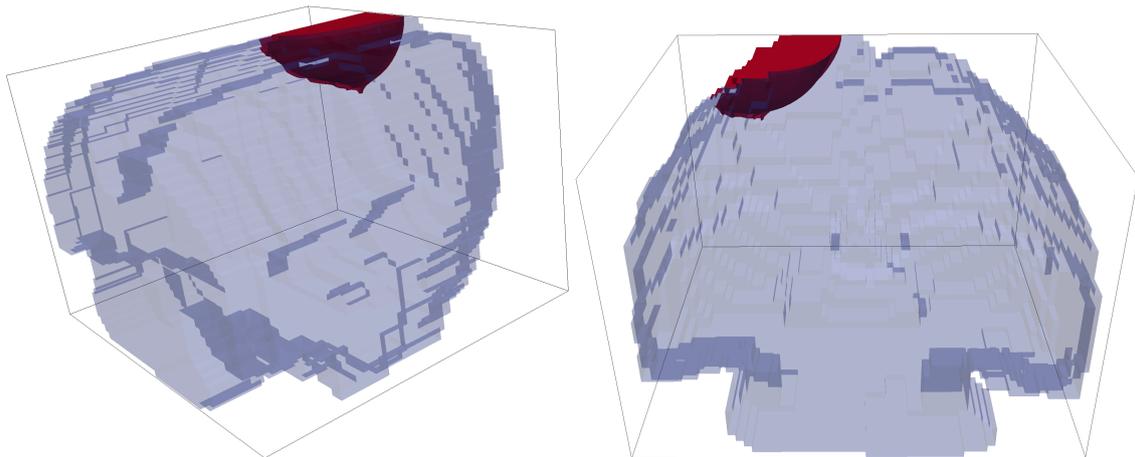
Early support of work on tumor growth modeling by the National Science Foundation under grant DMS 1115865. Work on predictive modeling and uncertainty quantification is supported by U.S. Department of Energy Office of Science, Office of Advanced Scientific Computing Research, Applied Mathematics program under Award Number DE-5C0009286. The authors also acknowledge the support of the Cancer Prevention Research Institute of Texas



(a)  $t = 12$  days.

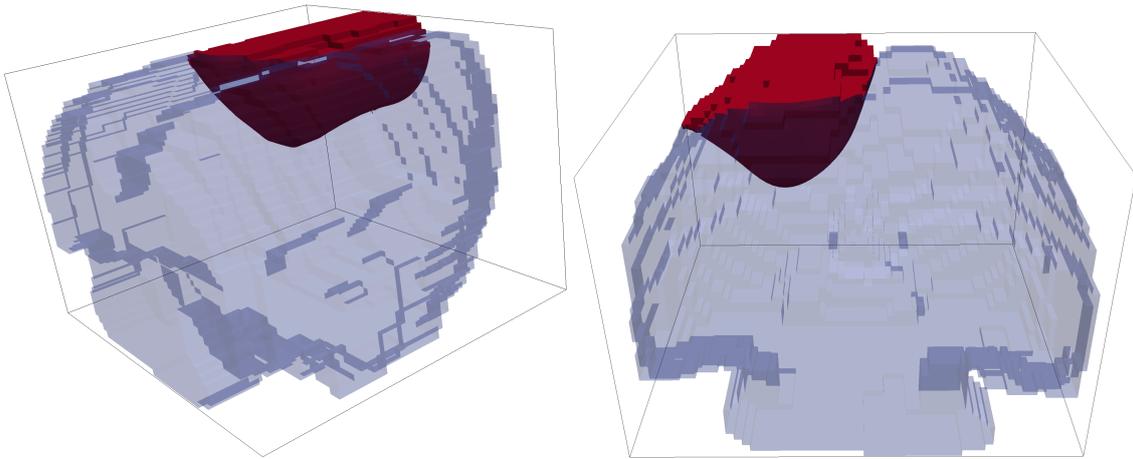


(b)  $t = 14$  days.

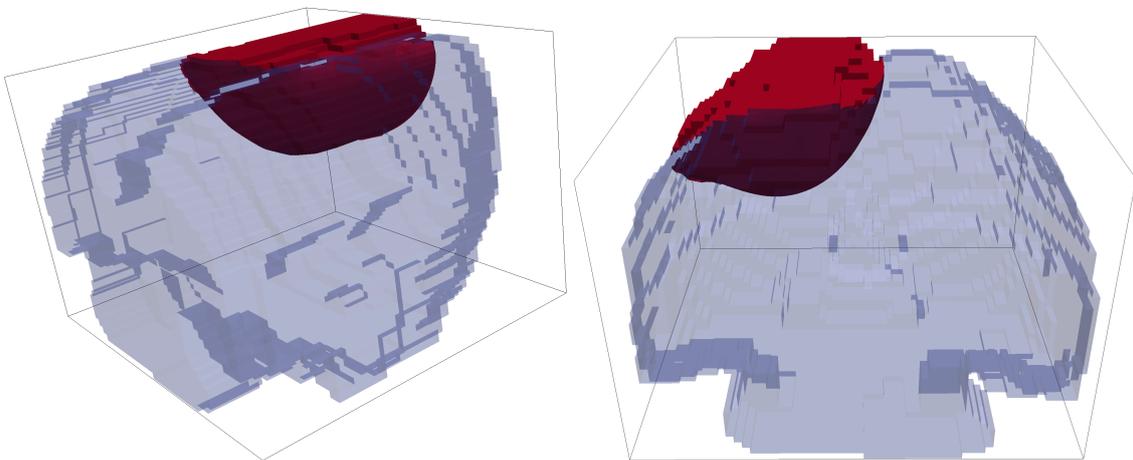


(c)  $t = 16.5$  days.

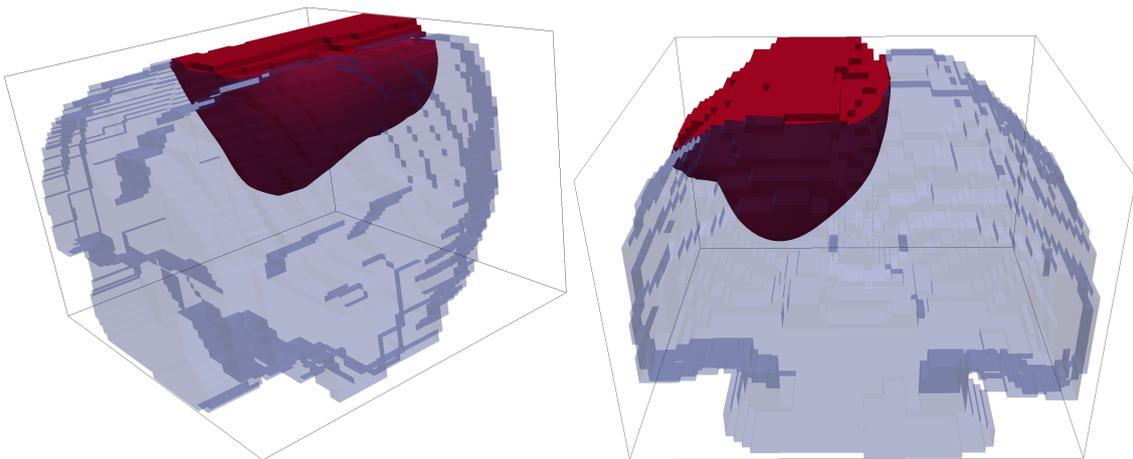
Figure 15: Numerical simulation of the RD01T01 model at the calibration scenario ( $t = 12$ ,  $t = 14$  and  $t = 16.5$  days), with treatment being applied at  $t = 14.5$  days. In red the tumor and in light blue the brain domain. The computational domain is bounded by a  $15 \times 15 \times 10$  mm box. The results are presented with two different perspectives (left and right figures) in order to help the visualization of the whole 3D tumor.



(a)  $t = 18.5$  days.



(b)  $t = 20.5$  days.



(c)  $t = 22$  days.

Figure 16: Numerical simulation of the RD01T01 model at the validation scenario ( $t = 18.5$  and  $t = 20.5$  days) and at the prediction scenario  $t = 22$ , with treatment being applied at  $t = 14.5$  days. In red the tumor and in light blue the brain domain. The computational domain is bounded by a  $15 \times 15 \times 10$  mm box. The results are presented with two different perspectives (left and right figures) in order to help the visualization of the whole 3D tumor.

(CPRIT) for funding part of this work through grant RR160005, the NCI for funding through grants U01CA174706, and R01CA186193, and the German Science Foundation (DFG, WO-671 11-1).

The authors gratefully acknowledge the helpful discussions with Drs. Angela Jarrett, Regina C. Almeida, Renato S. Silva and Xinzeng Feng about the presentation of the results and helpful hints to some references. The authors thank Dr. Zou Yue for performing the animal surgeries and Drs. Michael L. Freeman and Sekhar R. Konjeti for assistance and usage of the X-ray machine.

## References

- [1] N. C. Atuegwu, J. C. Gore, and T. E. Yankeelov. The integration of quantitative multi-modality imaging data into mathematical models of tumors. *Phys Med Biol*, 55(9):2429–49, 2010.
- [2] N. C. Atuegwu, L. R. Arlinghaus, X. Li, E. B. Welch, B. A. Chakravarthy, J. C. Gore, and T. E. Yankeelov. Integration of diffusion-weighted MRI data and a simple mathematical model to predict breast tumor cellularity during neoadjuvant chemotherapy. *Magn Reson Med*, 66(6):1689–96, 2011.
- [3] N. Bellomo, N. K. Li, and P. K. Maini. On the foundations of cancer modelling: selected topics, speculations, and perspectives. *Mathematical Models and Methods in Applied Sciences*, 18(4):593–646, 2008.
- [4] N. Bellomo, A. Bellouquid, Y. Tao, and M. Winkler. Toward a mathematical theory of Keller–Segel models of pattern formation in biological tissues. *Mathematical Models and Methods in Applied Sciences*, 25(09):1663–1763, 2015.
- [5] L. Bodgi, A. Canet, L. Pujo-Menjouet, A. Lesne, J.-M. Victor, and N. Foray. Mathematical models of radiation action on living cells: From the target theory to the modern approaches. A historical and critical review. *Journal of Theoretical Biology*, 394:93–101, 2016.
- [6] G. Borasi and A. Nahum. Modelling the radiotherapy effect in the reaction-diffusion equation. *Physica Medica*, 32(9):1175–1179, 2016.
- [7] R. M. Bowen. Theory of mixtures. In A. C. Eringen, editor, *Continuum Physics III*. Academic Press, New York, USA, 1976.
- [8] D. C. Colvin, M. E. Loveless, M. D. Does, Z. Yue, T. E. Yankeelov, and J. C. Gore. Earlier detection of tumor treatment response using magnetic resonance diffusion imaging with oscillating gradients. *Magnetic resonance imaging*, 29(3):315–323, 2011.
- [9] V. Cristini and J. Lowengrub. *Multiscale modeling of cancer: an integrated experimental and mathematical modeling approach*. Cambridge University Press, 2010.
- [10] J. Crowther. Some considerations relative to the action of x-rays on tissue cells. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 96(674):207–211, 1924.
- [11] T. S. Deisboeck and G. S. Stamatakos. *Multiscale Cancer Modeling*. Chapman & Hall/Crc Mathematical and Computational Biology Series. Taylor & Francis, 2010.
- [12] B. Douglas and J. Fowler. The effect of multiple small doses of X rays on skin reactions in the mouse and a basic interpretation. *Radiation Research*, 66(2):401–426, 1976.
- [13] M. M. Elkind and G. F. Whitmore. *Radiobiology of cultured mammalian cells*. Gordon and Breach, 1967.
- [14] K. C. Estacio-Hiros, E. E. Prudencio, N. P. Malaya, M. Vohra, and D. McDougall. The QUESO Library, User’s Manual. *arXiv preprint arXiv:1611.07521*, 2016.
- [15] D. J. Eyre. Unconditionally gradient stable time marching the Cahn-Hilliard equation. In J. W. Bullard and L.-Q. Chen, editors, *Computational and Mathematical Models of Microstructural Evolution*, volume 529 of *MRS Proceedings*, pages 39–46. Cambridge University Press, November 1998.
- [16] K. Farrell, J. T. Oden, and D. Faghihi. A Bayesian framework for adaptive selection, calibration, and validation of coarse-grained models of atomistic systems. *Journal of Computational Physics*, 295:189–208, 2015.
- [17] M. E. Gurtin. Generalized Ginzburg-Landau and Cahn-Hilliard equations based on a microforce balance. *Physica D: Nonlinear Phenomena*, 92(3-4):178–192, 1996.
- [18] D. Hanahan and R. A. Weinberg. The hallmarks of cancer. *Cell*, 100:57–70, 2000.
- [19] D. Hanahan and R. A. Weinberg. Hallmarks of cancer: The next generation. *Cell*, 144:646 – 674, 2011.
- [20] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
- [21] A. Hawkins-Daarud, S. Prudhomme, K. G. van der Zee, and J. T. Oden. Bayesian calibration, validation, and uncertainty quantification of diffuse interface models of tumor growth. *Journal of Mathematical Biology*, 67(6-7):1457–1485, 2013.
- [22] D. A. Hormuth II, J. A. Weis, S. L. Barnes, M. I. Miga, E. C. Rericha, V. Quaranta, and T. E. Yankeelov. Predicting in vivo glioma growth with the reaction diffusion equation constrained by quantitative magnetic resonance imaging data. *Physical Biology*, 12(4), 2015.
- [23] D. A. Hormuth II, J. A. Weis, S. L. Barnes, M. I. Miga, E. C. Rericha, V. Quaranta, and T. E. Yankeelov. A mechanically coupled reaction-diffusion model that incorporates intra-tumoural heterogeneity to predict in vivo glioma growth. *Journal of The Royal Society Interface*, 14(128):20161010, 2017.
- [24] D. A. Hormuth II, J. A. Weis, S. B. Eldridge, M. I. Miga, E. C. Rericha, V. Quaranta, and T. E. Yankeelov. Abstract a09: Predicting response to whole brain radiotherapy in a murine model of glioma. 77(2):A09 LP-09, 2017.
- [25] E. T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, Cambridge, 2003.
- [26] A. M. Kellerer and H. H. Rossi. The theory of dual radiation action. *Current Topics in Radiation Research Quarterly*, pages 85–158, 1974.
- [27] B. S. Kirk, J. W. Peterson, R. H. Stogner, and G. F. Carey. libMesh: A C++ Library for Parallel Adaptive Mesh Refinement/Coarsening Simulations. *Engineering with Computers*, 22(3-4):237–254, 2006.
- [28] E. A. B. F. Lima, J. T. Oden, and R. C. Almeida. A Hybrid Ten-Species Phase-Field Model of Tumor Growth. *Mathematical Models and Methods in Applied Sciences*, 24(13):2569–2599, 2014.
- [29] E. A. B. F. Lima, J. T. Oden, D. A. Hormuth II, T. E. Yankeelov, and R. C. Almeida. Selection, calibration, and validation of models of tumor growth. *Mathematical Models and Methods in Applied Sciences*, 26(12):2341–2368, 2016.

- [30] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE transactions on medical imaging*, 16(2):187–198, 1997.
- [31] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- [32] J. T. Oden, A. Hawkins, and S. Prudhomme. General diffuse-interface theories and an approach to predictive tumor growth modeling. *Mathematical Models and Methods in Applied Sciences*, 20(3):477–517, 2010.
- [33] J. T. Oden, I. Babuska, and D. Faghihi. Predictive Computational Science: Computer Predictions in the Presence of Uncertainty. In *Encyclopedia of Computational Mechanics*. John Wiley & Sons, to appear.
- [34] J. T. Oden. Foundations of Predictive Computational Sciences. *ICES Reports*, 2017.
- [35] J. T. Oden, E. A. B. F. Lima, R. C. Almeida, Y. Feng, M. N. Rylander, D. Fuentes, D. Faghihi, M. M. Rahman, M. DeWitt, M. Gadde, and J. C. Zhou. Toward Predictive Multiscale Modeling of Vascular Tumor Growth. *Archives of Computational Methods in Engineering*, 23(4):735–779, 2016.
- [36] A. R. Padhani, G. Liu, D. Mu-Koh, T. L. Chenevert, H. C. Thoeny, T. Takahara, A. Dzik-Jurasz, B. D. Ross, M. Van Cauteren, D. Collins, et al. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. *Neoplasia*, 11(2):102–125, 2009.
- [37] E. E. Prudencio and K. W. Schulz. The Parallel C++ Statistical Library 'QUESO': Quantification of Uncertainty for Estimation, Simulation and Optimization. In M. Alexander, P. D'Ambra, A. Belloum, G. Bosilca, M. Cannataro, M. Danelutto, B. Di Martino, M. Gerndt, E. Jeannot, R. Namyst, J. Roman, S. Scott, J. Traff, G. Vallée, and J. Weidendorfer, editors, *Euro-Par 2011: Parallel Processing Workshops*, volume 7155 of *Lecture Notes in Computer Science*, pages 398–407. Springer Berlin Heidelberg, 2012.
- [38] E. Prudencio and S. H. Cheung. Parallel adaptive multilevel sampling algorithms for the Bayesian analysis of mathematical models. *International Journal for Uncertainty Quantification*, 2(3):215–237, 2012.
- [39] R. Rockne, E. Alvord, J. K. Rockhill, and K. Swanson. A mathematical model for brain tumor response to radiation therapy. *Journal of Mathematical Biology*, 58(4-5):561, 2009.
- [40] R. H. Swendsen and J.-S. Wang. Replica Monte Carlo simulation of spin-glasses. *Physical Review Letters*, 57(21):2607, 1986.
- [41] J. A. Weis, M. I. Miga, L. R. Arlinghaus, X. Li, A. B. Chakravarthy, V. Abramson, J. Farley, and T. E. Yankeelov. A mechanically coupled reaction–diffusion model for predicting the response of breast tumors to neoadjuvant chemotherapy. *Physics in medicine and biology*, 58(17):5851, 2013.
- [42] J. A. Weis, M. I. Miga, L. R. Arlinghaus, X. Li, V. Abramson, A. B. Chakravarthy, P. Pendyala, and T. E. Yankeelov. Predicting the response of breast cancer to neoadjuvant therapy using a mechanically coupled reaction–diffusion model. *Cancer research*, 75(22):4697–4707, 2015.
- [43] J. A. Weis, M. I. Miga, and T. E. Yankeelov. Three-dimensional image-based mechanical modeling for predicting the response of breast cancer to neoadjuvant therapy. *Computer Methods in Applied Mechanics and Engineering*, 314:494–512, 2017.
- [44] T. E. Yankeelov and J. C. Gore. Dynamic contrast enhanced magnetic resonance imaging in oncology: theory, data acquisition, analysis, and examples. *Current Medical Imaging Reviews*, 3(2):91–107, 2009.
- [45] T. E. Yankeelov, N. Atuegwu, D. Hormuth II, J. A. Weis, S. L. Barnes, M. I. Miga, E. C. Rericha, and V. Quaranta. Clinically Relevant Modeling of Tumor Growth and Treatment Response. *Science Translational Medicine*, 5:187ps9, 2013.
- [46] T. E. Yankeelov, V. Quaranta, K. J. Evans, and E. C. Rericha. Toward a Science of Tumor Forecasting for Clinical Oncology. *Cancer Research*, 75(6):918–923, 2015.