

Cancer Biology & Therapy

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/kcbt20

A mathematical model for predicting the spatiotemporal response of breast cancer cells treated with doxorubicin

Hugo J. M. Miniere, Ernesto A. B. F. Lima, Guillermo Lorenzo, David A. Hormuth II, Sophia Ty, Amy Brock & Thomas E. Yankeelov

To cite this article: Hugo J. M. Miniere, Ernesto A. B. F. Lima, Guillermo Lorenzo, David A. Hormuth II, Sophia Ty, Amy Brock & Thomas E. Yankeelov (2024) A mathematical model for predicting the spatiotemporal response of breast cancer cells treated with doxorubicin, *Cancer Biology & Therapy*, 25:1, 2321769, DOI: [10.1080/15384047.2024.2321769](https://doi.org/10.1080/15384047.2024.2321769)

To link to this article: <https://doi.org/10.1080/15384047.2024.2321769>



© 2024 The Author(s). Published with
license by Taylor & Francis Group, LLC.



[View supplementary material](#)



Published online: 27 Feb 2024.



[Submit your article to this journal](#)



Article views: 3050



[View related articles](#)



[View Crossmark data](#)



Citing articles: 1 [View citing articles](#)

A mathematical model for predicting the spatiotemporal response of breast cancer cells treated with doxorubicin

Hugo J. M. Miniere^a, Ernesto A. B. F. Lima^{ID b}, Guillermo Lorenzo^{ID b,c}, David A. Hormuth II^{ID b,d}, Sophia Ty^{a,b}, Amy Brock^{ID a,d}, and Thomas E. Yankeelov^{ID a,b,d,e,f,g}

^aDepartment of Biomedical Engineering, The University of Texas at Austin, Austin, USA; ^bOden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, USA; ^cDepartment of Civil Engineering and Architecture, University of Pavia, Lombardy, Italy; ^dLivestrong Cancer Institutes, The University of Texas at Austin, Austin, USA; ^eDepartment of Diagnostic Medicine, The University of Texas at Austin, Austin, USA; ^fDepartment of Oncology, The University of Texas at Austin, Austin, USA; ^gDivision of Diagnostic Imaging, The University of Texas M.D. Anderson Cancer Center, Houston, USA

ABSTRACT

Tumor heterogeneity contributes significantly to chemoresistance, a leading cause of treatment failure. To better personalize therapies, it is essential to develop tools capable of identifying and predicting intra- and inter-tumor heterogeneities. Biology-inspired mathematical models are capable of attacking this problem, but tumor heterogeneity is often overlooked in *in-vivo* modeling studies, while phenotypic considerations capturing spatial dynamics are not typically included in *in-vitro* modeling studies. We present a data assimilation-prediction pipeline with a two-phenotype model that includes a spatiotemporal component to characterize and predict the evolution of *in-vitro* breast cancer cells and their heterogeneous response to chemotherapy. Our model assumes that the cells can be divided into two subpopulations: **surviving cells unaffected by the treatment, and irreversibly damaged cells** undergoing treatment-induced death. MCF7 breast cancer cells were previously cultivated in wells for up to 1000 hours, treated with various concentrations of doxorubicin and imaged with time-resolved microscopy to record spatiotemporally-resolved cell count data. Images were used to generate cell density maps. Treatment response predictions were initialized by a training set and updated by weekly measurements. Our mathematical model successfully calibrated the spatiotemporal cell growth dynamics, achieving median [range] concordance correlation coefficients of $>.99$ [.88, $>.99$] and .73 [.58, .85] across the whole well and individual pixels, respectively. Our proposed data assimilation-prediction approach achieved values of .97 [.44, $>.99$] and .69 [.35, .79] for the whole well and individual pixels, respectively. Thus, our model can capture and predict the spatiotemporal dynamics of MCF7 cells treated with doxorubicin in an *in-vitro* setting.

ARTICLE HISTORY

Received 11 May 2023
Revised 19 September 2023
Accepted 18 February 2024

KEYWORDS

Time-resolved microscopy;
data assimilation;
mechanism-based modeling;
mathematical oncology

Introduction

Chemoresistance, the ability of cancer cells to evade or endure the cytotoxic effects of drugs and therapies, is a fundamental contributing factor to treatment failure and is associated with poor prognosis, disease recurrence, and metastasis in patients.^{1,2} Driven by a variety of phenomena such as phenotypic resistance (i.e., phenotypic traits blocking drugs' mechanisms of action), altered metabolism, or genetic changes caused by the treatment itself,^{3–6} a key component of chemoresistance and tumor relapse following therapy lies in the heterogeneous genetic and phenotypic structure of solid neoplasms, allowing regrowth from even a small number of resistant cells.^{4,7–9}

Both genetic and non-genetic factors contribute to the development of intra- and intertumoral heterogeneity. Among the latter, the degree to which stromal cells infiltrate the microenvironment as well as how new vasculature is recruited to the tumor may result in a heterogeneous delivery of both nutrients and treatments. This heterogeneity can lead

to the formation of habitats; i.e., subregions within the tumor that are distinguished by unique phenotypic signatures and properties.^{10–12} The spatially varying habitats contribute to varied treatment response across the tumor as the local environment can affect both the delivery and efficacy of the selected treatment.¹³ Thus, the design of treatment protocols accounting for these intra- and inter-tumoral variations would enable the optimal improvement of therapeutic outcomes for each individual patient. To achieve this goal, it is fundamental to leverage a rigorous methodology that explicitly accounts for spatiotemporal heterogeneities in individual tumors.¹⁴

Mathematical models that account for the underlying biology of the spatial and temporal development of tumors – as well as their response to therapy – have been shown to accurately predict the response of individual patient tumors.^{15–17} In particular, there has been much progress in using imaging-based measurements to calibrate and constrain model predictions in both the pre-clinical and clinical settings.^{18–28} Importantly, it is difficult to develop and validate biology

CONTACT Thomas E. Yankeelov  tey@utexas.edu  Department of Biomedical Engineering, The University of Texas at Austin, 107 W Dean Keeton Street Stop C0800, Austin, TX 78712, USA

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/15384047.2024.2321769>

© 2024 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

based mathematical models in the clinical setting where there are fundamental limitations to the amount and quality of imaging data that can be acquired on an individual patient. Furthermore, **models that account for different cell types that can be calibrated to available data are rare**^{29–31}; and although studies employing compartmental models have been developed over the past decade,^{32–34} phenotypic heterogeneities are not typically incorporated into mathematical modeling studies of clinical data. **Models that seek to capture the dynamics of multiple interacting phenotypes are typically investigated in *in vitro* or *in silico* settings.** Furthermore, such models typically **disregard the spatial dynamics** and heterogeneity at play in cells cultivated *in vitro*; rather, they emphasize characterizing the **average cell population growth** observed across the whole assay.³⁵

Time-resolved microscopy is a high-resolution imaging technique that captures images of biological samples at specific intervals over extended durations with minimal intervention. It allows frequent, noninvasive, cell-preserving imaging, and therefore does not suffer from the data scarcity issues of clinical studies. For example, A coupled model of pharmacokinetics and pharmacodynamics was created and validated by McKenna *et al.* and was able to forecast the reaction of a uniform cell population to a treatment schedule that changed over time. The researchers used fluorescence microscopy over a period of time to examine the absorption of doxorubicin and how the cell population responded to this medication across a range of triple negative breast cancer cell lines.³⁶ Tyson *et al.* calibrated a model using immunofluorescence labeling and time-lapse automated microscopy, allowing them to estimate the lifespan of individual lung cancer cells.³⁷ Lima *et al.* used time-resolved microscopy with a phase-field tumor growth model to calibrate rates of proliferation, apoptosis, necrosis and mobility on *in vitro* human liver carcinoma cell subpopulations.³⁸ Additionally, by raster scanning an entire well-plate, dynamic spatial information on cell distribution can also be obtained over time. Most relevant to our work, Howard *et al.* employed time-resolved microscopy to image several breast cancer cell lines treated with various doses and regimens of doxorubicin, and then developed a family of mathematical models to characterize their growth and response to therapy.³⁹ Expanding on this approach, Yang *et al.* constructed a set of ordinary differential equations that were able to recapitulate each treatment regimen regardless of the concentration or number of doxorubicin doses.¹⁸ In particular, Yang *et al.* modeled *in vitro* breast cancer colonies treated with doxorubicin as compartmentalized subpopulations of resistant and irreversibly damaged cells, allowing for phenotypic switching of cells in response to the application of an additional dose. While successfully capturing the temporal evolution of MCF7 cancer cells and their response to a variety of treatment regimens, their study did not include spatial considerations of cell distribution. Moreover, they did not establish a pipeline allowing for the prediction of growth and treatment response, a key step of crucial clinical importance.

In this contribution, **we seek to broaden the work of Yang *et al.* in two directions.** First, we incorporate a spatial component to their modeling framework that aims at explicitly capturing heterogeneities in growth and response to

treatment in terms of tumor cell density. **Second, we develop a computational pipeline** to recapitulate and predict of the spatiotemporal heterogeneous dynamics of tumor cell density with our model. Toward this end, we leverage the data assimilation methodology proposed by Liu *et al.* to provide model updates as more data becomes available from time-resolved microscopy measurements.⁴⁰ Then, **we further demonstrate that our spatiotemporal modeling approach can recapitulate and forecast tumor cell density dynamics in a two-dimensional *in vitro* setting** by leveraging the same dataset of MCF7 cells treated with doxorubicin used by Yang *et al.* Once a mathematical model has been shown to accurately predict the spatial and temporal development of a tumor, it can potentially be used to predict and optimize therapeutic responses. If the model is accurate, then one can simulate a range of interventions and select the one that provides the greatest tumor control.^{41–43} For example, in the present case, one could use our pipeline to identify a doxorubicin treatment schedule to maintain tumor control for as long as possible.

The remainder of this study is organized as follows. We first summarize the experimental methods and derivation of the model introduced by Yang *et al.*¹⁸ We then describe the model calibration approach, how the initial predictions are made, and how we update our forecasts during the experiment. We conclude with a discussion of the significance of the study, its limitations, and possible directions for future investigation.

Materials and methods

Cell culture and experimental conditions

As the data employed in this study have been previously published, we highlight only the salient details. (For a complete description of the experimental methods, please see ref. 39) ER-positive MCF7 breast cancer cells (ATCC HTB-22) were grown in a Minimum Essential Media (Gibco)/10% fetal bovine serum (Gibco)/1% penicillin-streptomycin (Gibco) mix and maintained at 37°C in a 5% CO₂ incubator. Cells were genetically modified to express EGFP (enhanced green fluorescence protein) along with a nuclear localization signal (MCF7-EGFPNLS1) integrated using the Sleeping Beauty transposon system. The EGFP-NLS sequence, obtained as a gBlock (IDT), was cloned into the optimized Sleeping Beauty transfer vector pSBbi-Neo (gifted by Eric Kowarz, Addgene plasmid #60525).⁴⁴ Finally, the two-plasmid system was transfected into the cell population using Lipofectamine 2000.⁴⁵ EGFP-expressing cells were then collected through fluorescence activated cell sorting and cultured in G418 (Caisson Labs)-supplemented media (200 ng/ml).

Cells were then seeded onto a 96-well plate at 2000 cells per well and allowed to grow untreated in 100 µl of culture media for 48 hours. Next, 100 µl of media containing doxorubicin hydrochloride (Cayman Chemical 15,007, Ann Arbor, Michigan) at 2× the desired final concentration was added into each well for a 24-hour duration and then washed with fresh growth media containing no drug. Nine different doxorubicin concentrations were investigated: 10, 20, 35, 50, 75, 100,

125, 150, and 300 nM with $n = 6$ replicates per concentration, along with six untreated replicates as a control.

Data gathering and image processing

Fluorescent and phase contrast images (Figure 1(a)) were collected every 2–4 hours using an IncuCyte S2 Live Cell Analysis System (Essen/Sartorius, Goettingen, Germany) for 1000 hours. Cell quantification and centroid positions were gathered through background subtraction, thresholding, edge detection and minimum area filtering. The resulting time courses of cell count were truncated once confluence was reached.¹⁸ Cell centroid positions were used to generate 2D tumor cell density maps (see Figure 1(a)), which are then used to calibrate the mathematical model described in the next section. The resolution of the cell maps was limited by signal-to-noise considerations and chosen to be $430 \mu\text{m} \times 430 \mu\text{m}$, corresponding to a grid size of 15×15 pixels over a $6450 \mu\text{m} \times 6450 \mu\text{m}$ field-of-view of each well.

Mathematical model

The model used in this study corresponds to a two-dimensional reaction-diffusion extension of the single-dose model presented by Yang *et al.* (Figure 1(b)).^{18,46} Such models

are routinely used in cancer modeling. Hence, our model formulation accounts for cell migration and characterizes spatial heterogeneities arising both in the pre- and post-treatment phase of the experiment.

Prior to exposure to doxorubicin, the tumor cell population is assumed to be phenotypically homogeneous, expanding by diffusion, and growing freely according to exponential growth:

$$\frac{\partial N_{tot}(\bar{x}, t)}{\partial t} = .(D(\bar{x}) N_{tot}(\bar{x}, t)) + g(\bar{x})N_{tot}(\bar{x}, t) \quad (1)$$

$$N_{tot}(\bar{x}, 0) = N_0(\bar{x}), \quad (2)$$

where $N_{tot}(\bar{x}, t)$ is the total tumor cell count at location \bar{x} and time t , and $D(\bar{x})$ and $g(\bar{x})$ are the spatially defined diffusion coefficient and proliferation rate of the untreated cells, respectively. To reflect the homogeneous nature of the cell population in the absence of treatment, we consider that the cells diffuse and grow independent of their position \bar{x} ; thus, we set $D(\bar{x}) \equiv D_0$, and $g(\bar{x}) \equiv g_0$.

After doxorubicin treatment is applied at time $t_{dox} = 48$ hours, the previously homogeneous cell population (as described by Equations (1) and (2)) is now modeled as two subpopulations distinguishable by their reaction to the drug: a “surviving” subpopulation, with cell count $N_s(\bar{x}, t)$, that will evade doxorubicin-induced death, and an “irreversibly

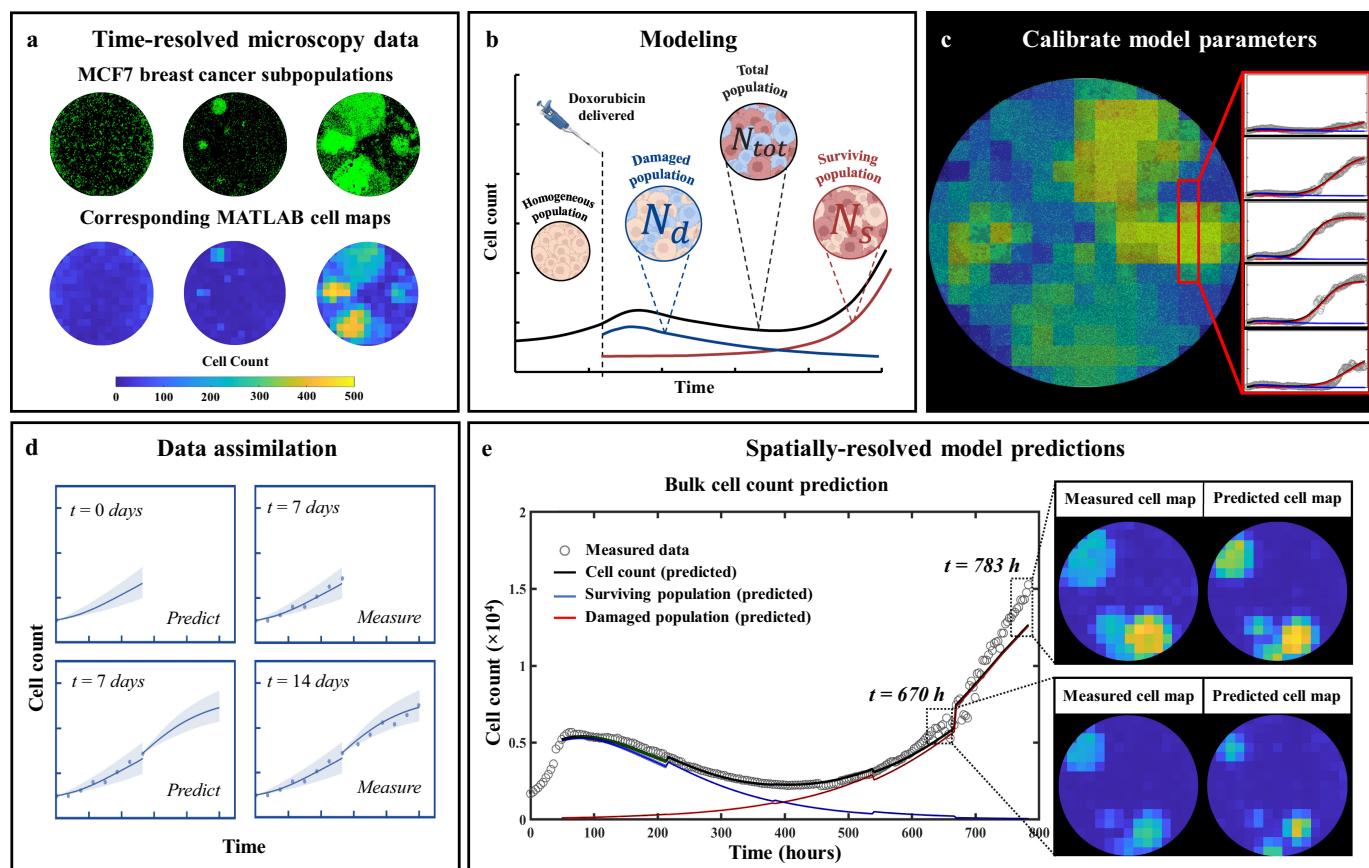


Figure 1. General flowchart of predicting MCF7 growth through data assimilation. (a) MCF7 breast cancer cells are grown and treated with varying doses of doxorubicin. Fluorescent microscopy is used to gather images every four hours which are converted into cell density maps. (b) A two-subpopulation model is used to describe the spatial and temporal development of MCF7 cells. (c) The model is calibrated on the cell density maps, yielding parameter values for each individual pixel. (d) We use a data assimilation pipeline consisting of successive short-term predictions guided by serial measurements. e. Cell density maps and bulk cell counts (i.e., total cell count across the entire well) are predicted at weekly intervals.

damaged” subpopulation, with cell count $N_d(\bar{x}, t)$, that will eventually die after a short period of growth. The fractions of the surviving and irreversibly damaged subpopulations at time $t = t_{dox}$ are defined as f_s and $(1 - f_s)$, respectively.

Thus, we have:

$$N_s(\bar{x}, t_{dox}) = f_s(\bar{x})N_{tot}(\bar{x}, t_{dox}), \quad (3)$$

$$N_d(\bar{x}, t_{dox}) = (1 - f_s(\bar{x}))N_{tot}(\bar{x}, t_{dox}), \quad (4)$$

with the total cell count in the post-treatment phase given by:

$$N_{tot}(\bar{x}, t) = N_s(\bar{x}, t) + N_d(\bar{x}, t). \quad (5)$$

The surviving subpopulation is assumed to follow logistic growth within a reaction-diffusion equation:

$$\frac{\partial N_s(\bar{x}, t)}{\partial t} = \cdot (D_s(\bar{x}) \cdot N_s(\bar{x}, t)) + g_s(\bar{x}) \left(1 - \frac{N_{tot}(\bar{x}, t)}{\theta} \right), \quad (6)$$

where $g_s(\bar{x})$ is the post-treatment proliferation rate, $D_s(\bar{x})$ the diffusion coefficient of surviving cells (where we again assumed that the cells diffuse independent of their position, i.e., $D_s(\bar{x}) = D_s$), and θ the carrying capacity (i.e., the maximum admissible cell density). The irreversibly damaged subpopulation is assumed to initially follow logistic growth, which progressively transitions into treatment-induced death as continued drug uptake triggers DNA damage and cell cycle disruption leads to cell death^{36,47–49}:

$$\begin{aligned} \frac{\partial N_d(\bar{x}, t)}{\partial t} = & \cdot (D_d(\bar{x}) \cdot N_d(\bar{x}, t)) \\ & + ((g_d(\bar{x}) + k_d(\bar{x})) \exp(-\gamma_d(\bar{x})(t - t_{dox})) \\ & - k_d(\bar{x}))N_d(\bar{x}, t) \left(1 - \frac{N_{tot}(\bar{x}, t)}{\theta} \right), \end{aligned} \quad (7)$$

where $g_d(\bar{x})$ and $k_d(\bar{x})$ respectively represent the proliferation and doxorubicin-induced death rates, $\gamma_d(\bar{x})$ denotes the doxorubicin-induced death delay rate (which characterizes the time it takes for the cells to transition from proliferating to dying following drug exposure), $D_d(\bar{x})$ is the diffusion coefficient of the damaged cells (also assumed independent of position; i.e., $D_d(\bar{x}) = D_d$), and θ is the same carrying capacity as in Equation (6). Note that the pre-treatment population followed exponential growth (Equation (1)) while the post-treatment subpopulations are modeled using logistic growth (Equations (6) and (7)). This is explained by the low cell density in the pre-treatment phase, allowing a simplification from a logistic to an exponential form. As the cell count is susceptible to rise sharply with cells reaching maximum capacity in the post-treatment phase, a logistic model is now used.

Moreover, for simplicity, we assume that the untreated, surviving, and irreversibly damaged cells diffuse similarly regardless of their treatment status i.e., $D_s = D_d = D_0$. As such, diffusion is calibrated in pre-treatment phase only, and is then fixed post-treatment. Figure 1(b) presents a graphical representation of the model compartments during the two experiment phases, and Table 1 summarizes the model parameters. The partial differential equations were solved using the finite difference method over the pixel grid, with fully explicit time differentiation (time step $\Delta t = 1$ hr), along with central differences for spatial discretization and no flux boundary conditions.

Global and local parameters

The model presented in Equations (1), (6), and (7) contains eight parameters (i.e. g_0 , D_0 , g_s , f_s , g_d , k_d , γ_d , and θ). Given the number of pixels per cell map (see ‘Data gathering and image processing’ section above), one parameter value per pixel for each of the eight parameters would require calibrating several hundred parameters and almost certainly lead to overfitting. To avoid this issue, a global sensitivity analysis was carried out to determine the contribution of each model parameter on the spatiotemporal dynamics of tumor cells. This was performed by setting all parameters global and randomly sampling values from their respective parameter spaces (see Table 1) to generate 5000 model evaluations and calculate the total effect index for each parameter. The results (see Figure S1 of the Supplementary Materials) indicated that the two parameters having the greatest effect on the tumor cell density (i.e., small changes in their values leading to the greatest changes on the cell number) were the proliferation rate of the surviving subpopulation, g_s and the carrying capacity, θ . We therefore decided to calibrate all other parameters in a global fashion, meaning that a single parameter value will be calibrated over the whole well and shared by all pixels. Moreover, the carrying capacity θ is assumed to be a physical parameter characterizing the experimental conditions (e.g., type of cell, well dimensions), which may only exhibit narrow variations emanating from diversity in seeding, dosing, and ensuing growth. Thus, the carrying capacity θ was also calibrated globally over a narrow range of admissible values. The surviving growth rate g_s is thus the only parameter with a different value in each pixel. As will be shown in the ‘Results’ section below, this modeling choice suffices to capture the experimentally observed spatial heterogeneity.

Table 1. List of model parameters.

Parameter	Definition	Notes
g_0	Proliferation rate of untreated cells (hr^{-1})	Global, calibrated in pre-treatment phase (range: [.02 .035])
D_0	Diffusion coefficient ($\text{mm}^2/\text{hr}^{-1}$)	Global, calibrated in pre-treatment phase (range: [0.005])
g_s	Proliferation rate of surviving cells (hr^{-1})	Local, calibrated (range: [.0001 .035])
f_s	Surviving cell fraction at time of treatment	Global, calibrated (range: [0 1])
k_d	Drug-induced death rate of irreversibly-damaged cells (hr^{-1})	Global, calibrated (range: [.0001 .01])
g_d	Proliferation rate of irreversibly damaged cells (hr^{-1})	Global, calibrated (range: [.001 .05])
γ_d	Drug-induced death delay rate (hr^{-1})	Global, calibrated (range: [1/120 1/10])
θ	Carrying capacity	Global, calibrated within a narrow range (range: [88000 90,000])

Model fitting

We fit our model to each replicate from the time-resolved data obtained as described above in the ‘*Cell culture and experimental conditions*’ section. Model calibration (Figure 1(c)) was performed using a trust-region reflective algorithm via MATLAB’s (R2021a, The Mathworks, Natick, MA) nonlinear least-squares fitting function, *lsqnonlin*. Parameter bounds and initial guesses were inspired by ref. 18. Calibration precision is evaluated using a bootstrapping approach (100 bootstrap samples). This is performed by calibrating the model to the data (yielding a first parameter set), and using the calibration residual as an estimate of the normally distributed noise. A new synthetic dataset can now be created by sampling a new residual from the noise distribution and adding it to the model calibration. The model can now be calibrated onto the synthetic dataset, yielding another parameter set. This process is repeated 100 times, providing normal distributions for each parameter corresponding to the calibration precision as quantified by calculating the interquartile range.

Prediction scheme

Predictions of the spatiotemporal tumor cell dynamics were performed through a leave-one-out approach employing a data assimilation pipeline inspired by Liu *et al.*³⁷ (Figure 1(d,e)). In brief, with this method we seek to make short-term predictions which are then updated at regular intervals as more data becomes available for the replicate whose dynamics we are predicting. As we are most interested in accounting for spatial heterogeneities, and these are not observed prior to treatment (i.e., nearly no inter- or intra-well heterogeneity), the pre-treatment phase is excluded from our prediction approach; that is, our predictions begin at the time the drug is applied. We next describe the details of each step in our data assimilation-prediction approach.

For the $n = 6$ replicates receiving the same dose of doxorubicin, a training set of $n_{\text{training}} = 5$ replicates is selected, and the model is calibrated to each of these five datasets using the whole time course of measurements collected during the experiment. The five resulting parameter sets are then averaged into a parameter set P_{training} ($15 \times 15 \times 6$ matrix, corresponding to the grid size and number of post-treatment model parameters). The remaining replicate is placed in the testing set ($n_{\text{testing}} = 1$). For this replicate, the prediction of the dynamics of the tumor cell population follows a series of subsequent data assimilation-prediction steps over time periods of one week. During the first week after treatment, we run the model forward using the cell count of that replicate at the time the treatment is applied as initial conditions and leveraging the parameter set P_{training} (Figure 2(a,b)). The first week of post-treatment data is then measured for the replicate, allowing calibration of Equation (6)–(7) and yielding parameter set P_{measure} (Figure 2(c)). An individualized prediction over the subsequent week initialized using the last time point measured can now be performed on each pixel through the following three steps. First, for each pixel p within the testing set, the $5 \times 15 \times 15$ pixels in the training set are ranked based on the similarity of their time course of cell number to the time course observed in p (from the time of the treatment to the last

available time point in p). The similarity measure is the concordance correlation coefficient (CCC).⁵⁰ Second, the model parameters associated with the pixels in the training set with the 25% most similar time courses to that of pixel p in the testing set are averaged (in a weighted fashion) to form the parameter set $P_{\text{training},p}$ according to their degree of similarity:

$$P_{\text{training},p} = \sum_{k=1}^N \alpha_k P_k, \quad (8)$$

where $N = 282$ (i.e., $.25 \times 5 \times 15 \times 15$), P_k the calibrated parameter set of pixel k from the training set, and α_k the weight assigned to this pixel. The weights α_k are determined by dividing the CCC of the time course of pixel k by the sum of the CCCs of all 282 pixels, that is:

$$\alpha_k = \frac{\text{CCC}_k}{\sum_{j=1}^N \text{CCC}_j}. \quad (9)$$

Note that, as per Eq. (8), all the six post-treatment parameters used for the predictions can now vary pixel-wise, as opposed to the calibrated parameters of which only one (surviving proliferation rate g_s) varied pixelwise. This was necessary during the calibration step for reasons of overfitting mentioned above, but is no longer necessary when predicting weekly segments as this step only requires forward model evaluations. Third, P_{training} (now a new $15 \times 15 \times 6$ matrix containing all $P_{\text{training},p}$ from Eq. (8)) is weighted against P_{measure} to yield the final parameter set, $P_{\text{prediction}}$, allowing the prediction of the dynamics of the tumor cell population over the following week (Figure 2(d)):

$$P_{\text{prediction}} = \omega P_{\text{training}} + (1 - \omega) P_{\text{measure}}. \quad (10)$$

As time progresses and more data is acquired for the test dataset, ω decreases to give more weight to the calibrated parameter set, P_{measure} , which is specific to the replicate for which we are obtaining predictions of tumor cell dynamics. Inspired by ref. 37, we use:

$$\omega = 0.6\hat{t}^2 - 1.2\hat{t} + 1.0 \quad (11)$$

with \hat{t} the time in hours divided by the time at which the final time point is gathered on the longest experiment (963 hours) so that \hat{t} ranges from 0 to 1. The form of ω allows $\omega(0) = 1$ giving greater weight to the training set on the early time points, followed by a decrease shifting the weight to the individual replicate’s parameter set. The quadratic form was chosen because it decreases slower than a linear function during the second half of the experiment and is virtually unchanged in the neighborhood of the terminal timepoint (i.e., $\hat{t} = 1$), thus maintaining the weight of the training set approximately constant on the final time points. The three steps described above are then repeated each week until reaching the end of the experiment. Notice that the surviving fraction, f_s (previously defined as the fraction of surviving cells at time $t = t_{\text{dox}}$), is also updated through the prediction pipeline and is now used to split the cells at each weekly timepoint.

Note that, as we use bootstrapping to evaluate the precision of our model calibrations in the replicates in the training set and in each weekly calibration for the replicate in the testing set, we can obtain a distribution of values for

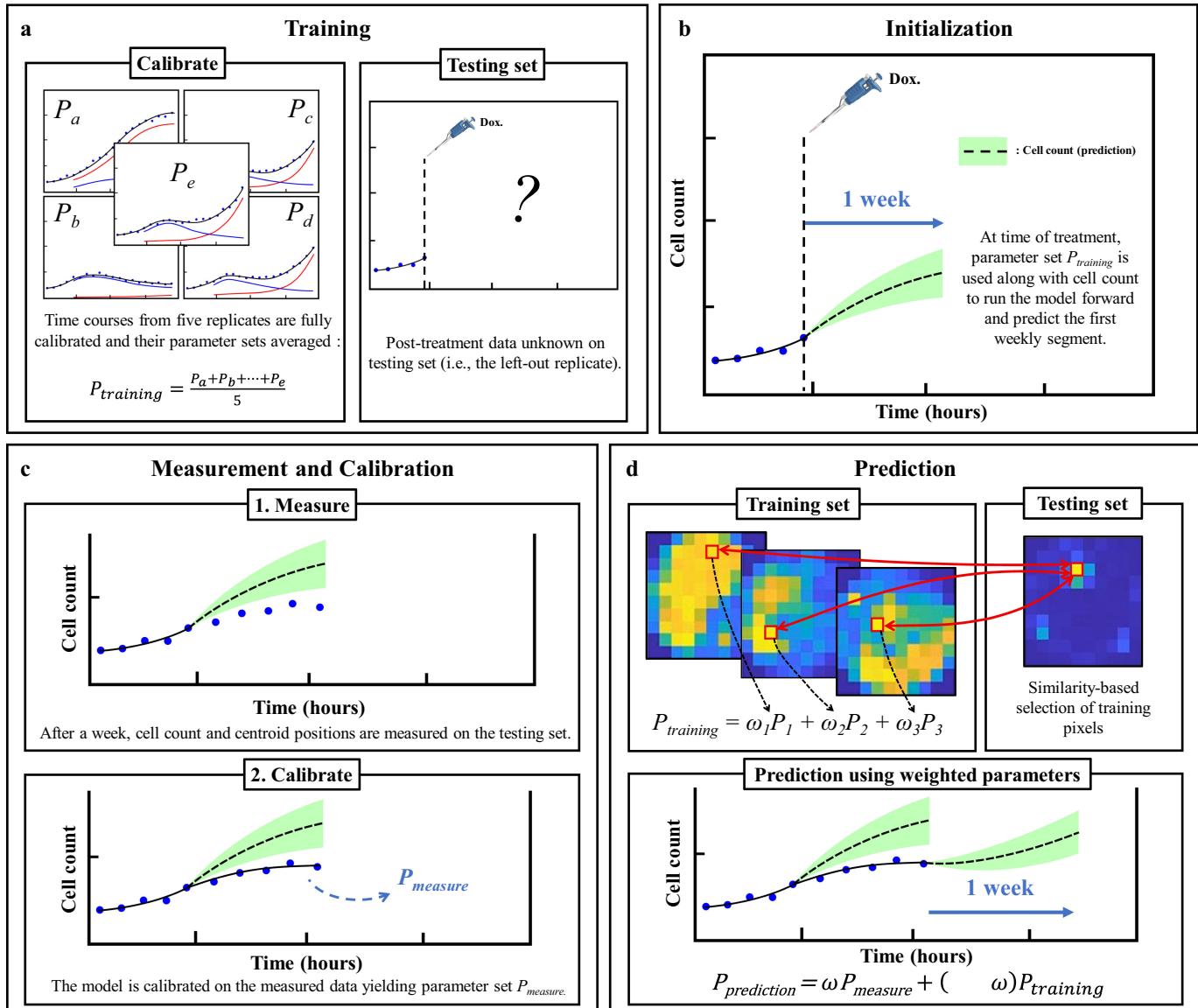


Figure 2. Predicting MCF7 cell growth and response to doxorubicin through leave-one-out data assimilation-prediction pipeline. (a) Out of six replicates treated with similar doxorubicin doses, five are selected to form the training set and fully calibrated. Their parameter sets are averaged into parameter set $P_{training}$, the post-treatment data on the left-out replicate are *a priori* “hidden”. (b) for the replicate in the testing set, the data at the time of the treatment are used to run the model forward with parameter set $P_{training}$, thereby yielding a first prediction. (c) After one week, the model is informed by the data collected during this timeframe and calibrated accordingly, thereby yielding parameter set $P_{measure}$. (d) to predict tumor cell dynamics over the subsequent week, we first estimate the new values of the model parameters based on the degree of similarity between the tumor cell dynamics in each pixel of the replicate in the testing set and all the pixels in the training set. This comparison yields $P_{training}$, which we combine with $P_{measure}$ to obtain the estimated parameter values for model prediction (i.e., $P_{prediction}$). The green shaded regions in panels b-d represent the bootstrap interquartile range of the predictions, around the dashed line representing the bootstrapped mean prediction.

each model parameter in $P_{training}$ and $P_{measure}$. Thus, combining these parameter distributions via Equation 10, we can also obtain the corresponding distributions in $P_{prediction}$. Then, we can sample the parameter distributions for $P_{prediction}$ and quantify the precision of our model predictions. Toward this end, we randomly draw 1000 samples and obtain an equivalent number of predictions of tumor cell dynamics during each week of the data assimilation-prediction pipeline as outlined above. In particular, we focus on analyzing the mean and interquartile range of the predicted tumor cell counts in each pixel of the replicate in the testing set.

Statistical analysis

To quantify the quality of the model calibrations and predictions we use the CCC (concordance correlation coefficient).⁴⁷ This quantity measures the agreement between the model calibration or prediction results and matching experimental data. We employ CCC to assess model performance in recapitulating and predicting tumor cell counts both at global and local level. Hence, we denote by CCC_{well} the use of this metric to compare the measured time course of the total cell count of the entire well to that obtained from the model calibration or prediction. Similarly, we denote by CCC_{pixel} the use of this metric to measure the quality of the local, pixel-wise calibration or prediction by calculating the average over the well of

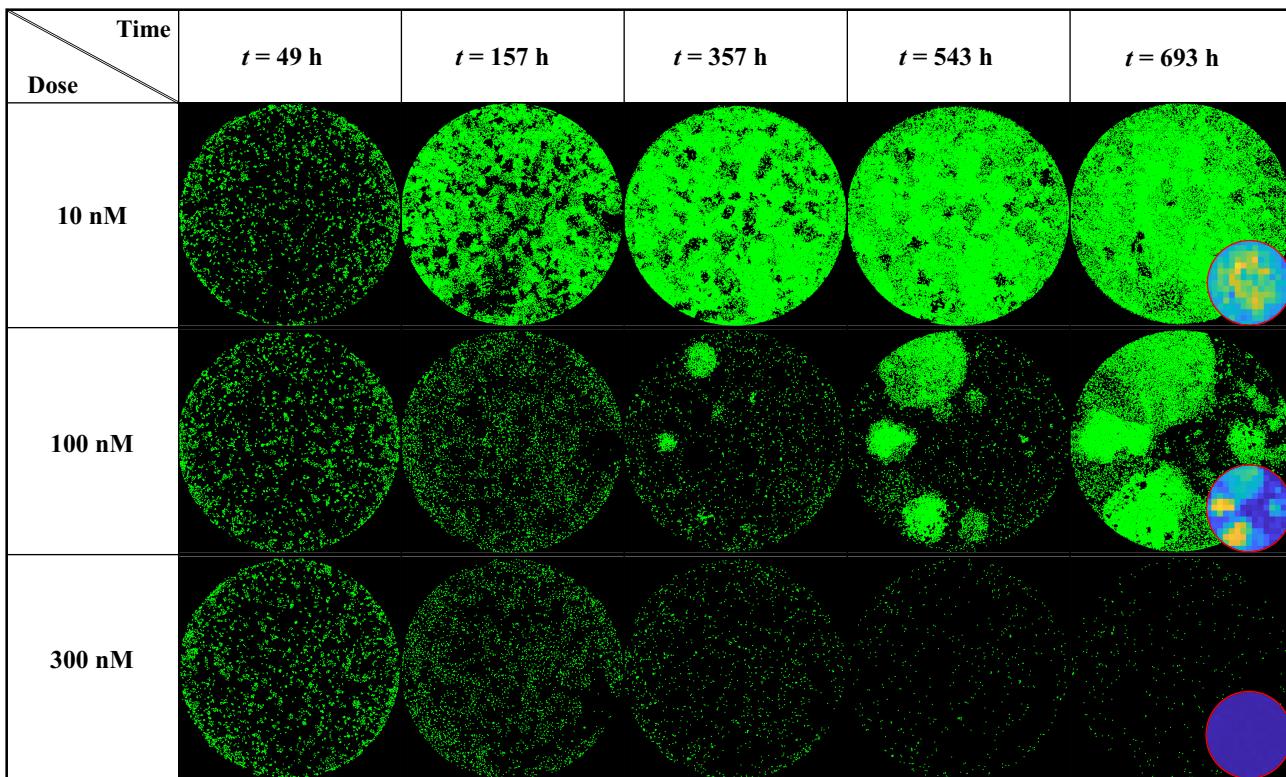


Figure 3. MCF7 breast cancer cells at various time points and treated with 10, 100 and 300 nM of Doxorubicin. Data show that low (10 nM) and high (300 nM) doses do not generate significant spatial heterogeneity post treatment, as opposed to a medium dosage where clusters of highly proliferating cells are responsible for uneven cell distribution after approximately two weeks of growth. Corresponding MATLAB tumor cell density maps can be seen on the last column.

the CCCs that compare each pixel's time course of total cell count to that obtained from model calibration or prediction.

Results

Fluorescent microscopy of MCF7 cells

Figure 3 shows a dish treated at low (10 nM), medium (100 nM), and high (300 nM) doses of doxorubicin at various time points (49, 157, 357, 543 and 693 hours after seeding), exhibiting clear patterns of spatial heterogeneity at a medium dosage. At a low dosage (10 nM, top row), cell growth is uniform across the well leading to cells covering the entirety of the dish without any distinctive pattern of spatial distribution. Thus, the effect of the treatment is negligible. At a high dosage (300 nM, bottom row), the powerful effect of the chemotherapy prevents any cell growth or recovery and the entire population homogeneously dies over the course of weeks. When a medium dosage (100 nM, middle row) is applied, rapid cell recovery is observed in localized areas forming clusters of highly proliferative cells responsible for uneven distribution. This intra-tumor spatial heterogeneity then translates into temporal inter-replicate heterogeneity, with each replicate showing distinct cell growth distribution patterns that lead to a diverse set of longitudinal time courses (see Supplemental Figure S2 from ref. 18 for complete dataset).

Fitting phenotypic model to MCF7 cells treated with doxorubicin

Figure 4 shows the tumor cell counts for representative replicates for each dose along with their associated model fit, demonstrating the ability of our spatial calibration to accurately characterize the dynamics of the bulk cell count (i.e., total cell count across the entire well) of treated MCF7 subpopulations. The calibration is performed for the whole duration of the experiment, and provides the model's estimate of total, surviving, and damaged tumor cell counts along with their respective precisions displayed as shaded areas showing greater calibration uncertainty on the damaged population's estimate. The quality of fit for the well and the pixels resulted in median and range $CCC_{\text{well}} > .99$ [.88, >.99] and $CCC_{\text{pixel}} = .73$ [.58, .85], respectively. Moreover, Figure 5(a) displays details of a local pixel calibration on a replicate treated with 100 nM of doxorubicin. Panel A shows the bulk cell counts obtained from the global fit, with $CCC_{\text{well}} = .99$. On panel B, the pixel map is superimposed on the fluorescence image at time $t = 782$ hours. Examples of calibrated pixels are shown, reflecting the model's ability handle a wide variety of temporal dynamics and capture the elevated heterogeneity found within these *in vitro* tumor cell populations ($CCC_{\text{pixel}} = .69$) with only a single parameter calibrated locally (i.e., g_s). The upper pixel exhibits a decay in cell count for approximately 500 hours before a slow cell recovery, contrasting with the middle pixel where a sharp proliferation phase begins around $t = 300$ hours. Lastly, the bottom pixel exhibits more unique dynamics, where a long decay of approximately 600 hours is followed by a sudden and

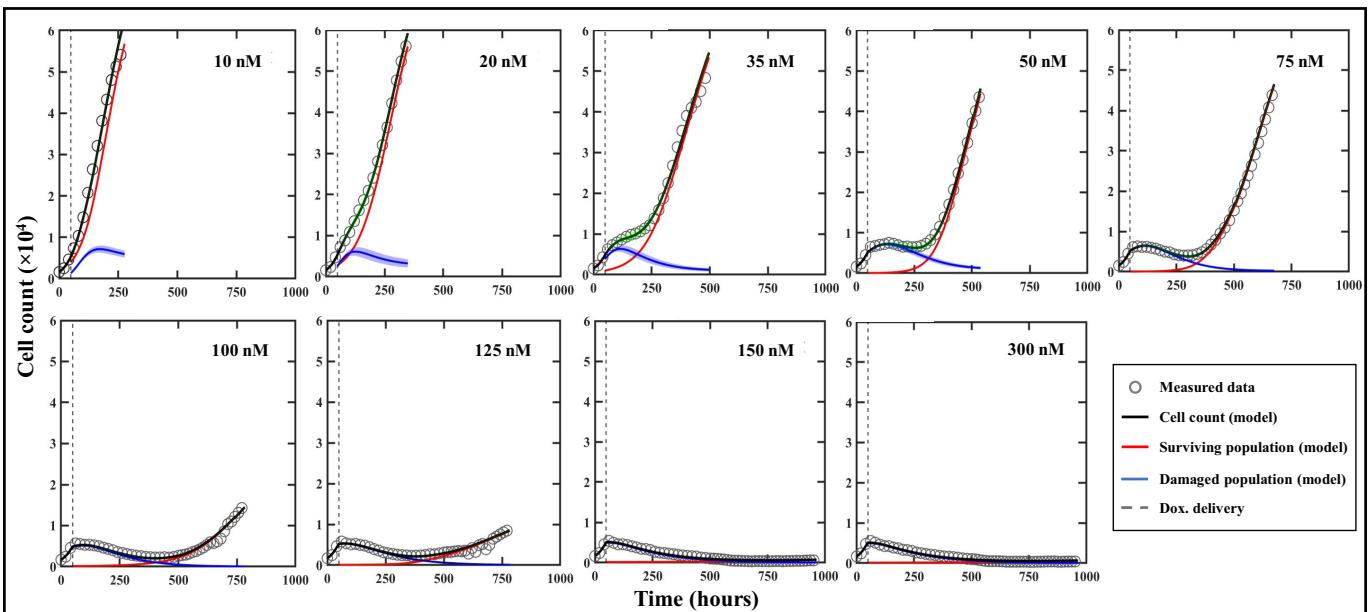


Figure 4. Model calibrations for representative replicates of MCF7 cells treated with varying doses of doxorubicin after 48 hours of growth. Measured bulk cell counts (gray circles) are shown along with the model's estimates of the surviving (red), damaged (blue) and total (black) populations and their respective precisions. For clearer visualization, only one out of four data points are shown. The vertical dashed line represents the time of delivery of doxorubicin.

brief increase in cell count, indicating a cell infiltration from neighboring pixels. As will be discussed below, such phenomena are poorly captured by the modeling approach.

Figure 6 provides an example of the bootstrap parameter distributions resulting from global model calibration in a representative replicate treated with 35 nM of doxorubicin. Panel A shows the time-resolved cell counts of the replicate, along with the corresponding fit obtained from the global model calibration. The result of model calibration on an individual pixel is shown on panel B with the associated bootstrapped parameter distributions on panel 6C. Note that, as all parameters are globally calibrated with the exception of g_s , the distributions for f_s , k_d , g_d , γ_d and θ are the same for all the pixels of each individual replicate. The normalized standard

deviations (standard deviations divided by the means, referred to as NSD) for those parameters indicate that the parameters with the largest spread are k_d (NSD = .074), g_d (NSD = .069) and γ_d (NSD = .11); that is, the parameters describing the damaged population and thus the effects of the treatment (NSDs for g_s , f_s and θ are respectively .011, .032 and .00021). This observation can be made across all the replicates and doses, and it is aligned with the larger bootstrap interquartile range of the damaged populations depicted on Figure 4.

Predictions of spatial tumor dynamics

Representative predictions of total cell counts obtained through our data assimilation-prediction pipeline are shown

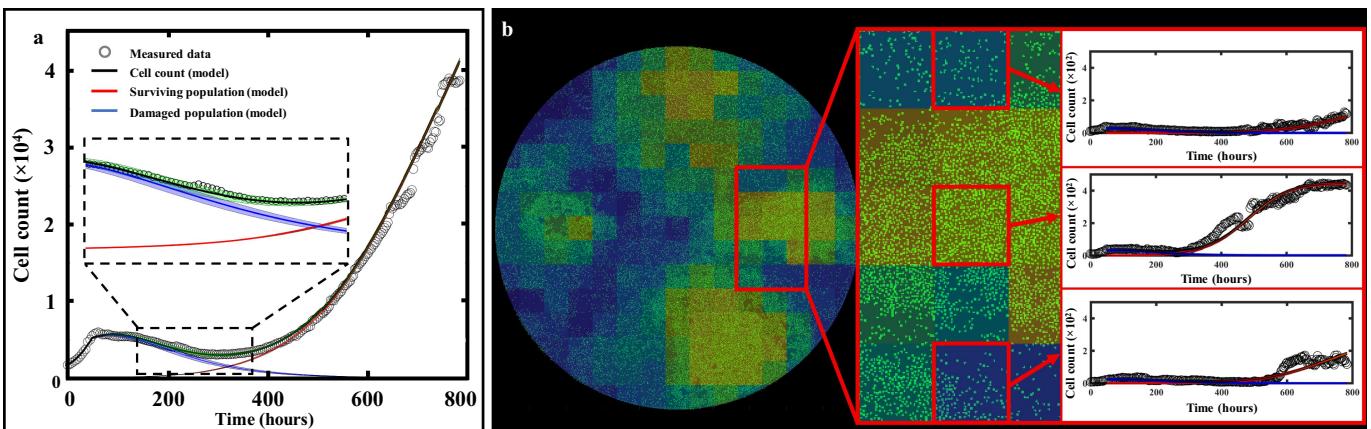


Figure 5. Spatial calibration of the two-subpopulation model on a replicate treated with 100 nM of doxorubicin. a. Bulk cell counts obtained for the global calibration of the model to the whole time course of microscopy measurements collected for the replicate, including the model's estimate of the surviving, damaged, and total population cell counts. The calibration precision is shown on all three populations via the shaded areas. b. Our model is calibrated to fit the time course on each pixel of the well, thereby capturing the spatial heterogeneity of the tumor cell population dynamics. The cell density map is shown superimposed on the corresponding cell microscopy image, taken at the final time point ($t = 782$ h). The plots on the right illustrate the local counts of total, surviving, and damaged cells for three representative pixels exhibiting different temporal dynamics.

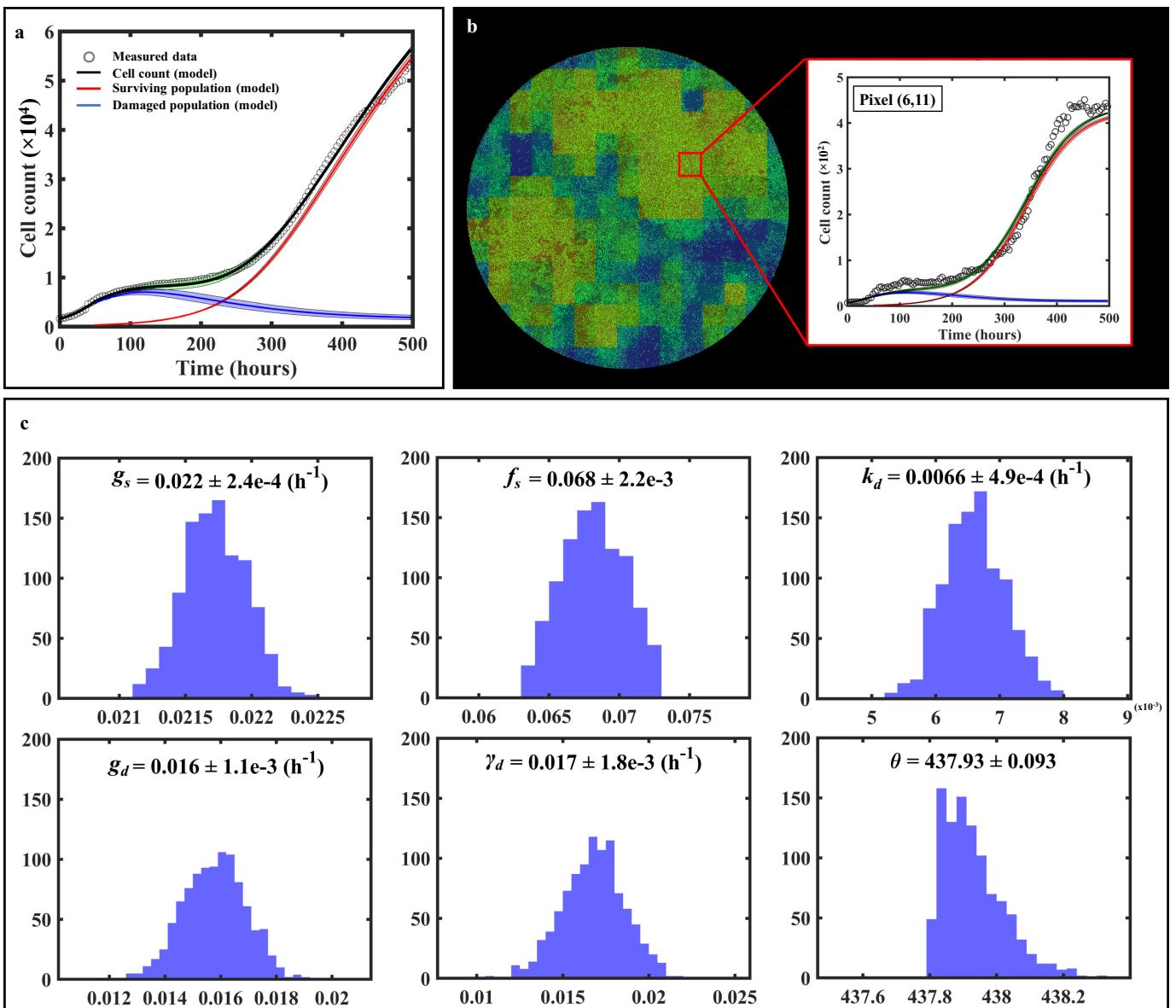


Figure 6. Spatial calibration of the two-phenotype model on a replicate treated with 35 nM of doxorubicin. (a) Bulk model calibration providing the model's estimate of the surviving, damaged, and total population. The calibration precision is shown on all three populations (b) Cell map at final time point along with model calibration on an individual pixel at final time point ($t = 497$ h). (c) Distributions of each six post-treatment model parameters for pixel (6,11). The largest spreads are found on k_d and γ_d , indicative of a greater uncertainty on the estimate of the damaged population as reflected by its larger shaded area on panel a.

in Figure 7. Again, predictions of total, surviving, and irreversibly damaged cell counts are shown along with the measured data. As outlined in the Methods section, the first prediction (week 0 to 1) is performed using the population parameter set from the training set, and predictions from week 1 to week 2 are then performed by calibrating the model from week 0 to week 1 and running the model forward in time from week 1 to 2. Then, predictions from week 2 to week 3 are performed by calibrating the model from week 0 to week 2, and running the model forward from week 2 to 3. This process is repeated until measurements stop. The gray dashed lines indicate the weekly timeframes of data assimilation and ensuing prediction; that is, the moments at which the data are measured and the forecast updated. To guide the reader, we make clear that jumps seen in the model estimate of the cell count (such as those observed at

$t = 384$ h for 35 nM, or at $t = 540$ h for 75 nM) reflect model updates following assimilation of the measurements from the preceding week, resulting in a correction of the predicted cell count that had been previously over- (in the 35 nM case) or underestimated (75 nM case). The quality of our model predictions for the whole well and individual pixels resulted in $CCC_{\text{well}} = .97$ [.44, $>.99$] and $CCC_{\text{pixel}} = .69$ [.35, .79], respectively. Figure 8 shows a prediction performed on an individual replicate treated with 100 nM of doxorubicin. Panel 8A presents predictions of the total, surviving, and damaged populations updated every week as indicated by the vertical dashed lines ($CCC_{\text{well}} = .97$). Panel 8B further provides predictions of spatial dynamics and cell distribution. The top and bottom rows present the predicted cell maps and the corresponding measured data, respectively, at the time point preceding the

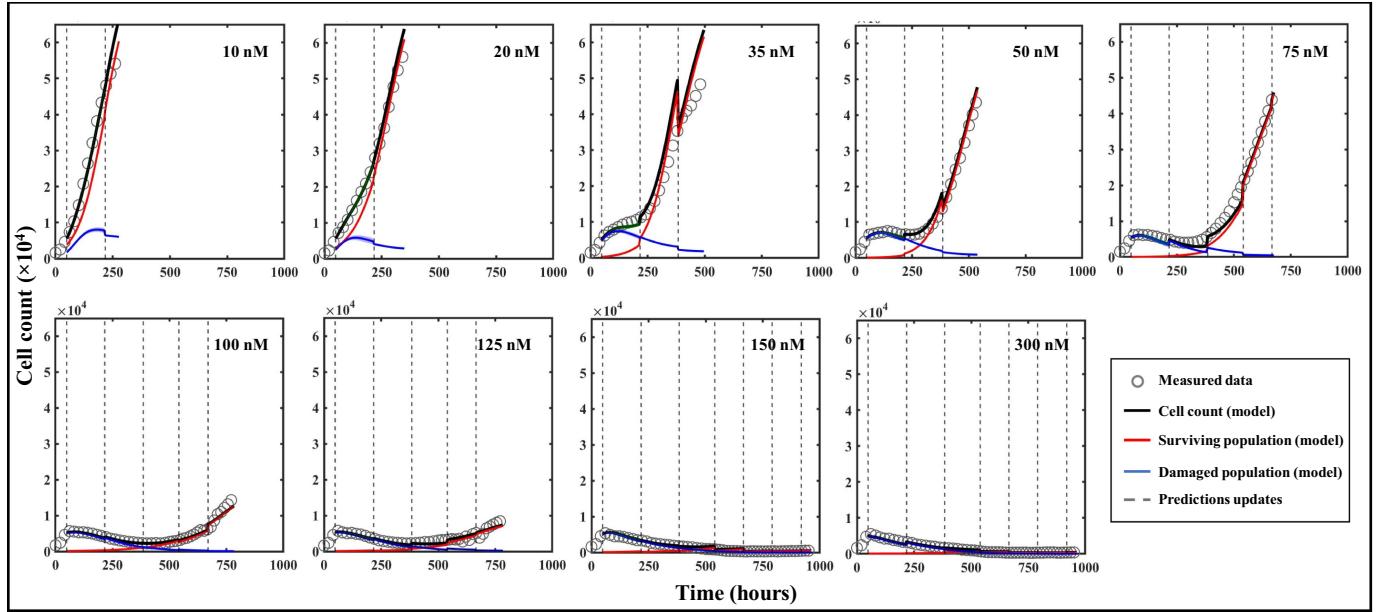


Figure 7. Model predictions of bulk cell count on representative replicates of MCF7 subpopulations treated with varying doses of doxorubicin after 48 hours of growth. Measured bulk cell count (gray circles) is shown along with the model's predictions of the surviving (red), damaged (blue) and total (black) populations as well as their respective precisions. The predictions are performed on weekly segments indicated by the gray dashed lines. Our prediction pipeline is able to successfully forecast the total cell count of treated MCF7 colonies on weekly segments. For clearer visualization, only one out of four data points are shown.

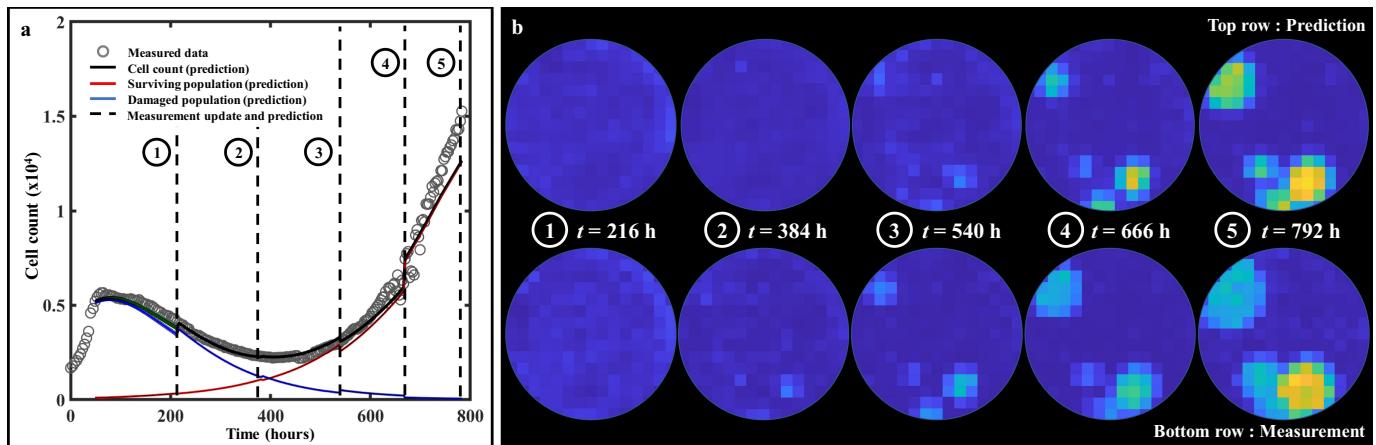


Figure 8. Spatiotemporal prediction of cell dynamics on a replicate treated with 100 nM of doxorubicin. (a) Prediction of bulk cell count. The prediction is performed on weekly segments and updated by measurements indicated by the vertical black dashed lines. (b) Comparisons between our model's predictions of cell distribution and the map measured through microscopy at the time points indicated by the vertical dashed lines in panel A. The main patterns of heterogeneous cell distribution following doxorubicin treatment are successfully captured through data assimilation.

weekly updates. After an initial phase of treatment-induced cell death and its associated global downward trend in total cell count for approximately two weeks, clusters of recovering, highly-proliferative cells become identifiable (bottom row, $t = 384$ h). These clusters remain unpredictable at this stage and can only be identified at later time points (top row, $t > 384$ h). More specifically, the prediction from week 1 ($t = 216$ h) to week 2 ($t = 384$ h) is performed by calibrating the model to the data from week 0 ($t = t_{dox} = 48$ h) to week 1 ($t = 216$ h) – when those clusters were not yet apparent (bottom row, $t = 216$ h) – and running the model forward in time. However, when predicting from week 2 to week 3 ($t = 540$ h), and therefore calibrating from week 0 to week 2, these clusters are now part of the calibrated data and the personalized prediction process

enables an accurate forecast for weeks 4 and 5, where the general trend and main patterns of cell distribution and spatial growth are successfully predicted ($CCC_{pixel} = .66$).

Lastly, Figure 9 quantifies how the bootstrapped parameter distributions evolve after each weekly prediction on a replicate treated with 50 nM of doxorubicin. Panel 9A shows the bulk cell count prediction ($CCC_{well} = .99$), while the prediction on an individual pixel is illustrated on panel 9B. The parameter distributions used to predict each of the three weekly segments of this pixel are presented in panel 9C. In particular, we observe that f_s experiences a significant shift in mean and standard deviation after each week. This behavior is a consequence of the surviving population gradually outnumbering the damaged population. Moreover, the changes in the

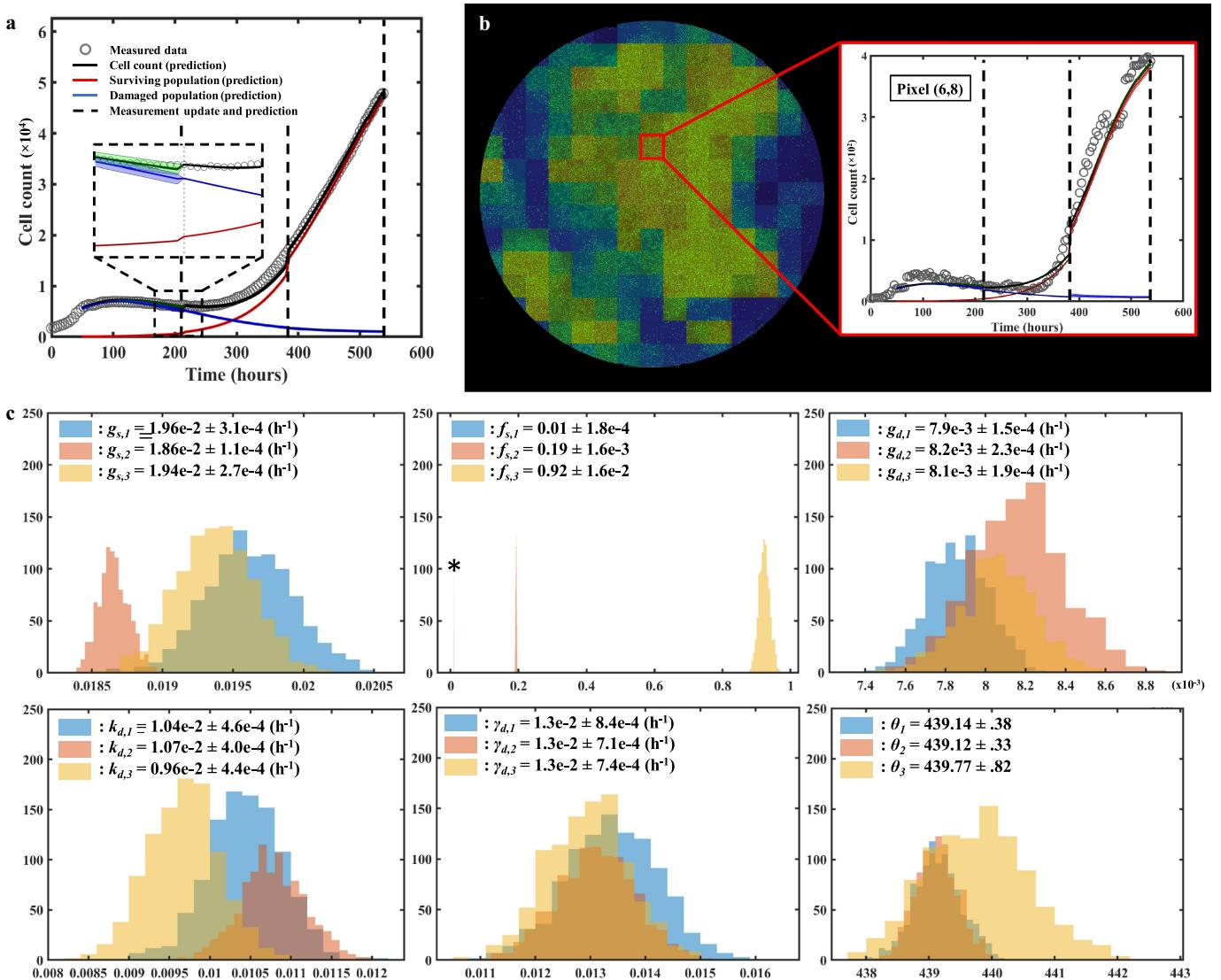


Figure 9. Parameter distributions of consecutive predictions following weekly data assimilation. (a) Prediction of bulk cell count on a replicate treated with 50 nM of doxorubicin. (b) Prediction of cell count on pixel (6,8). (c) Bootstrapped parameter distributions for model predictions on pixel (6,8), showing the shift in the mean value and variance of the post-treatment parameters between each weekly timeframe. The distribution for the first weekly segment of parameter f_s is indicated by a star as it is so narrow and may be difficult to see. The changes in the width of the parameter distributions is representative of changes in the uncertainty in the parameter value, which emanate from the quality of the model fit in the preceding weeks as well as from the wide variety of parameter values in similar pixels in the training set. The changes in the mean value point toward changes in the overall dynamics identified by the model as more data is progressively assimilated on a weekly basis.

bootstrapped distributions of the proliferation rate of the surviving population (g_s) shown in panel 9C illustrate the impact of assimilating individualized measurements of tumor cell dynamics in the model to render more accurate parameter estimations and forecasts.

When predicting the first week of data using the parameter values from the training set, the model is able to capture the usual initial growth followed by an incipient downward trend in the total tumor cell count observed after the delivery of doxorubicin in the replicates (e.g., see Figure 7). When calibrating on the first week of data (and therefore predicting the second week), the model is only exposed to the aforementioned downward trend, which translates into a narrow distribution for $g_{s,2}$ around lower values than $g_{s,1}$ (i.e., the assimilation of the observed cell dynamics over the first week post-treatment suggests lower surviving cell proliferation rates). However, this update to g_s provides a poor estimate of

the dynamics of the surviving and total population over the second week because the model has not yet been informed by the sharp increase in total tumor cell counts after $t = 300$ h; Then, once the data collected during week 2 is fed to the model, the calibration provides a distribution of $g_{s,3}$ that is similar to $g_{s,1}$, exhibiting larger values for this parameter that match the rising trend in total tumor cell counts for this replicate. Additionally, once the tumor cell population enters a growing phase after treatment, the parameters extracted from the training set (P_{training}) to calculate $g_{s,3}$ correspond to pixels in replicates exhibiting post-therapeutic relapse as well. Since these dynamics of relapse can vary widely across the dosage levels in this dataset (see Figure 7 and ref. ¹⁸), the variance of the values of g_s corresponding to P_{training} will be large, and this translates into $g_{s,3}$ having larger variance than $g_{s,2}$. For g_d and k_d , the predictions over the second week are once again performed by exposing the model to the downward

trend of the first week. This leads to an overestimation of $g_{d,2}$ and $k_{d,2}$ over $g_{d,1}$ and $k_{d,1}$ (calibrated on the entire training set) due to the model calibration on this portion of the temporal measurements and the selection of pixels from wells with higher values for those two parameters. After assimilating the relapse of the second week, $g_{d,3}$ and $k_{d,3}$ are readjusted to lower values. Note that γ_d (the death-delay constant) experiences little shift or change in spread week after week. The rationale for this behavior is that this parameter typically influences only the first week of growth, and can thus be precisely and consistently estimated early after treatment.

Discussion

One way to improve the treatment and care of women with breast cancer relies on a rigorous understanding of chemoresistance through the lens of tumor heterogeneity and its patient-specific nature. In this study, we have proposed a novel experimental-mathematical pipeline to model, characterize, and predict the total cell number and spatial distribution of *in vitro* colonies of MCF7 breast cancer cells over an extended time frame (up to 6 weeks). This duration is marked by an elevated amount of spatial and temporal heterogeneity generated by the cells' treatment with doxorubicin. Our work includes an explicit separation between tumor cells surviving treatment and those irreversibly damaged, making tumor heterogeneity the main focus of our longitudinal spatiotemporally-resolved modeling and predicting effort. The addition of a spatial component to the model previously developed by Yang *et al.*¹⁸ represents progress in the ongoing effort to gain quantitative insights on the heterogeneous aspects of tumors which can reduce the efficacy of drugs.^{4,7–9} Utilizing a mechanism-based model offers the advantage of straightforward interpretations due to its explicit consideration of underlying biological mechanisms of (in the present case) tumor cell movement, proliferation, death due to treatment with doxorubicin and relative fractions of damaged and resistant cells. Moreover, employing a mechanism-based model provides valuable guidance for conducting additional experiments and analyses. By establishing clear relations between experimental conditions and model parameters, such studies allow identification of (for example) the effective treatment protocols and experimental designs (see, e.g., refs. ^{18, 41–43}). The model was first calibrated using full time-resolved microscopy datasets yielding parameter sets that were then leveraged to forecast growth dynamics and treatment response through a data assimilation pipeline relying on a succession of short-term predictions and measurements. This approach was able to successfully recapitulate and predict the main spatiotemporal dynamics of the tumor cells following treatment with a wide range of doses of doxorubicin. A data assimilation process, as presented in this work, could be efficiently translated to clinical applications in which the microscopy data is replaced by (for example) longitudinal imaging data.⁵¹

Previous studies have shown the potential of mathematical oncology for predicting tumor response by developing models capturing cell growth, movement, and treatment response in both the pre-clinical and clinical settings.^{18–28} For example,

Yang *et al.* investigated the interactions between glucose availability and cancer cell dynamics by leveraging a model informed by time-resolved microscopy data and consisting of a time-resolved system of ordinary differential equations.²⁰ At the genomic level, Johnson *et al.* performed transcriptomic data analysis on single cells to predict treatment response through a machine learning framework incorporated into a two-phenotype ODE (ordinary differential equations)-based modeling platform.⁵² In the pre-clinical, *in vivo* setting, quantitative magnetic resonance imaging (MRI) data were acquired on murine models of glioma by Hormuth *et al.* to parametrize a reaction-diffusion model that includes angiogenesis.⁵³ We sought to build on these efforts by refining the prediction approach through data assimilation, a methodology that is commonly used in weather forecasting (see ref.⁵⁴ and implemented to predict *in vitro* glioma cell growth by Liu *et al.*⁴⁰ Zahid *et al.* also designed prediction pipelines including a dynamic tumor carrying capacity undergoing stepwise reductions depending on the radiation dose given,⁵⁵ or featuring adaptative dosing allowing treatment escalation and de-escalation.⁵⁶ First initialized with population data, the models were updated with newly-collected patient-specific information to perform individual predictions of outcome. In addition, other studies have aimed at capturing heterogeneities within tumors through various approaches. Slavkova *et al.* employed a similar experimental platform as Hormuth *et al.* (quantitative MRI data of murine models) to describe the dynamics of "habitats" within gliomas subjected to radiotherapy.³¹ In our work, we have extended these studies by using a set of partial differential equations inspired by the model in ref. 18 along with a data assimilation-prediction approach adopted from ref. 40 to characterize the spatial and temporal development of MCF7 colonies over weeks.

As a novel approach to characterize spatiotemporal cell dynamics over extended timeframes, our study contains a number of limitations deserving of greater scrutiny. Firstly, we adopted a fixed spatial resolution for all the pixel cell density maps. Segmentation differences between images along with perturbations of the microscope's field of view caused by vibrations and media changes generate noise on the signal (i.e., the cell count) of a given pixel. As such, as the grid size is made finer, the signal-to-noise ratio on a pixel basis deteriorates, leading to fitting inaccuracies and greater parameter uncertainty. While grid sizes between 3×3 up to 22×22 were investigated, the grid size of 15×15 used heuristically in this work allowed clear identification of the spatial patterns while limiting signal-to-noise deterioration. Moreover, given the time frame of the experiments (from 277 hours for the 10 nM dose to 963 hours for 300 nM), we decided to update the predictions on a weekly basis (168 hours). This choice was motivated by a compromise between quality of prediction and update frequency. Another approach could include, for example, updating the prediction once a specific cell count uncertainty threshold has been crossed. Lastly, while the weighting scheme for ω was chosen based on prior work from our team, the question of the optimal weighting scheme remains open. While other forms of weighting were investigated for this study (including a linear decrease or modifying the weights to give greater influence to $P_{measure}$), Equation (11)

was selected as it generally outperformed all other forms tried. Nevertheless, a more thorough investigation into pixel cell density generation, temporal model updating, and the weighting scheme for the training and measurement-driven parameter estimates is warranted to ensure consistency and set definitive standards on *in vitro* implementations of data assimilation techniques.

Given the number of potential free parameters in our model (i.e., Equations (1)–(7)), we had to determine if each parameter should be calibrated on either a local (i.e., one value per pixel) or a global (i.e., one value shared by all pixels across the map) fashion to reduce complexity and overfitting. To identify which parameters should be calibrated locally, a sensitivity analysis was performed and indicated that the most influential parameters on cell number were the proliferation rate of the surviving cells, g_s , and the carrying capacity, θ . However, given the well-defined geometry of the experimental assay, we elected to interpret θ as a physical carrying capacity (i.e., the maximum number of cells that can physically fit within a fixed region of space) to be calibrated within a narrow range to further guard against potential parameter unidentifiability issues. To confirm this modeling choice other scenarios were investigated, including greater boundaries for θ (Figure S2 of Supplementary Materials) as well as sequentially allowing each post-treatment parameter to be calibrated locally, while calibrating the remaining five globally (Figure S3 in the Supplementary Materials). The results of these analyses indicated that the scenarios allowing g_s to be calibrated locally frequently outperformed the others, and that restricting θ to narrow range did not impact the calibration and prediction performance of our model. We also investigated scenarios in which multiple parameters were locally calibrated (Figure S2 of Supplementary Materials). These attempts unsurprisingly yielded better model calibrations, but the added complexity and parameter unidentifiability led to less accurate predictions. As obtaining the most accurate predictions was central to this study, such scenarios were discarded. We also note that it is indeed expected for the parameters controlling the dynamics of the surviving population to have a greater impact on the output than the parameters controlling the damaged population. As the damaged population is destined to die (no matter the values given to g_d , k_d and γ_d within their boundaries), the vast majority of randomly generated scenarios in the sensitivity analysis will include a surviving population vastly outnumbering the damaged population, thus giving greater weight to the proliferation rate g_s . While it is certainly possible to fix some of the parameters that have less impact on the model output, this would, however, impact the model's flexibility and decrease the calibration accuracy.

As the model characterized by Equations (1)–(7) is built on biological phenomena, the calibration and prediction results can also be readily interpreted from a biological perspective. For example, as f_s is calibrated globally, any subset of cells at any location of the well will be divided into the same fractions of surviving and damaged cells at the time of treatment. Thus, we did not investigate specific patterns of cell distribution for the unaffected and damaged populations separately. The heterogeneity in cell count and spatial dynamics is then captured by some of the surviving cells

exhibiting much higher proliferation rate than other surviving cells depending on their location (i.e., given that $g_s = g_s(x)$). This approach contrasts with a calibration scheme involving, for instance, a globally-calibrated g_s , but a locally-calibrated surviving fraction f_s . In this scenario, the compartmentalization of surviving and damaged subpopulations at treatment onset would vary over the whole well, and hence lead to spatial heterogeneities, but post-treatment dynamics would be controlled by globally-defined parameters. As previously mentioned, the various modeling choices considered in this work were made to maximize our prediction accuracy while minimizing model complexity, and further studies are warranted to thoroughly investigate the underlying biological causes of the observed heterogeneity. We chose to calibrate diffusion in the pre-treatment phase, and then set that value to both phenotypes in the post-treatment phase. While this choice was motivated by seeking to avoid overfitting and parameter identifiability issues, we also did not have access to data that would allow us to pursue a more comprehensive parameterization of this phenomena. Additionally, our model could accommodate temporally varying parameters. As Panel 9C displays, changes in cell population dynamics between consecutive model updates can lead to difficulties predicting the different trends of treatment response observed within a single time course with temporally-fixed parameters. The appearance of highly proliferative and localized clusters of surviving cells after days or weeks of decline in the total tumor cell count points to clear changes in cellular machinery that ought to be reflected in the modeling approach. Thus, temporally varying growth rates represent a promising direction of investigation to ameliorate the model's flexibility and better characterize such phenomena, while dynamic drug-induced death and delay rates (k_d and γ_d) could characterize the various phases of drug intake and action.¹⁸

As both the initial cell seeding and the doxorubicin treatment are applied in a homogeneous fashion (i.e., evenly distributed across the entire well), one may hypothesize that the *in vitro* tumor cell populations would develop in a spatially-homogeneous manner. While this is observed at low (where the effects of the treatment are negligible) and high doses (where no cell recovery was observed), it is decidedly not the case within the intermediate range (i.e., 50 to 150 nM). The underlying biological causes of this heterogeneous development are not interrogated in this study and tracking more specific metrics (e.g., time-resolved live and dead cell counts, or “cell barcoding” technology for capturing phenotypic dynamics)^{57,58} could lead to greater insights on this topic. As predictions are performed and updated on a weekly basis using only cell counts and centroids positions, such phenomena may be missed by the current realization of the prediction scheme. Understanding the causes behind the rapid changes in cell number driving spatial heterogeneity (e.g., in panel B of Figure 9) would allow to capture these events more accurately. Predictions could then be further improved by anticipating changes by leveraging temporally-varying model parameters as mentioned above.

While the data employed for this study possess spatial and temporal resolution not available in the clinical setting, it

allows for the development of techniques that are readily translated provided the requisite data are available. We propose that our prediction pipeline can also be implemented on clinical data available from (for example) MRI data obtained before and during treatment. Indeed, such data are already used for predicting treatment outcomes^{23–29}; thus, the notion of employing a data assimilation technique accounting for spatial heterogeneity in response to systemic or local therapy is a natural next step for optimizing interventions on a patient specific basis. In particular, some radiation therapy studies do provide the opportunity for more frequent imaging during the course of therapy, thereby allowing for the implementation of a data assimilation pipeline similar to that presented in the present study.⁵⁹ Lastly, we note that trying to predict an accurate estimate of the cell count is not necessarily required for treatment optimization in the clinical setting. For example, techniques such as functional diffusion mapping⁶⁰ coupled with our data assimilation approach could be used to predict areas of increased proliferation, quiescence, or death to guide the delivery of radiation treatment plans.

Conclusion

We have presented a computational-experimental platform that efficiently recapitulates and forecasts highly heterogeneous spatiotemporal dynamics of tumor cell populations before and after chemotherapy with high accuracy. In particular, our methodology can forecast the development of major spatial patterns of MCF7 cell populations after their formation following treatment with doxorubicin *in vitro*. These results represent an important step in the effort of quantifying the relationship between chemoresistance and tumor heterogeneity as well as their effect on treatment efficacy.

Acknowledgment

The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing high-performance computing resources that have contributed to the research results reported within this paper, as well as Grant Howard for providing MCF7 breast cancer microscopy data. GL acknowledges funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 838786.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Cancer Institute under Grants R01CA240589, U01CA174706, R01CA186193, U24CA226110, U01CA253540 and by CPRIT via RR160005 and RP220225. TEY is a CPRIT Scholar of Cancer Research.

Notes on contributors

Hugo Miniere received his M.Eng from Arts et Métiers ParisTech in 2018 and his M.S. from Texas Tech University in 2019. He is currently a Ph.D. student in the Department of Biomedical Engineering at the University of Texas at Austin.

Ernesto Lima received his B.S. and M.S. in 2008 and 2010 from the Universidade Estadual Paulista and his Ph.D. in Computational Modeling from the National Laboratory for Scientific Computing (LNCC) in Brazil in 2014. He is currently working on the development of numerical methods, novel tumor growth models, and model selection and calibration as a Research Associate at the Center for Computational Oncology.

Guillermo Lorenzo received his M.S. and Ph.D. in civil engineering from the University of A Coruña in 2013 and 2018. He is currently a postdoctoral researcher at the Health Research Institute of Santiago de Compostela.

David Hormuth II received his B.S. from the Rose-Hulman Institute of Technology in 2010, and his M.S. and Ph.D. at Vanderbilt University in 2012 and 2016. He is now a Research Scientist at the Oden Institute for Computational Engineering and Sciences and his research interests are in the integration of predictive mathematical models with patient-specific data (imaging or otherwise) to help guide cancer care.

Sophia Ty received her B.S. in biomedical Engineering at the University of Texas at Austin in 2023. She is now a postbaccalaureate research fellow and the National Cancer Institute.

Amy Brock received her B.S. at the Massachusetts Institute of Technology in 1998 and her Ph.D. from Harvard University in 2004. She is now an Associate Professor at the University of Texas at Austin where her research integrates approaches from dynamic systems, computational sciences, microfabrication and molecular cell biology to tackle fundamental questions of tumor cell-state alterations and response to treatment.

Thomas Yankeelov received his M.S. in Applied Mathematics and Physics in 1998 and 2000 at Indiana University and his Ph.D. at SUNY Stony Brook in 2003. He is now the W.A. "Tex" Moncrief Chair of Computational Oncology and Professor of Biomedical Engineering, Diagnostic Medicine, and Oncology at the University of Texas at Austin. His research aims at developing tumor forecasting methods by integrating advanced imaging technologies with predictive, multi-scale models of tumor growth to optimize therapy.

ORCID

Ernesto A. B. F. Lima  <http://orcid.org/0000-0002-9219-8784>
 Guillermo Lorenzo  <http://orcid.org/0000-0001-9090-190X>
 David A. Hormuth II  <http://orcid.org/0000-0002-9643-1694>
 Amy Brock  <http://orcid.org/0000-0001-8255-9024>
 Thomas E. Yankeelov  <http://orcid.org/0000-0001-6201-3913>

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <https://doi.org/10.5281/zenodo.6973776>.

References

1. Holohan C, Van Schaeybroeck S, Longley D, Johnston P. Cancer drug resistance: an evolving paradigm. *Nat Rev Cancer*. 2013;13(10):714–726. doi:[10.1038/nrc3599](https://doi.org/10.1038/nrc3599).
2. Zheng HC. The molecular mechanisms of chemoresistance in cancers. *Oncotarget*. 2017;8(35):59950–59964. PMID: 28938696; PMCID: PMC5601792. doi:[10.18632/oncotarget.19048](https://doi.org/10.18632/oncotarget.19048).
3. Dagogo-Jack I, Shaw A. Tumour heterogeneity and resistance to cancer therapies. *Nat Rev Clin Oncol*. 2018;15(2):81–94. London. doi:[10.1038/nrclinonc.2017.166](https://doi.org/10.1038/nrclinonc.2017.166).
4. Easwaran H, Tsai H, Baylin S. Cancer epigenetics: Tumor heterogeneity, plasticity of stem-like states, and drug resistance. *Mol Cell*. 2014;54(5):716–727. doi:[10.1016/j.molcel.2014.05.015](https://doi.org/10.1016/j.molcel.2014.05.015).
5. Dong H, Wang W, Mo S, Chen R, Zou K, Han J, Zhang F, Hu J. RETRACTED ARTICLE: SP1-induced lncRNA AGAP2-AS1 expression promotes chemoresistance of breast cancer by

- epigenetic regulation of MyD88. *J Exp Clin Cancer Res.* **2018**;37(1):202–215. doi:[10.1186/s13046-018-0875-3](https://doi.org/10.1186/s13046-018-0875-3).
6. Ji X, Lu Y, Tian H, Meng X, Wei M, Cho W. Chemoresistance mechanisms of breast cancer and their countermeasures. *Biomed Pharmacother.* **2019**. 114:108800. doi:[10.1016/j.bioph.2019.108800](https://doi.org/10.1016/j.bioph.2019.108800).
 7. Polyak K. Heterogeneity in breast cancer. *J Clin Invest.* **2011**;121(10):3786–3788. doi:[10.1172/JCI60534](https://doi.org/10.1172/JCI60534).
 8. Alizadeh A, Aranda V, Bardelli A, Blanpain C, Bock C, Borowski C, Caldas C, Califano A, Doherty M, Elsner M, et al. Toward understanding and exploiting tumor heterogeneity. *Nat Med.* **2015**;21(8):846–853. doi:[10.1038/nm.3915](https://doi.org/10.1038/nm.3915).
 9. Sun XX, Yu Q. Intra-tumor heterogeneity of cancer cells and its implications for cancer treatment. *Acta Pharmacol Sin.* **2015**;36(10):1219–1227. doi:[10.1038/aps.2015.92](https://doi.org/10.1038/aps.2015.92).
 10. Gatenby RA, Grove O, Gillies RJ. Quantitative imaging in cancer evolution and ecology. *Radiology.* **2013**;269(1):8–14. doi:[10.1148/radiol.13122697](https://doi.org/10.1148/radiol.13122697).
 11. Syed A, Whisenant J, Barnes S, Sorace A, Yankeelov T. Multiparametric analysis of longitudinal quantitative MRI data to identify distinct tumor habitats in preclinical models of breast cancer. *Cancers (Basel).* **2020**;12(6):1–20. doi:[10.3390/cancers12061682](https://doi.org/10.3390/cancers12061682).
 12. Kazerouni A, Hormuth D 2nd, Davis T, Bloom M, Mounho S, Rahman G, Virostko J, Yankeelov T, Sorace A. Quantifying tumor heterogeneity via MRI habitats to characterize microenvironmental alterations in HER2+ breast cancer. *Cancers (Basel).* **2022**;14, 1837.
 13. Hanahan D. Hallmarks of cancer: new dimensions. *Cancer Discov.* **2022**;12(1):31–46. doi:[10.1158/2159-8290.CD-21-1059](https://doi.org/10.1158/2159-8290.CD-21-1059).
 14. Jarrett A, Faghihi D, Hormuth D 2nd, Lima E, Virostko J, Biros G, Patt D, Yankeelov T. Optimal control theory for personalized therapeutic regimens in oncology: background, history, challenges, and opportunities. *J Clin Med.* **2020**;9(5):1314. PMID: 32370195; PMCID: PMC7290915. doi:[10.3390/jcm9051314](https://doi.org/10.3390/jcm9051314).
 15. Yankeelov T, Atuegwu N, Hormuth D 2nd, Weis J, Barnes S, Miga M, Rericha E, Quaranta V. Clinically relevant modeling of tumor growth and treatment response. *Sci Transl Med.* **2013**;5(187):187ps9. doi:[10.1126/scitranslmed.3005686](https://doi.org/10.1126/scitranslmed.3005686).
 16. Anderson A, Quaranta V. Integrative mathematical oncology. *Nat Rev Cancer.* **2008**;8(3):227–234. doi:[10.1038/nrc2329](https://doi.org/10.1038/nrc2329).
 17. Ghaffari Laleh N, Loeffler C, Grajek J, Stankova K, Pearson A, Muti H, Trautwein C, Enderling H, Poleszczuk J, Kather J, et al. Classical mathematical models for prediction of response to chemotherapy and immunotherapy. *PLoS Comput Biol.* **2022**;18(2):e1009822. doi:[10.1371/journal.pcbi.1009822](https://doi.org/10.1371/journal.pcbi.1009822).
 18. Yang E, Howard G, Brock A, Yankeelov T, Lorenzo G. Mathematical characterization of population dynamics in breast cancer cells treated with doxorubicin. *Front Mol Biosci.* **2022**. 9:972146. doi:[10.3389/fmolb.2022.972146](https://doi.org/10.3389/fmolb.2022.972146).
 19. Liu J, Hormuth D 2nd, Davis T, Yang J, McKenna M, Jarrett A, Enderling H, Brock A, Yankeelov T. A time-resolved experimental-mathematical model for predicting the response of glioma cells to single-dose radiation therapy. *Integr Biol.* **2021**;202113(7):167–183. (Camb) PMID: 34060613; PMCID: PMC8271006. doi:[10.1093/intbio/zyab010](https://doi.org/10.1093/intbio/zyab010).
 20. Yang J, Virostko J, Hormuth D 2nd, Liu J, Brock A, Kowalski J, Yankeelov T, Carlier A. An experimental-mathematical approach to predict tumor cell growth as a function of glucose availability in breast cancer cell lines. *PloS One.* **2021**;16(7):e0240765. PMID: 34255770; PMCID: PMC8277046. doi:[10.1371/journal.pone.0240765](https://doi.org/10.1371/journal.pone.0240765).
 21. Pozzi G, Grammatica B, Chaabane L, Catucci M, Mondino A, Zunino P, Ciarletta P. T cell therapy against cancer: a predictive diffuse-interface mathematical model informed by pre-clinical studies. *J Theor Biol.* **2022**;547:111172. ISSN 0022-5193. doi:[10.1016/j.jtbi.2022.111172](https://doi.org/10.1016/j.jtbi.2022.111172).
 22. Luo M, Nikolopoulou E, Gevertz J. 2022. From fitting the average to fitting the individual: a cautionary tale for mathematical modelers. *Front Oncol.* 12. doi: [10.3389/fonc.2022.793908](https://doi.org/10.3389/fonc.2022.793908).
 23. Hormuth D 2nd, Weis J, Barnes S, Miga M, Rericha E, Quaranta V, Yankeelov T. A mechanically coupled reaction–diffusion model that incorporates intra-tumoural heterogeneity to predict *in vivo* glioma growth. *J R Soc Interface.* **2017**;14(128):142016101020161010. doi:[10.1098/rsif.2016.1010](https://doi.org/10.1098/rsif.2016.1010).
 24. Wu C, Jarrett A, Zhou Z, Elshafeey N, Adrada B, Candelaria R, Mohamed R, Boge M, Huo L, White J, et al. MRI-Based digital models forecast patient-specific treatment responses to neoadjuvant chemotherapy in triple-negative breast cancer. *Cancer Res.* **2022**; 82(18):3394–3404. PMID: 35914239; PMCID: PMC9481712. doi:[10.1158/0008-5472.CAN-22-1329](https://doi.org/10.1158/0008-5472.CAN-22-1329).
 25. Hormuth D 2nd, Al Feghali K, Elliott A, Yankeelov T, Chung C. Image-based personalization of computational models for predicting response of high-grade glioma to chemoradiation. *Sci Rep.* **2021**;11(1):8520. PMID: 33875739; PMCID: PMC8055874. doi:[10.1038/s41598-021-87887-4](https://doi.org/10.1038/s41598-021-87887-4).
 26. Rockne R, Trister A, Jacobs J, Hawkins-Daarud A, Neal M, Hendrickson K, Mrugala M, Rockhill J, Kinahan P, Krohn K, et al. A patient-specific computational model of hypoxia-modulated radiation resistance in glioblastoma using ⁽¹⁸⁾F-FMISO-PET. *J R Soc Interface.* **2015**;12(103):20141174. doi:[10.1098/rsif.2014.1174](https://doi.org/10.1098/rsif.2014.1174).
 27. Hogaë C, Davatzikos C, Biros G. An image-driven parameter estimation problem for a reaction-diffusion glioma growth model with mass effects. *J Math Biol.* **2008**;56(6):793–825. doi:[10.1007/s00285-007-0139-x](https://doi.org/10.1007/s00285-007-0139-x).
 28. Clatz O, Sermesant M, Bondiau P-Y, Delingette H, Warfield S, Malandain G, Ayache N. Realistic simulation of the 3-D growth of brain tumors in MR images coupling diffusion with biomechanical deformation. *IEEE Trans Med Imaging.* **2005**;1(10):1334–1346. doi:[10.1109/TMI.2005.857217](https://doi.org/10.1109/TMI.2005.857217).
 29. O'Connor J, Rose C, Waterton J, Carano R, Parker G, Jackson A. Imaging intratumor heterogeneity: Role in therapy response, resistance, and clinical outcome. *Clin Cancer Res.* **2015**;21:249–257. doi:[10.1158/1078-0432.CCR-14-0990](https://doi.org/10.1158/1078-0432.CCR-14-0990).
 30. Hu L, Hawkins-Daarud A, Wang L, Li J, Swanson K. Imaging of intratumoral heterogeneity in high-grade glioma. *Cancer Lett.* **2020**;477:97–106. PMID: 32112907; PMCID: PMC7108976. doi:[10.1016/j.canlet.2020.02.025](https://doi.org/10.1016/j.canlet.2020.02.025).
 31. Slavkova K, Patel S, Cacini Z, Kazerouni A, Gardner A, Yankeelov T, Hormuth DA, Hormuth 2nd D. Mathematical modelling of the dynamics of image-informed tumor habitats in a murine model of glioma. *Sci Rep.* **2023**;13(1):2916. doi:[10.1038/s41598-023-30010-6](https://doi.org/10.1038/s41598-023-30010-6).
 32. Zhang J, Cunningham J, Brown J, Gatenby R. Integrating evolutionary dynamics into treatment of metastatic castrate-resistant prostate cancer. *Nat Commun.* **2017**; 8, 1816. [10.1038/s41467-017-01968-5](https://doi.org/10.1038/s41467-017-01968-5).
 33. Brady-Nicholls R, Enderling H. Range-bounded adaptive therapy in metastatic prostate cancer. *Cancers.* **2022**;14(21):5319. doi:[10.3390/cancers14215319](https://doi.org/10.3390/cancers14215319).
 34. Brady-Nicholls R, Zhang J, Zhang T, Wanh A, Butler R, Gatenby R, Enderling H. Predicting patient-specific response to adaptive therapy in metastatic castration-resistant prostate cancer using prostate-specific antigen dynamics. *Neoplasia.* **2021**;23(9):851–858. doi:[10.1016/j.neo.2021.06.013](https://doi.org/10.1016/j.neo.2021.06.013).
 35. Kazerouni A, Gadde M, Gardner A, Hormuth D 2nd, Jarrett A, Johnsn K, Lina E, Lorenzo G, Philips C, Brock A, et al. Integrating quantitative assays with biologically based mathematical modeling for predictive oncology. *iScience.* **2020**;23(12):101807. doi:[10.1016/j.isci.2020.101807](https://doi.org/10.1016/j.isci.2020.101807).
 36. McKenna M, Weis J, Quaranta V, Yankeelov T. Variable cell line pharmacokinetics contribute to non-linear treatment response in heterogeneous cell populations. *Ann Biomed Eng.* **2018**;46(6):899–911. doi:[10.1007/s10439-018-2001-2](https://doi.org/10.1007/s10439-018-2001-2).
 37. Tyson D, Garbett S, Frick P, Quaranta V. Fractional proliferation: a method to deconvolve cell population dynamics from single-cell data. *Nat Methods.* **2012**;9(9):923. doi:[10.1038/nmeth.2138](https://doi.org/10.1038/nmeth.2138).
 38. Lima E, Ghousifam N, Ozkan A, Oden J, Shahmoradi A, Rylander M, Wohlmuth B, Yankeelov T. Calibration of multi-parameter models of avascular tumor growth using time resolved microscopy data. *Sci Rep.* **2018**;8(1):14558. doi:[10.1038/s41598-018-32347-9](https://doi.org/10.1038/s41598-018-32347-9).

39. Howard G, Jost T, Yankeelov T, Brock A, Finley S. Quantification of long-term doxorubicin response dynamics in breast cancer cell lines to direct treatment schedules. *PLoS Comput Biol*. **2022**;18(3):e1009104. doi:[10.1371/journal.pcbi.1009104](https://doi.org/10.1371/journal.pcbi.1009104).
40. Liu J, Hormuth D 2nd, Yang J, Yankeelov T. A data assimilation framework to predict the response of glioma cells to radiation. *Math Biosci Eng*. **2023**;20(1):318–336. doi:[10.3934/mbe.2023015](https://doi.org/10.3934/mbe.2023015).
41. Wu C, Hormuth D 2nd, Lorenzo G, Jarrett A, Pineda F, Howard F, Karczmar G, Yankeelov T. Towards patient-specific optimization of neoadjuvant treatment protocols for breast cancer based on image-guided fluid dynamics. *IEEE Trans Biomed Eng*. **2022**;69(11):3334–3344. doi:[10.1109/TBME.2022.3168402](https://doi.org/10.1109/TBME.2022.3168402).
42. Jarrett A, Hormuth D 2nd, Wu C, Kazerouni A, Ekrut D, Virostko J, Sorace A, DiCarlo J, Kowalski J, Patt D, et al. Evaluating patient-specific neoadjuvant regimens for breast cancer via a mathematical model constrained by quantitative magnetic resonance imaging data. *Neoplasia*. **2020**;12(12):820–830. doi:[10.1016/j.neo.2020.10.011](https://doi.org/10.1016/j.neo.2020.10.011).
43. Lima E, Wyde R, Sorace A, Yankeelov T. Optimizing combination therapy in a murine model of HER2+ breast cancer. *Comput Methods Appl Mech Eng*. **2022**. 402:115484. doi:[10.1016/j.cma.2022.115484](https://doi.org/10.1016/j.cma.2022.115484).
44. Kowarz E, Löscher D, Marschalek R. Optimized sleeping beauty transposons rapidly generate stable transgenic cell lines. *Biotechnol J*. **2015**;10(4):647–653. doi:[10.1002/biot.201400821](https://doi.org/10.1002/biot.201400821).
45. Mátés L, Chuah M, Belay E, Jerchow B, Manoj N, Acosta-Sánchez A, Grzela D, Schmitt A, Becker K, Matrai J, et al. Molecular evolution of a novel hyperactive sleeping beauty transposase enables robust stable gene transfer in vertebrates. *Nat Genet*. **2009**;41(6):753–761. doi:[10.1038/ng.343](https://doi.org/10.1038/ng.343).
46. Jarrett A, Lima E, Hormuth D 2nd, McKenna M, Feng X, Ekrut D, Resende A, Brock A, Yankeelov T. Mathematical models of tumor cell proliferation: a review of the literature. *Expert Rev Anticancer Ther*. **2018**;18(12):1271–1286. doi:[10.1080/14737140.2018.1527689](https://doi.org/10.1080/14737140.2018.1527689).
47. Gewirtz D. A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin. *Biochem Pharmacol*. **1999**;57(7):727–741. doi:[10.1016/s0006-2952\(98\)00307-4](https://doi.org/10.1016/s0006-2952(98)00307-4).
48. Carvalho C, Santos R, Cardoso S, Correia S, Oliveira P, Santos M, Moreira PD. Doxorubicin: the good, the bad and the ugly effect. *Curr Med Chem*. **2009**;16(25):3267–3285. doi:[10.2174/092986709788803312](https://doi.org/10.2174/092986709788803312).
49. McKenna M, Weis J, Barnes S, Tyson D, Miga M, Quaranta V, Yankeelov T. A predictive mathematical modeling approach for the study of doxorubicin treatment in triple negative breast cancer. *Sci Rep*. **2017**;7(1):5725. doi:[10.1038/s41598-017-05902-z](https://doi.org/10.1038/s41598-017-05902-z).
50. Lin L. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. **1989**;45(1):255–268. JSTOR. doi:[10.2307/2532051](https://doi.org/10.2307/2532051).
51. Kostelich E, Kuang Y, McDaniel J, Moore N, Martirosyan N, Preul M. Accurate state estimation from uncertain data and models: an application of data assimilation to mathematical models of human brain tumors. *Biol Direct*. **2011**;6(1):64. doi:[10.1186/1745-6150-6-64](https://doi.org/10.1186/1745-6150-6-64).
52. Johnson K, Howard G, Morgan D, Brenner E, Gardner A, Durrett R, Mo W, Al-Khafaji A, Sontag E, Jarret A, et al. Integrating transcriptomics and bulk time course data into a mathematical framework to describe and predict therapeutic resistance in cancer. *Phys Biol*. **2020**;18(1):016001. doi:[10.1088/1478-3975/abb09c](https://doi.org/10.1088/1478-3975/abb09c).
53. Hormuth D 2nd, Philips C, Wu C, Lima E, Lorenzo G, Jha P, Jarrett A, Oden J, Yankeelov T. Biologically-based mathematical modeling of tumor vasculature and angiogenesis via time-resolved imaging data. *Cancers*. **2021**;13(12, Art. no. 12):3008. doi:[10.3390/cancers13123008](https://doi.org/10.3390/cancers13123008).
54. Navon I. Data assimilation for numerical weather prediction: a review. In: Park SK Xu L, editors. *Data assimilation for atmospheric, oceanic and hydrologic applications*. Berlin, Heidelberg: Springer; **2009**. pp. 21–65. [10.1007/978-3-540-71056-1_2](https://doi.org/10.1007/978-3-540-71056-1_2)
55. Zahid M, Mohsin N, Mohamed A, Caudell J, Harrison L, Fuller C, Moros E, Enderling H. Forecasting individual patient response to radiation therapy in head and neck cancer with a dynamic carrying capacity model. *Int J Radiat Oncol Bio Phys*. **2021**;111(3):693–704. doi:[10.1016/j.ijrobp.2021.05.132](https://doi.org/10.1016/j.ijrobp.2021.05.132).
56. Zahid M, Mohamed A, Caudell J, Harrison L, Fuller C, Moros E, Enderling H. Dynamics-Adapted Radiotherapy Dose (DARD) for head and neck cancer radiotherapy dose personalization. *J Pers Med*. **2021**;11(11):1124. doi:[10.3390/jpm1111124](https://doi.org/10.3390/jpm1111124).
57. Hu C, He S, Lee Y, HE Y, Kong E, Li H, Anastasio M, Popesco G. Live-dead assay on unlabeled cells using phase imaging with computational specificity. *Nat Commun*. **2022**;13(1):713. doi:[10.1038/s41467-022-28214-x](https://doi.org/10.1038/s41467-022-28214-x).
58. Gutierrez C, Al'khafaji A, Brenner E, Johnson K, Gohil S, Lin Z, Knisbacher B, Durrett R, Li S, Parvin S, et al. Multifunctional barcoding with ClonMapper enables high-resolution study of clonal dynamics during tumor evolution and treatment. *Nat Cancer*. **2021**;2(7):758–772. doi:[10.1038/s43018-021-00222-8](https://doi.org/10.1038/s43018-021-00222-8).
59. Hormuth D 2nd, Farhat M, Bronk J, Langshaw H, Yankeelov T. Effect of chemoradiation on high-grade gliomas can be forecasted by mid-treatment images via image-driven mathematical modeling. *Proc Intl Soc Mag Reson Med*. **2023**;31:0136.
60. Moffat B, Chenevert T, Meyer C, McKeever P, Hall D, Hoff B, Johnson T, Rehemtulla A, Ross B. The functional diffusion map: an imaging biomarker for the early prediction of cancer treatment outcome. *Neoplasia*. **2006**;8(4):259–267. doi:[10.1593/neo.05844](https://doi.org/10.1593/neo.05844).