An experimental-mathematical approach to predict tumor cell growth as a function of glucose availability in breast cancer cell lines

Jianchen Yang[1], Jack Virostko[2,3,5], David A. Hormuth II[4,5], Junyan Liu[1], Amy Brock[1,3,5], Jeanne Kowalski[3,5], Thomas E. Yankeelov[1-6*]

[1]Departments of Biomedical Engineering, The University of Texas at Austin, Austin, Texas, United States of America

[2]Departments of Diagnostic Medicine, The University of Texas at Austin, Austin, Texas, United States of America

[3]Departments of Oncology, The University of Texas at Austin, Austin, Texas, United States of America

[4]Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, Texas, United States of America

[5]Livestrong Cancer Institutes, The University of Texas at Austin, Austin, Texas, United States of America

[6]Departments of Imaging Physics, The University of Texas MD Anderson Cancer Center, Houston, Texas, United States of America

[*] Corresponding author

Email: thomas.yankeelov@utexas.edu (TEY)

Abstract

We present the development and validation of a mathematical model that predicts how glucose dynamics influence metabolism and therefore tumor cell growth. Glucose, the starting material for glycolysis, has a fundamental influence on tumor cell growth. We employed time-resolved microscopy to track the temporal change of the number of live and dead tumor cells under different initial glucose concentrations and seeding densities. We then constructed a family of mathematical models (where cell death was accounted for differently in each member of the family) to describe overall tumor cell growth in response to the initial glucose and confluence conditions. The Akaikie Information Criteria was then employed to identify the most parsimonious model. The selected model was then trained on 75% of the data to calibrate the system and identify trends in model parameters as a function of initial glucose concentration and confluence. The calibrated parameters were applied to the remaining 25% of the data to predict the temporal dynamics given the known initial glucose concentration and confluence, and tested against the corresponding experimental measurements. With the selected model, we achieved an accuracy (defined as the fraction of measured data that fell within the 95% confidence intervals of the predicted growth curves) of 77.2 ± 6.3% and 87.2 ± 5.1% for live BT-474 and MDA-MB-231 cells, respectively.

Key words: metabolism, time-resolved microscopy, computational, BT-474, MDA-MB-231

1. Introduction

The major source of energy for most cancer cells comes from a high rate of glycolysis followed by lactate fermentation in the cytosol, even in the presence of sufficient oxygen—a phenomenon known as the Warburg effect [1,2]. This contrasts with normal cells that exhibit a comparatively low rate of glycolysis followed by oxidative phosphorylation in the mitochondria. Additionally, high concentrations of oxygen can lead to a reduction of glycolytic activity, known as the Pasteur effect [3]. Also, an observation by Sonveaux [4] supports the claim that well-oxygenated tumor cells utilize lactate, leaving glucose available for hypoxic cells. This phenomenon has stimulated numerous efforts to investigate the underlying mechanisms [2,5,6] of altered metabolism and has identified potential targets including glucose transporters [7], lactate transporters [8], and enzymes like hexokinase and pyruvate kinase in the pathway of glycolysis [9,10] for the development of new therapeutics. Efforts have been made to rigorously model the development of tumor subpopulations, nutrient dynamics, and tumor-environment interactions [11–18]. For example, in the model developed by Mendoza-Juez *et al*. [12], tumor cells were divided into two subpopulations, the oxidative cells that undergo aerobic oxidation of glucose and glycolytic cells that undergo glycolysis and produce lactate, corresponding to an oxidative phenotype and a Warburg phenotype. Proliferation and conversion between the two subpopulations was described by a set of ordinary differential equations. This study also considered the nutrient concentrations of glucose and lactate as a result of consumption and production by tumor cells, which in return, can cause conversion between phenotypes. Mendoza-Juez *et al*. [12] further provided preliminary validation of their model by comparing it to metabolic data available from several previously published studies [4,19,20]. However, as no direct calibration of this model to experimental data was performed, it was not possible to capture

specific parameter values that could be used to characterize cell lines [11,16], or make predictions of tumor cell dynamics as a function of glucose availability or utilization. Additionally, the reliance on a large number of unmeasured parameters makes further applications challenging. Therefore, in this work, we aim to simplify this model with a smaller set of parameters that can be estimated or calibrated from experimental data and recast the associated models we developed with these estimates to predict tumor growth given initial conditions.

We designed a set of experiments employing time-resolved microscopy to track the temporal change of the number of live and dead tumor cells *in vitro* given a set of initial confluences (i.e., seeding density) and glucose concentrations. To quantitatively characterize those observations, we developed a family of mathematical models to describe the proliferation and death of tumor cells as a function of glucose availability and consumption. These models, which are based on those of Mendoza-Juez [12], take the form of systems of nonlinear, ordinary differential equations to describe the collective temporal behavior of tumor cells. We aim to identify the most parsimonious model from that family to optimally characterize tumor cell growth as a function of glucose dynamics. After the optimal model is selected, we quantify the proliferation rate, death rate due to glucose depletion, death rate due to the bystander effect, and the consumption rate of glucose in a training set. We then use this calibrated model to predict tumor cell growth given prescribed initial conditions in a validation set.

2. Materials and methods

Throughout the following text, the reader is encouraged to refer to Fig 1 which provides an overview of the experimental and computational modeling components of the study.

Fig 1. A flow chart indicating the data acquisition and analysis steps. Based on the phase-contrast (panels a and b) and fluorescent (panels e and f) images acquired from the time-resolved microscopy studies, we perform cell segmentation of total and dead cells (panels c and g, respectively) and generate time courses of confluence for both live and dead cells (panel d). The data are then used for selecting the most parsimonious mathematical model which estimates model parameters. Finally, the data are divided into subsets for training and validation of the predictive accuracy of the model.

2.1 Cell culture

We applied our experimental-mathematical approach in two breast cancer subtypes to quantitatively characterize cell types known to have distinct phenotypes, molecular profiles, and metabolic activities. Triple negative breast cancer [21] (TNBC) is defined by the absence of the expression of the estrogen, progesterone, and HER2 (human epidermal growth factor receptor 2) receptors, while in HER2+ breast cancer [22], HER2 is overexpressed.

BT-474 (a model of HER2+ breast cancer) and MDA-MB-231 (a model of triple negative breast cancer) cell lines were obtained from the American Type Culture Collection (ATCC, Manassas, VA) and maintained in culture according to ATCC recommendation. Ninety-six well-plates were seeded with either BT-474 or MDA-MB-231 cells at initial confluences ranging from 10% to 80% in Dulbecco's modified eagle medium (DMEM without glucose, sodium pyruvate, HEPES, L-glutamine and phenol red, Thermo Fisher Scientific, Waltham, MA) one day before imaging experiments began. On day zero, media were changed to DMEM with different glucose concentrations (0 mM, 0.1 mM, 0.2 mM, 0.5 mM, 0.8 mM, 1 mM, 2 mM, 5 mM, 8 mM and 10

mM). Each initial condition had four replicates. Cells were cultured in 5% CO_2 and air at 37°C for 4 days.

2.2 Image acquisition

Cells were incubated in the IncuCyte S3 live cell imaging system (Essen BioScience, Ann Arbor, MI). Images were acquired with a 4× objective and stitched together to obtain whole well images (2400 × 2400 pixels) for each well of the 96-well plates *via* the device's whole-well imaging function. IncuCyte Cytotox Red Reagents (Essen BioScience, Ann Arbor, MI), a cyanine nucleic acid dye, was added to the medium on day 0 before the first scan to enable quantification of cell death. Once cells become unhealthy, the plasma membrane begins to lose integrity allowing entry of the IncuCyte Cytotox Reagent and yielding a 100-1000-fold increase in fluorescence upon binding to DNA. Phase-contrast and red fluorescent (excitation wavelength: 585 nm and emission wavelength: 635 nm) images were acquired every 3 hours for 4 days.

2.3 Image segmentation to quantify confluence over time

All cell segmentation was performed in Matlab (The Mathworks, Inc., Natick, MA). The segmentation approaches were developed based on the particular morphological features of the two cells lines. In particular, the BT-474 cells are mass cells with robust cell-cell adhesion that form cell clusters, while the MDA-MB-231 cells are elongated cells [23].

To segment the BT-474 cells within the phase-contrast images at each time point, a predetermined mask corresponding to the size of 96-well-plate from IncuCyte Software (Essen BioScience, Ann Arbor, MI) was first applied to the images so the region of interest (ROI) only included the area within each well and not the surrounding area of the plate in each square image.

The masked image was then converted from the RGB (red, green, blue) format to grayscale and the Matlab function 'colfilt' was used to calculate the standard deviation of signal intensities within each 3-by-3 sliding block of the image to detect the edge of cell clusters. Next, a Gaussian filter was used to smooth the image returned from 'colfilt' to reduce the variance of signal intensities within each cell cluster. The resulting image was then normalized (by dividing the signal intensity of each pixel by the highest signal intensity from each image) between 0 and 1. After normalization, the morphological operator 'imerode' was used to make the clusters shrink in size and enlarge the holes to avoid losing open space within clusters. Next the returned image was converted to a binary image by the Matlab function 'im2bw'. The morphological operator 'imclose' was used to fill holes in the interior of cell clusters. The morphological operator 'imopen' was used to smooth object contours, break thin connections and remove thin protrusions. Finally, 'bwareaopen' was used to remove small objects like cell debris or noise. Please see S1 Fig of the Supporting Information for details and example images from each step.

While BT-474 cells form clusters that have clear boundaries, MDA-MB-231 cells are elongated and do not form clusters. This results in a much higher edge-area ratio in MDA-MB-231 images compared to BT-474. Thus, the segmentation scheme just described was adjusted to handle these differences in cell morphology. In particular, once the ROI was identified, 'histcount' was used to count the number of pixels for each signal intensity (256 possible signal intensity values in grayscale image) within the ROI. The pixels with signal intensities in the top 10% were assigned a 0, while the remaining pixels were assigned a 1 to binarize the image. All other steps were the same as the BT-474 segmentation. Please see S2 Fig of the Supporting Information for example images.

159 The fluorescent images were used to quantify the Cytotox Red signal (which marks the
160 dead cells) for both cell lines. Since MDA-MB-231 cells change from an elongated to a circular
161 morphology when they die, the differences in morphology of the two cell lines observed in phase-
162 contrast images of the living cells vanishes. Thus, we applied the same approach segmenting the
163 phase-contrast images of BT-474 cells to the florescent images of both cell lines.

164 The resulting segmented and binarized phase-contrast and fluorescent images were then
165 analyzed to determine confluence at each time point. Confluence was defined as the percentage of
166 the well covered by cells and was calculated by counting the number of pixels in the segmented
167 images and dividing by the area of the field of view. Thus, our time-resolved microscopy data
168 provided time courses of both living and dead cell number.

169 Tumor cell growth time courses were obtained from 4 experiments for each set of initial
170 conditions, and each point in each time course consisted of a mean ± 95% confidence interval (a
171 one-sample Kolmogorov-Smirnov test confirmed normality). One-way ANOVA was used to
172 compare the average number of live cells for each experiment at the end of day 4 between the
173 groups with different initial conditions.

174

175 2.4 Mathematical models

176 We developed a family of mathematical models to quantitatively and temporally describe
177 the change in tumor cell number as function of glucose levels. To do so, we started with the model
178 developed by Mendoza-Juez *et al*. [12] which describes the tumor as consisting of two
179 subpopulations undergoing either aerobic oxidation of glucose or glycolysis, corresponding to
180 Warburg and oxidative phenotypes, respectively. In our work, we first simplified the model to
181 account for only one metabolic phenotype, and then extended it to account for the accumulation

of dead tumor cells due to glucose depletion and the bystander effect [24,25]. Accordingly, we modeled the change of glucose concentration as a result of consumption by all live tumor cells. Our complete model is described by a coupled set of ordinary differential equations shown below (the reader is encouraged to refer to Table 1 through the following discussion):

$$\frac{dN(t)}{dt} = k_p N(t)\left(1-\frac{N(t)}{\theta}\right) S_p\left(G(t)\right) - k_d N(t) S_d\left(G(t)\right) - k_{bys} N(t)\left(\frac{D(t)}{D(t)+N(t)}\right) \quad [1]$$

$$\frac{dD(t)}{dt} = k_d N(t) S_d\left(G(t)\right) + k_{bys} N(t)\left(\frac{D(t)}{D(t)+N(t)}\right) \quad [2]$$

$$\frac{dG(t)}{dt} = -\nu N(t)\left(\frac{G(t)}{G(t)+G^*}\right), \quad [3]$$

where $N(t)$, $D(t)$, and $G(t)$ describe the live tumor cell number, dead tumor cell number, and glucose concentration, respectively, at time t. The first term on the right-hand side of Eq. [1] describes logistic growth of tumor cells where k_p is the proliferation rate, and θ is the carrying capacity. Here we define the carrying capacity as the limitation on the number of tumor cells that can physically fit within the environment. The logistic term is also modified by the state function, $S_p(G(t))$, that scales the proliferation rate as a function of glucose concentration. The second term on the right-hand side of Eq. [1] describes the death of tumor cells due to glucose depletion at the rate k_d. This term is also modified by the state function, $S_d(G(t))$, that scales the rate of cell death as a function of glucose concentration. We assume that the dead tumor cells are accumulating and releasing factors [24,25] which may be sensed by the remaining live cells and, potentially, induce cell death. This is referred to as the bystander effect [24,25] and it is captured by the third term on the right-hand side of Eq. [1] which induces cell death at the rate k_{bys}. Eq. [2] models the rate of change in number of dead cells, with the first term on the right-hand side describing death due to glucose depletion, and the second term accounting for the death due to the bystander effect. Eq.

[3] describes the change of glucose concentration due to the consumption by tumor cells at the rate v and a Michaelis-Mentens constant, G^*. The state functions for tumor cell proliferation and tumor cell death are given as:

$$S_d\left(G(t)\right) = \left(1 - \frac{G(t)}{G(t) + G_{\min}}\right)\tanh(t) \quad [4]$$

$$S_p\left(G(t)\right) = 1 - \left(1 - \frac{G(t)}{G(t) + G_{\min}}\right)\tanh(t), \quad [5]$$

where $G_{\min}$ is the minimum glucose level required for proliferation. The parenthetical term on the right-hand side of Eq. [4] describes the dependence of cell fate (proliferation or death) on glucose availability. Observe that as $G \to 0$, $S_d(G(t)) \to 1$, which maximizes the death rate due to glucose depletion. Conversely, as $G \to \infty$, $S_d(G(t)) \to 0$, which minimizes cell death. As tumor cells may keep proliferating for some time even in a glucose free medium (please see S3 Fig of the Supporting Information), we introduced a hyperbolic tangent function of time. We hypothesize the tumor cell population is composed of two sub-populations, one that has passed the restriction point [26–29], is committed to divide, and thus does not need to be checked by the state function; and a second subpopulation that has not passed the restriction point, and thus has to be checked by the state function. The hyperbolic tangent function increases from 0 to 1 as time increases from 0 to infinity; thus, the hyperbolic tangent function on the right-hand side of Eq. [4] introduces a delay due to the duration of mitosis [30,31]. At time 0, the effect of glucose concentration described by the parenthetical term is multiplied by tanh(0), and becomes 0. This means the effect of glucose concentration is not sensed by cells immediately. At a later time, as tanh(t) increases to 1, the effect of glucose concentration increases until fully sensed by the cells. Afterwards, any further mitosis

is fully affected by glucose concentration through the state function. Note that we have $S_d(G(t)) + S_p(G(t)) = 1$.

Table 1. The definitions, units, and source for the model parameters.

Parameter	Definitions	Units	Source
k_p	Proliferation rate	day^{-1}	Calibrated
k_d	Death rate due to starvation	day^{-1}	Calibrated
k_{bys}	Death rate due to bystander effect	day^{-1}	Calibrated
θ	Carrying capacity	cells	Assigned from literature [32]
v	General glucose consumption	$mM \cdot cell^{-1} \cdot day^{-1}$	Calibrated
G^*	Michaelis-Menten constant	mM	Assigned from literature [12]
$G_{\min}$	Minimum glucose level for uptake	mM	Assigned from literature [12]

Eqs. [1] – [5] can then be used to generate a family of three models by making a small set of simplifying assumptions. If we remove cell death due to the bystander effect in Eqs. [1] and [2], we create another coupled system. Similarly, if we remove cell death due to glucose depletion in Eqs. [1] and [2], we construct a third coupled system. These three sets of equations provide our three-member model family which we then subject to a model selection operation to identify the most parsimonious model.

2.5 Model calibrations

The model described in the previous section was calibrated to experimentally measured live and dead cell time courses (described in Section 2.3), with the initial glucose concentration and confluence serving as the initial conditions. Recall that the overall goal was to calibrate model parameters against a test data set, and then use the subsequent parameterized model to predict tumor cell numbers in a validation cohort. To achieve this goal we performed a series of three

240 calibrations for each cell line: one in which the parameters were calibrated for each individual time
241 course, another in which the parameters were calibrated globally (i.e., a single set of parameters
242 for the entire cohort/test set), and in the third in which we combined the results from the first two
243 approaches so that some parameters were calibrated globally and others calibrated individually as
244 a function of initial conditions.

245 In the first calibration scenario, the measured live and dead tumor cell time courses were
246 independently fit to the model (i.e., Eqs. [1] – [5]) to produce separate estimates for each model
247 parameter within each cell line. The resulting parameter values were then further analyzed to
248 determine if their value was a function of initial glucose level and confluence. In the second
249 calibration scenario, the measured live and dead tumor cell time courses were fit by assuming
250 model parameters were independent of initial conditions; i.e., a single set of model parameters
251 were determined to simultaneously fit all time courses (for each cell line). This approach assumed
252 that the parameter values were not affected by initial conditions and are specific to each cell line.
253 In the third calibration scenario, we assumed (based on the results of the first two calibration
254 scenarios) that the proliferation rate, k_p, the consumption rate of glucose, v, and the death rate due
255 to glucose depletion, k_d, were specific for each cell line, while the other parameter, k_{bys} was a
256 function of initial confluence and glucose levels. A Student's t-test was used to test for statistical
257 differences, between the two cell lines, of each global model parameter (i.e., k_p, v, and k_d) estimated.

258 To perform each of the above calibrations, we employed a non-linear, least squares
259 approach which seeks to minimize the residual sum of square (RSS) errors between the measured
260 data and the model described in section 2.4. We defined the system of ODEs, initial conditions,
261 and time steps in Matlab using the built-in ODE solver 'ode45' to estimate the model parameters.
262 We used a least square optimization algorithm 'lsqcurvefit' to update the parameter estimates and

minimize the RSS errors. To avoid local minima, we used Matlab's 'MultiStart' to run, in parallel, 10 optimization problems with different initial parameter guesses to identify the set of parameters that minimized the RSS error. The initial parameter guesses that led to the solution point with the lowest (best) RSS error were recorded and set to be the single-start initial points for a second round of 'lsqcurvefit' to calculate the residuals and Jacobian matrix, which cannot be acquired during the first round fitting with multiple starting points. The residuals and Jacobian matrix were used to determine the confidence interval for each parameter by calling the function, 'nlparci'. Before the fitting procedure, the initial live and dead tumor cell numbers were assigned as the average of the first three timepoints to reduce error in the estimation of the initial conditions.

2.6 Model selection

As the three models (described in Section 2.4) with the different fitting strategies (described in Section 2.5) are phenomenological in nature (i.e., they are not derived from first principles), we do not know which one, *a priori*, provides the best description of the experimental data. To address this limitation, we performed model selection *via* the Akaike Information Criteria (AIC) [33]. The AIC seeks to select the most parsimonious model by balancing goodness of fit with the number of free parameters. Given our data set, we will employ the AIC_c [34,35] which includes a correction for small sample size and is given as follows:

$$AIC_c = n\ln(RSS) + 2p + \frac{2p(p+1)}{n-p-1} \quad , \qquad [6]$$

where n is the number of data samples and p is the number of model parameters. The model with the lowest AIC_c value is selected as the most parsimonious.

2.7 Determining the dependence of model parameters on initial conditions

The results of the third calibration scenario showed that k_{bys} for the BT-474 line increased with higher initial confluence (see S4 Fig of the Supporting Information), but decreased with higher initial glucose level, while k_{bys} for the MDA-MB-231 line was not affected by initial confluence (see S5 Fig of the Supporting Information), but decreased with higher initial glucose level. The dependence of local parameter (i.e., parameters calibrated for individual time courses) values on initial conditions were determined by Pearson's partial correlation coefficient. Given this relationship, we sought to determine if there was a simple functional relation between model parameters and initial conditions. We were able to find one such relation for k_{bys} for the BT-474 cells:

$$k_{bys} = k_{bys,0} N_0 \exp(-\alpha G_0), \quad [7]$$

where N_0 is the initial confluence, G_0 is the initial glucose concentration, $k_{bys,0}$ is the maximum k_{bys} rate, and α is a decay parameter. We then fit Eq. [7] to the set of initial conditions and associated parameter estimates (with their confidence intervals) obtained from the training data set to estimate $k_{bys,0}$, α, and their respective 95% confidence interval. Thus, Eq. [7] determines a parameter surface where k_{bys} can be estimated given the initial confluence and glucose. This death rate, combined with estimates of the other global parameters (i.e., k_p, v, and k_d), can then be substituted into the Eqs. [1] – [5] to predict tumor cell number at future time points. Using an analogous procedure, a similar relation was determined for the MDA-MB-231 cells:

$$k_{bys} = k_{bys,0} \exp(-\alpha G_0) + \beta, \quad [8]$$

where the parameters are as indicated for Eq. [7], with β being a base death rate which is present in this cell line even when sufficient glucose is present. Note that N_0 does not appear in Eq. [8] as this death rate for MDA-MB-231 cells is not affected by the initial confluence. Thus, Eq. [8] also determines a parameter curve where k_{bys} can be estimated given the initial glucose level. Again, this death rate, combined with estimates of the other global parameters (i.e., k_p, v, and k_d), can then be substituted into the Eqs. [1] – [5] to predict tumor cell number at future time points.

2.8 Training and validation

The data measured from the time-resolved microscopy experiments were divided into training (75% of the data) and validation sets by random sampling. The training subset was used to calibrate the global parameters k_p, k_d, and v. We calculated the absolute value of the error between the best fit curve and measured data across the whole training set to provide an estimate of the error in the measurement (i.e., uncertainty) of the initial number of live and dead tumor cells, as required for forming a prediction on the validation set. Then, given these global parameters, and the initial conditions (i.e., G_0 and N_0) from each time course in the validation set, k_{bys} was calculated using Eq. [7] for the BT-474 line or [8] for the MDA-MB-231 line. Next, k_{bys} was combined with the global parameters and initial conditions to run the forward model *via* Eqs. [1] – [5]. The resulting predicted tumor cell number time courses (with confidence intervals) for live and dead tumor cells were compared to the corresponding measured data and the errors were tabulated. We defined 'prediction accuracy' as the fraction of measured data that fell within the 95% confidence intervals of the predicted growth curves, while accuracy for the whole validation set was determined as the average 'prediction accuracy' over all measured time courses. This training and validation process was repeated 50 times, and the average error for predicted time

courses and average overall accuracy was recorded. To evaluate the model's performance, we report the averages of the RSS, mean percent error over the time course percent error at the end of experiment, mean error over the time course, error at the end of experiment (explicit matrices are presented in S1 Table of the Supporting Information).

3. Results

3.1 Tumor cell growth with different initial conditions

Example time courses for the BT-474 cell line, with different initial confluences (i.e., seeding density) and four glucose concentrations, are shown in Fig 2a-c. For wells with low initial confluence (23.8 ± 0.5%), the number of live cells changed from day 0 to day 4 by -34.3 ± 12.3% (mean ± 95% confidence interval), -6.5 ± 10.5%, +31.4 ± 8.4%, and +35.7 ± 1.8% for the 0.2 mM, 0.5 mM, 2 mM, and 5 mM initial glucose concentrations, respectively (Fig 2a). For wells with intermediate initial confluence (35.9 ± 1.8%), the number of live cells changed from day 0 to day 4 by -63.7 ± 9.3%, -55.6 ± 3.1%, -10.4 ± 19.6%, and +14.9 ± 8.3% for the 0.2 mM, 0.5 mM, 2 mM, and 5 mM initial glucose concentrations, respectively (Fig 2b). For wells with high initial confluence (51.7 ± 1.4%), the number of live cells changed from day 0 to day 4 by -76.0 ± 1.0%, -76.3 ± 1.5%, -43.9 ± 12.4%, and -17.6 ± 7.6% for the 0.2 mM, 0.5 mM, 2 mM, and 5 mM initial glucose concentrations, respectively (Fig 2c). The average number of live cells for each experiment at the end of day 4 was significantly different among the groups with different initial conditions ($p < 1e\text{-}5$).

Fig 2. Time courses of tumor cell confluence in media with varying initial glucose levels, grouped by initial confluence. Panels a-c present confluence time courses for the BT-474 cell

line, while panels d-f present confluence time courses for the MDA-MB-231 cell line. Different colors represent the four initial glucose concentrations, and the error bars were calculated from four replicates with similar initial conditions. In each panel, cells represented by each color were seeded at the same initial confluence, but yielded significant differences in confluence at the end of the experiment. These time courses provide quantitative and dynamic data on the effects of glucose availability and confluence on tumor cell growth.

Example time courses for the MDA-MB-231 cell line, with different initial confluences and four glucose concentrations are shown in Fig 2d-f. For wells with low initial confluence (36.9 ± 1.0%), the number of live cells changed from day 0 to day 4 by -68.3 ± 7.9%, -47.7 ± 12.7%, +30.2 ± 13.5% and +32.9 ± 7.9% for the 0.2 mM, 0.5 mM, 2 mM, and 5 mM initial glucose concentrations, respectively (Fig 2d). For wells with intermediate initial confluence (56.2 ± 1.4%), the number of live cells changed from day 0 to day 4 by -63.7 ± 10.3%, -46.5 ± 10.6%, -1.8 ± 3.4%, and +13.3 ± 2.1% for the 0.2 mM, 0.5 mM, 2 mM, and 5 mM initial glucose concentrations, respectively (Fig 2e). For wells with high initial confluence (71.9 ± 1.0%), the number of live cells changed from day 0 to day 4 by -41.5 ± 11.6%, -33.6 ± 9.2%, -10.0 ± 3.2%, and -1.6 ± 2.9% for the 0.2 mM, 0.5 mM, 2 mM, and 5 mM initial glucose concentrations, respectively (Fig 2f). The average number of live cells for each experiment at the end of day 4 was significantly different among the groups with different initial conditions ($p < 1e\text{-}5$).

3.2 Model calibration

The model characterized by Eqs. [1] – [5] featuring three global parameters (k_p, k_d, and v), and one local parameter dependent on initial conditions (k_{bys}) was selected by the AIC_c as the most

parsimonious and employed for all subsequent analysis (details provided in S2 Table of the Supporting Information). The estimates for the three global parameters and their 95% confidence intervals for both BT-474 and MDA-MB-231 cells are shown in Table 2. The proliferation and glucose consumption rates of the BT-474 cells were significantly lower than the MDA-MB-231 cells ($p < 1e\text{-}4$), while the death rate due to glucose depletion of the BT-474 cells was higher than MDA-MB-231 cells ($p < 1e\text{-}4$).

Table 2. Parameter estimates obtained from the global calibration procedure

Parameters	Cell Line		p-value
	BT-474	MDA-MB-231	
k_p (day^{-1})	0.092 ± 0.002	0.14 ± 0.003	$< 10^{-4}$
k_d (day^{-1})	0.13 ± 0.013	0.041 ± 0.006	$< 10^{-4}$
v ($\times 10^{-5}$ $mM \cdot cell^{-1} \cdot day^{-1}$)	2.68 ± 0.10	4.48 ± 0.15	$< 10^{-4}$

As described in Section 2.5, the selected model (i.e., the model with globally calibrated k_p, k_d, and v and locally calibrated k_{bys}) was fit to the measured experimental data. The mean percent error across all timepoints, mean percent error at the end of experiment mean error across all timepoints, and mean error at the end of experiment are reported in Table 3. The model was able to provide an accurate description of the time course data over a wide range of initial conditions with mean percent error and mean percent error at the end of experiments below 7% for live cells in both cell lines (Table 3). For the dead cells, the model performs more modestly with mean percent error between 16% and 67% over all initial conditions. Importantly, the mean error, either across all timepoints or at the end of experiment, was < 2% for both live and dead cells in both cell lines. This suggests the higher percent error of dead cells is due to the small number of dead cells as compared to the number of live cells.

Table 3. Evaluation of fitting quality with selected model for both cell lines

	Cell Line			
	BT-474		MDA-MB-231	
	Live	Dead	Live	Dead
RSS	1.35	1.13	1.87	1.31
Mean % Error	0.09 ± 0.23	66.01 ± 4.13	0.59 ± 0.22	18.17 ± 1.66
% Error EoE*	-0.78 ± 3.44	44.37 ± 30.63	6.22 ± 2.31	16.92 ± 18.43
Mean Error	-0.01 ± 0.06	0.61 ± 0.05	-0.03 ± 0.07	0.87 ± 0.05
Error EoE*	0.24 ± 0.57	0.56 ± 0.42	1.65 ± 0.42	0.42 ± 0.43

*EoE = End of Experiment

3.3 Relationship between bystander effect death rate (k_{bys}) and initial conditions

For the BT-474 cells, the death rate due to the bystander effect, k_{bys}, was found to increase with increasing initial confluence, with a partial correlation coefficient of 0.66 ($p < 1e-4$). For 8 of 10 initial glucose levels tested (0, 0.1, 0.2, 0.5, 0.8, 1, 2, and 5 mM), the bystander effect death rate was positively correlated with initial confluence, with correlation coefficients all above 0.74 ($p < 0.01$). Estimates of k_{bys} were plotted against the initial confluence level (Fig 3a). The correlation coefficients between k_{bys} and initial confluence level were 0.75, 0.94, 0.81, 0.81, and 0.84 for initial glucose level of 0.2 mM, 0.5 mM, 1 mM, 2 mM, and 5 mM, respectively. For the highest two initial glucose levels (8 and 10 mM), there was no significant correlation between the bystander effect death rate and the initial confluence ($p > 0.1$). The bystander effect was found to decrease with increasing initial glucose level, with a partial correlation coefficient of -0.74 ($p < 1e-4$). For low (23.8 ± 0.5%), intermediate (35.9 ± 1.8%), and high (51.7 ± 1.4%) initial

410 confluences, there are significant correlations between k_{bys} and the initial glucose level (Fig 3b),
411 with correlation coefficient of -0.44 ($p < 0.01$), -0.80 ($p <$ 1e-4), and -0.91 ($p <$ 1e-4). Given these
412 relationships, k_{bys} was fit to each initial condition as described in Section 2.7 (see Eq. [7]), yielding
413 a $k_{bys,0}$ of 2.37 ± 0.13 × 10^{-5} mM·cell^{-1}·day^{-1} and an α of 0.13 ± 0.029 mM^{-1}. With $k_{bys,0}$ and α
414 identified, Eq. [7] defines a parameter surface where we can obtain the value of k_{bys} for any initial
415 confluence and glucose level within the experimentally measured range (Fig 3d).
416
417 **Fig 3. Relationship between bystander effect death rate (k_{bys}) and initial conditions**. Panel a
418 presents estimates of the death rate due to the bystander effect, k_{bys}, as a function of different initial
419 confluence and glucose levels for BT-474 cells. For each glucose level, k_{bys} increases with higher
420 initial confluence, where the lowest initial glucose level increases k_{bys} the most. Panel b indicates
421 that k_{bys} increases with higher initial confluence and decreases with higher initial glucose level for
422 the BT-474 line. (Error bars were calculated from the four wells with similar initial conditions.)
423 Panel c shows that k_{bys} decreases with higher initial glucose level for the MDA-MB-231 cell line.
424 (Error bars were calculated from the twelve wells with same initial glucose level.) Panel d shows
425 the parameter surface for the BT-474 cell line, where k_{bys} is displayed as function of initial
426 confluence and glucose level, with blue dots representing calibrated estimates of k_{bys}. Panel e
427 indicates how k_{bys} decreases with initial glucose level for the MDA-MB-231 line, with shaded area
428 between solid red curves showing the 95% confidence interval. The blue dots represent the
429 calibrated estimates of k_{bys}. The fitted surface and curve in panels d and e, respectively, is used to
430 assign k_{bys} as a function of initial confluence and glucose concentration in the validation data set.
431

For the MDA-MB-231 cells, there was no significant correlation between the death rate due to the bystander effect and the initial confluence with a partial correlation coefficient of -0.03 ($p = 0.73$). The bystander effect was found to decrease with increasing initial glucose level (Fig 3c), with a partial correlation coefficient of -0.72 ($p < 1e\text{-}4$). For the low (36.9 ± 1.0%), intermediate (56.2 ± 1.4%), and high (71.9 ± 1.0%) initial confluences, there are significant correlations between the death rate due to the bystander effect and the initial glucose level, with correlation coefficient of -0.76 ($p < 1e\text{-}4$), -0.76 ($p < 1e\text{-}4$), and -0.66 ($p < 1e\text{-}4$). Given these relationships, k_{bys} was fit to each initial condition as described in Section 2.7 (see Eq. [8]), yielding a $k_{bys,0}$ of $0.71 \pm 0.067 \times 10^{-5}$ mM·cell^{-1}·day^{-1} , an α of 0.98 ± 0.23 mM^{-1} and a β of 0.22 ± 0.053 mM·cell^{-1}·day^{-1}. With $k_{bys,0}$, α, and β identified, Eq. [8] defines a parameter curve where we can obtain the value of k_{bys} for any initial glucose level within the experimentally measured range (Fig 3e).

3.4 Evaluation of model performance through training and validation

In each round of training and validation, 75% of the whole dataset was randomly selected for a training set, with the remainder assigned to the validation set. The selected model (i.e., the model with globally calibrated k_p, k_d, and v and locally calibrated k_{bys}) was calibrated to each time course in the training set to obtain estimates and confidence intervals for the model parameters.

For the BT-474 cells, we reported the model performance during training (Table 4). The average mean percent error across all timepoints, and the average percent error at the end of experiment were < 1% for live cells. Although the average mean percent error across all timepoints, and the average percent error at the end of experiment were > 45% for dead cells, the average mean error across all timepoints and average error at the end of experiment were < 1% for both

455 live and dead cells. The average uncertainty across 50 training sets for live and dead cells were
456 6.88 ± 0.09% and 30.83 ± 0.15%, respectively.

457 **Table 4. Summary of model calibration across 50 training sets**

	Cell Line			
	BT-474		MDA-MB231	
	Live	Dead	Live	Dead
RSS	1.09 ± 0.02	0.88 ± 0.01	2.22 ± 0.02	0.96 ± 0.02
Mean % Error	0.24 ± 0.04	72.42 ±1.23	-0.26 ± 0.06	12.29 ± 0.29
% Error EoE*	-0.90 ± 0.28	49.88 ± 2.41	6.01 ± 0.24	11.25 ±1.62
Mean Error	0.07 ± 0.01	0.69 ± 0.01	-0.52 ± 0.03	0.19 ± 0.02
Error EoE*	0.25 ± 0.04	0.69 ±0.03	1.44 ±0.05	-0.62 ± 0.05
Uncertainty	6.88 ± 0.09	30.83 ± 0.15	5.17 ± 0.05	16.78 ± 0.12

458 *EoE = End of Experiment

459

460 For the BT-474 cells, the parameters $k_{bys,0}$ and α in Eq. [7] were estimated as described in
461 section 2.7 and a specific parameter surface of k_{bys} was determined. The uncertainty calculated
462 from fitting the data of the training set to the model was used to estimate the confidence interval
463 of the initial confluence from the validation set. The initial conditions (i.e., initial glucose level
464 and confluence) from the validation set were used with Eq. [7] to identify the value of k_{bys} to be
465 used, in conjunction with the three global parameters (k_p, v, and k_d and their respective confidence
466 intervals) in Eqs. [1] – [5] to run the forward model. This process was repeated 50 times to obtain
467 an average RSS, average mean percent error, average percent error at the end of experiment,
468 average mean error, average error at the end of experiment, and accuracy (Table 5). The accuracy
469 was defined as the percent of data points falling within the 95% confidence interval of the predicted
470 values. The average RSS was 1.45 ± 0.09 and 1.22 ± 0.09 for live and dead cells, respectively,

while the accuracy was 77.2 ± 6.3% and 50.5 ± 7.5% for live and dead cells, respectively. The average mean percent error across all timepoints and the average percent error at the end of experiment were both < 2% for live cells. Although the average mean percent error across all timepoints and average percent error at the end of the experiment for dead cells can be as high as > 150%, the average mean error across all timepoints and average error at the end of experiment were < 3% for both live and dead cells. Fig 4 presents representative prediction results compared with measured data on BT-474 cells from the same round of training and validation (Fig 4).

Table 5. Evaluation of prediction across 50 rounds of training and validation

	Cell Line			
	BT-474		MDA-MB231	
	Live	Dead	Live	Dead
RSS	1.45 ± 0.09	1.22 ± 0.09	1.69 ± 0.10	1.35 ± 0.12
Mean % Error	-1.96 ± 0.54	153.18 ± 9.07	-0.59 ± 0.47	25.22 ± 1.27
% Error EoE*	-5.78 ±1.49	168.20 ±16.31	7.04 ± 1.62	47.54 ± 9.95
Mean Error	-0.78 ± 0.15	1.57 ± 0.13	-1.15 ± 0.15	0.87 ± 0.16
Error EoE*	-1.66 ± 0.33	2.67 ± 0.27	-0.12 ± 0.38	1.09 ± 0.43
Accuracy	77.2 ± 6.3	50.5 ± 7.5	87.2 ± 5.1	66.7 ± 7.0

*EoE = End of Experiment

Fig 4. Model predictions for BT-474 cells. Example model predictions from one validation set of BT-474 cells are shown in dashed lines. In each panel, data measured from experiments are shown in circles, while the 95% confidence intervals for the predicted tumor cell growth and dead cell accumulation numbers are shown with shaded regions between the solid curves; with blue indicating live cells, and red indicating dead cells. The initial glucose level is shown above each

plot. For this validation set, the model prediction accuracy was 72.2 ± 8.6% and 49.3 ± 10.0% for live and dead cells, respectively.

For the MDA-MB-231 cells, we reported the model performance during training (Table 4). The average mean percent error across all timepoints, and the average percent error at the end of experiment were < 13% for both live and dead cells. The average mean error across all timepoints and average error at the end of experiment were < 2% for both live and dead cells. The average uncertainty across 50 training sets for live cells and dead cells were 5.17 ± 0.05% and 16.78 ± 0.12% respectively.

For the MDA-MB-231 cells, the parameters $k_{bys,0}$, α, and β in Eq. [8] were estimated as described in section 2.7 and a specific parameter curve for k_{bys} was determined. The uncertainty calculated from fitting the data of the training set to the model was used to estimate the confidence interval of the initial confluence from the validation set. The initial condition (i.e., initial glucose level) from the validation set were used with Eq. [8] to identify the value of k_{bys} to be used, in conjunction with the three global parameters (k_p, v, and k_d and their respective confidence intervals) in Eqs. [1-5] to run the forward model. This process was repeated 50 times to obtain average RSS, average mean percent error, average percent error at the end of experiment, average mean error, average error at the end of experiment, and accuracy (Table 5). The accuracy was defined as the percent of data points falling within the 95% confidence interval of the predicted values. The average RSS was 1.69 ± 0.10 and 1.35 ± 0.12 for live and dead cells, respectively, while the accuracy was 87.2 ± 5.1% and 66.7 ± 7.0% for live and dead cells, respectively. The average mean percent error across all timepoints and the average of percent error at the end of experiment were both < 8% for live cells. Although the average percent error across all timepoints and average error

509 at the end of experiment for dead cells were > 25%, the average mean error across all timepoints
510 and average error at the end of experiment were < 2% for both live and dead cells. Fig 5 presents
511 representative prediction results compared with measured data on MDA-MB-231 cells from the
512 same round of training and validation (Fig 5).
513
514 **Fig 5. Model predictions for MDA-MB-231 cells.** Example model predictions from one
515 validation set of MDA-MB-231 cells are shown in dashed lines. In each panel, data measured from
516 experiments are shown in circles, while the 95% confidence intervals for the predicted tumor cell
517 growth and dead cell accumulation numbers are shown with shaded regions between the solid
518 curves; with blue indicating live cells, and red indicating dead cells. The initial glucose level is
519 shown above each plot. For this validation set, the model prediction accuracy was 86.9 ± 7.4% and
520 69.6 ± 9.4% for live and dead cells, respectively.
521

522 # 4. Discussion

523 This study sought to develop an experimental-mathematical approach to quantify tumor
524 cell proliferation as a function of glucose availability. This allowed us to quantify important cell
525 phenotypes related to proliferation and cell death, and then use the model to predict the temporal
526 change in tumor cell number. To accomplish this task, we proposed a family of three models in
527 which each member of the family consisted of a system of coupled ordinary differential equations
528 (ODEs) describing the rate of change of living and dead tumor cell number and glucose
529 concentration. The complete model considered tumor cell proliferation, cell death due to glucose
530 depletion, the bystander effect quantifying the effects of dead cells accumulated in the environment,
531 and the consumption of glucose. To calibrate the model, we acquired time-resolved microscopy

images to generate confluence time courses of both live and dead tumor cells over an array of initial glucose concentrations and confluences. We then fit the data to all the models and selected the most parsimonious model with the lowest AIC_c value. During the model selection, we determined that the proliferation rate, death rate due to glucose depletion, and consumption rate of glucose were three parameters that depended only on cell line and not initial conditions; thus, they could be fit as global parameters. Conversely, it was determined that the death rate due to the bystander effect was a local parameter that varied with the initial conditions. We therefore investigated the relationship between this parameter and the initial conditions for each cell line. Finally, we evaluated the performance of the selected model through training and validation.

Mathematical models have been developed to describe cancer cell metabolism from different perspectives. For example, Mendoza-Juez *et al*. [12] focused on glucose and lactate as the main nutrient resources, and thus concentrated on the dynamic development of two subpopulations with different metabolic behavior. Conversely, Astanin *et al*. [11], while similarly modeling oxidative and glycolytic subpopulations, also included oxygen consumption and ATP production in their model. Without measurement, this system was characterized with typical values of dimensionless parameters for simulation. These models, with a heavy reliance on a large number of unmeasured parameters, can be difficult to calibrate and therefore difficult to apply within an experimental-predictive framework.

The family of models proposed in this study was derived from the work of Mendoza-Juez *et al*. [12], but included two simplifications. First, we viewed the live tumor cells as a single population instead of several subpopulations with different metabolic phenotypes. At the cost of losing detailed phenotype or subpopulation dynamics, the simplification allowed us to practically connect the accessible experimental data and mathematical modeling. Given sufficient

555 experimental data (*via* the time-resolved microscopy data) and known initial conditions, we were
556 able to perform a direct calibration of our model. This experimental-computational approach was
557 applied in two cell lines, representing very different breast cancer subtypes. We found the
558 proliferation rate of the BT-474 cells was statistically lower than MDA-MB-231 cells. We found
559 the death rate due to glucose depletion for the BT-474 cells was statistically higher than that of the
560 MDA-MB-231, while the consumption rate of glucose for the BT-474 cells is statistically lower
561 than that of the MDA-MB-231 cells. We concluded that while MDA-MB-231 cells consume
562 glucose at a higher rate (thereby enabling more rapid growth and division), the glucose level
563 required for proliferation was lower than that of the BT-474 cells. These results serve to quantify
564 the well-established experimental observations that MDA-MB-231 is more aggressive than the
565 BT-474 cell line [36–38]. To the best of our knowledge, this represents the first time these
566 important cellular parameters have been quantified within a rigorous modeling framework. Once
567 calibrated, our model could be used to predict the number of live tumor cells, validated by direct
568 comparison with experimental data.

569 The death rate due to the bystander effect, k_{bys}, proved to be dependent on the initial
570 conditions. Therefore, k_{bys} was estimated individually for each set of initial conditions, and not
571 considered as a global parameter. The bystander effect parameter for both cell lines became
572 significantly lower when the initial glucose level increased. This parameter significantly increased
573 when the initial confluence for the BT-474 cell line increased, but was not affected by initial
574 confluence in the MDA-MB-231 line. This difference indicates there could be different
575 mechanisms underlying the bystander effect in different cell types. Studies concerning the
576 bystander effect can be identified into two separate groups. In the first case, the bystander effect is
577 proven to be mediated by degree of GJIC (gap junction intercellular communication) capacity [39–

41]. Since BT-474 cells are mass cells with robust cell-cell adhesion and close cell contact within clusters, they have high GJIC capacity. However, MDA-MB-231 cells do not form clusters with strong cell contact and exhibit low GJIC level. These are consistent with the results that the death rate of bystander effect for BT-474 increases with initial confluence, but is not affected by initial confluence for MDA-MB-231. In the second case, killing of the non-treated cells involves the release of one or more soluble factors, such as apoptosis inducing signals [42], extracellular vesicles [43], or oxidizing diffusive factors [44]. In our study, the death rate of bystander effect for MDA-MB-231 is not affected by initial confluence, implying there would be at least one soluble factor regulated by metabolism. Furthermore, there could be multiple mechanisms underlying the bystander effect for a given cell line, considering the death rate of bystander effect for BT-474 is affected by both confluence and glucose level. While further work including the identification and quantification of these factors is required to support our work, this experimental-computational approach allows us to analyze the characteristics of bystander effect for the cell line tested. This could provide guidance on choice of enhanced therapies utilizing the bystander effect (e.g., GJIC enhancement) for synergistic effect [41].

The present work assumed glucose consumption was entirely captured by the temporal change of tumor cell number, which is most likely an oversimplification. To address this limitation, the development of a method for time-resolved measurement of glucose dynamics is required. Further quantification and mathematical description of the glucose dynamics (e.g., a FRET nanosensor for glucose [45]) would provide additional time-resolved data that would enable extension of the model to more precisely describe glucose kinetics. The hyperbolic tangent function introduced in Eqs. [4] and [5] is sufficient to characterize the growth curves in our current research and we chose to keep it simple to avoid overfitting. However other sigmoidal functions

601 of time that are more directly related to phenomena affecting glucose dynamics should be explored
602 to refine to the model by introducing more biology. In particular, the state functions (i.e., Eqs. [4]
603 and [5]) have the potential to be extended to characterize glucose utilization as a function of cell
604 cycle [46–48]. Additionally, our model could then be extended to account for additional nutrients
605 of metabolic interest (e.g., lactate, intermediate products between glycolysis and oxidative
606 phosphorylation, and oxygen). Such extension would, of course, require additional time resolved
607 measurements to parameterize the model. We also limited the application of our model to only
608 two different breast cancer cell lines, but given their differences in parameter values, systematic
609 investigation of a range of cells lines is warranted.

610 In summary, the temporal change of tumor cell number with different initial glucose levels
611 and seeding densities was tracked with time-resolved microscopy. These data were used to
612 calibrate a mathematical model describing cell proliferation and death as a function of glucose
613 dynamics, which was then used to predict tumor cell dynamics in a separate validation set. This
614 approach yielded an accuracy of > 75% for predicting the change in the number of living cells over
615 time, and is readily extendable to account for and predict the effects of interventions designed to
616 affect glucose metabolism.

617

618 # 5. Conclusion

619 We have developed and validated an experimental-mathematical approach that is capable
620 of accurately predicting how glucose availability influences tumor cell proliferation. The method
621 was applied in two commonly studied breast cancer cells in which we were able to quantify rates
622 directly reporting on cell proliferation, death due to glucose starvation, death due to the bystander
623 effect, and overall glucose consumption. Furthermore, this framework is directly applicable to

624 other tumor cell lines. The integration of mechanism-based modeling and time-resolved
625 microscopy is a powerful, and flexible, approach to systematically investigate glucose dynamics
626 related tumor cell growth.
627

628 Acknowledgement

629 We are grateful for the members of the Center for Computational Oncology at The University of
630 Texas at Austin for many helpful and informative discussions. We thank Dr. Heiko Enderling and
631 Dr. Chad Quarles for constructive feedback at early stages of this project.
632

633 Author Contributions

634 **Conceptualization:** Jianchen Yang, Jack Virostko, Thomas E. Yankeelov
635 **Data curation:** Jianchen Yang
636 **Formal analysis:** Jianchen Yang
637 **Funding acquisition:** Thomas E. Yankeelov
638 **Investigation:** Jianchen Yang
639 **Methodology:** Jianchen Yang, David A. Hormuth II, Thomas E. Yankeelov
640 **Project administration:** Thomas E. Yankeelov
641 **Software:** Jianchen Yang, Junyan Liu
642 **Supervision:** Jack Virostko, Amy Brock, Jeanne Kowalski, Thomas E. Yankeelov
643 **Validation:** Jianchen Yang
644 **Writing - original draft:** Jianchen Yang, Thomas E. Yankeelov

Writing - review & editing: Jianchen Yang, Jack Virostko, David A. Hormuth II, Amy Brock, Thomas E. Yankeelov

References

1. Warburg O. On the Origin of Cancer Cells. Science. 1956;123: 309–314.

2. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. Science. 2009;324: 1029–1033. doi:10.1126/science.1160809

3. Dixon KC. The Pasteur Effect and Its Mechanism. Biol Rev. 1937;12: 431–460. doi:10.1111/j.1469-185X.1937.tb00641.x

4. Sonveaux P, Végran F, Schroeder T, Wergin MC, Verrax J, Rabbani ZN, et al. Targeting lactate-fueled respiration selectively kills hypoxic tumor cells in mice. 1 Dec 2008. doi:10.1172/JCI36843

5. DeBerardinis RJ, Lum JJ, Hatzivassiliou G, Thompson CB. The Biology of Cancer: Metabolic Reprogramming Fuels Cell Growth and Proliferation. Cell Metab. 2008;7: 11–20. doi:10.1016/j.cmet.2007.10.002

6. Pavlova NN, Thompson CB. The Emerging Hallmarks of Cancer Metabolism. Cell Metab. 2016;23: 27–47. doi:10.1016/j.cmet.2015.12.006

7. El Mjiyad N, Caro-Maldonado A, Ramírez-Peinado S, Muñoz-Pinedo C. Sugar-free approaches to cancer cell killing. Oncogene. 2010;30: 253.

8. Kennedy KM, Dewhirst MW. Tumor metabolism of lactate: the influence and therapeutic potential for MCT and CD147 regulation. Future Oncol. 2009;6: 127–148. doi:10.2217/fon.09.145

9. Tennant DA, Durán RV, Gottlieb E. Targeting metabolic transformation for cancer therapy. Nat Rev Cancer. 2010;10: 267.

10. Vander Heiden MG. Targeting cancer metabolism: a therapeutic window opens. Nat Rev Drug Discov. 2011;10: 671.

11. Astanin S, Preziosi L. Mathematical modelling of the Warburg effect in tumour cords. J Theor Biol. 2009;258: 578–590. doi:10.1016/j.jtbi.2009.01.034

12. Mendoza-Juez B, Martínez-González A, Calvo GF, Pérez-García VM. A Mathematical Model for the Glucose-Lactate Metabolism of in Vitro Cancer Cells. Bull Math Biol. 2012;74: 1125–1142. doi:10.1007/s11538-011-9711-z

677 13. McGillen JB, Kelly CJ, Martínez-González A, Martin NK, Gaffney EA, Maini PK, et al.
678 Glucose–lactate metabolic cooperation in cancer: Insights from a spatial mathematical
679 model and implications for targeted therapy. J Theor Biol. 2014;361: 190–203.
680 doi:10.1016/j.jtbi.2014.09.018

681 14. Phipps C, Molavian H, Kohandel M. A microscale mathematical model for metabolic
682 symbiosis: Investigating the effects of metabolic inhibition on ATP turnover in tumors. J
683 Theor Biol. 2015;366: 103–114. doi:10.1016/j.jtbi.2014.11.016

684 15. Kianercy A, Veltri R, Pienta KJ. Critical transitions in a game theoretic model of tumour
685 metabolism. Interface Focus. 2014;4: 20140014–20140014. doi:10.1098/rsfs.2014.0014

686 16. Chen Y, Wang H, Zhang J, Chen K, Li Y. Simulation of avascular tumor growth by agent-
687 based game model involving phenotype-phenotype interactions. Sci Rep. 2016;5: 17992.
688 doi:10.1038/srep17992

689 17. Robertson-Tessi M, Gillies RJ, Gatenby RA, Anderson ARA. Impact of Metabolic
690 Heterogeneity on Tumor Growth, Invasion, and Treatment Outcomes. Cancer Res. 2015;75:
691 1567–1579. doi:10.1158/0008-5472.CAN-14-1428

692 18. Anderson ARA, Weaver AM, Cummings PT, Quaranta V. Tumor Morphology and
693 Phenotypic Evolution Driven by Selective Pressure from the Microenvironment. Cell.
694 2006;127: 905–915. doi:10.1016/j.cell.2006.09.042

695 19. Elstrom RL, Bauer DE, Buzzai M, Karnauskas R, Harris MH, Plas DR, et al. Akt
696 Stimulates Aerobic Glycolysis in Cancer Cells. Cancer Res. 2004;64: 3892–3899.
697 doi:10.1158/0008-5472.CAN-03-2904

698 20. Voisin P, Bouchaud V, Merle M, Diolez P, Duffy L, Klint K, et al. Microglia in Close
699 Vicinity of Glioma Cells: Correlation Between Phenotype and Metabolic Alterations. Front
700 Neuroenergetics. 2010;2: 131. doi:10.3389/fnene.2010.00131

701 21. Foulkes WD, Reis-Filho JS. Triple-Negative Breast Cancer. N Engl J Med. 2010; 11.

702 22. Arteaga CL, Sliwkowski MX, Osborne CK, Perez EA, Puglisi F, Gianni L. Treatment of
703 HER2-positive breast cancer: current status and future perspectives. Nat Rev Clin Oncol.
704 2012;9: 16–32. doi:10.1038/nrclinonc.2011.177

705 23. Kenny PA, Lee GY, Myers CA, Neve RM, Semeiks JR, Spellman PT, et al. The
706 morphologies of breast cancer cell lines in three-dimensional assays correlate with their
707 profiles of gene expression. Mol Oncol. 2007;1: 84–96. doi:10.1016/j.molonc.2007.02.004

708 24. Freeman SM, Abboud CN, Whartenby KA, Packman CH, Koeplin DS, Moolten FL, et al.
709 The “Bystander Effect”: Tumor Regression When a Fraction of the Tumor Mass Is
710 Genetically Modified. : 11.

711 25. Li Bi W, Parysek LM, Warnick R, Stambrook PJ. In Vitro Evidence That Metabolic
712 Cooperation Is Responsible for the Bystander Effect Observed with HSV tk Retroviral
713 Gene Therapy. Hum Gene Ther. 1993;4: 725–731. doi:10.1089/hum.1993.4.6-725

714 26. Pardee AB. A Restriction Point for Control of Normal Animal Cell Proliferation. Proc Natl
715 Acad Sci. 1974;71: 1286–1290. doi:10.1073/pnas.71.4.1286

716 27. Zetterberg A, Larsson O, Wiman KG. What is the restriction point? Curr Opin Cell Biol.
717 1995;7: 835–842. doi:10.1016/0955-0674(95)80067-0

718 28. Planas-Silva MD, Weinberg RA. The restriction point and control of cell proliferation. Curr
719 Opin Cell Biol. 1997;9: 768–772. doi:10.1016/S0955-0674(97)80076-2

720 29. Blagosklonny MV, Pardee AB. The Restriction Point of the Cell Cycle. Cell Cycle. 2002;1:
721 102–109. doi:10.4161/cc.1.2.108

722 30. Mcintosh JR, Hepler PK, Wie DGV. Model for Mitosis. Nature. 1969;224: 659–663.
723 doi:10.1038/224659a0

724 31. King RW, Jackson PK, Kirschner MW. Mitosis in transition. Cell. 1994;79: 563–571.
725 doi:10.1016/0092-8674(94)90542-8

726 32. SOP: Thawing, Propagation and Cryopreservation of NCI-PBCF-HTB26 (MDA-
727 MB-231). : 25.

728 33. Akaike H. A New Look at the Statistical Model Identification. In: Parzen E, Tanabe K,
729 Kitagawa G, editors. Selected Papers of Hirotugu Akaike. New York, NY: Springer New
730 York; 1998. pp. 215–222. doi:10.1007/978-1-4612-1694-0_16

731 34. Cavanaugh JE. Unifying the derivations for the Akaike and corrected Akaike information
732 criteria. Stat Probab Lett. 1997;33: 201–208. doi:10.1016/S0167-7152(96)00128-9

733 35. Burnham KP, Anderson DR. Multimodel Inference: Understanding AIC and BIC in Model
734 Selection. Sociol Methods Res. 2004;33: 261–304. doi:10.1177/0049124104268644

735 36. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, et al. A collection of breast
736 cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell.
737 2006;10: 515–527. doi:10.1016/j.ccr.2006.10.008

738 37. Holliday DL, Speirs V. Choosing the right cell line for breast cancer research. Breast
739 Cancer Res. 2011;13: 215. doi:10.1186/bcr2889

740 38. Chavez KJ, Garimella SV, Lipkowitz S. Triple negative breast cancer cell lines: One tool in
741 the search for better treatment of triple negative breast cancer. Eng-Wong J, Zujewski JA,
742 editors. Breast Dis. 2011;32: 35–48. doi:10.3233/BD-2010-0307

743 39. Asklund T, Appelskog IB, Ammerpohl O, Langmoen IA, Dilber MS, Aints A, et al. Gap
744 junction-mediated bystander effect in primary cultures of human malignant gliomas with

recombinant expression of the HSVtk gene. Exp Cell Res. 2003;284: 183–193. doi:10.1016/S0014-4827(02)00052-6

40. Matono S, Tanaka T, Sueyoshi S, Yamana H, Fujita H, Shirouzu K. Bystander effect in suicide gene therapy is directly proportional to the degree of gap junctional intercellular communication in esophageal cancer. Int J Oncol. 2003;23: 1309–1315. doi:10.3892/ijo.23.5.1309

41. Yang Y, Qin S-K, Wu Q, Wang Z-S, Zheng R-S, Tong X-H, et al. Connexin-dependent gap junction enhancement is involved in the synergistic effect of sorafenib and all-trans retinoic acid on HCC growth inhibition. Oncol Rep. 2014;31: 540–550. doi:10.3892/or.2013.2894

42. Lyng FM, Seymour CB, Mothersill C. Production of a signal by irradiated cells which leads to a response in unirradiated cells characteristic of initiation of apoptosis. Br J Cancer. 2000;83: 1223–1230. doi:10.1054/bjoc.2000.1433

43. Xu S, Wang J, Ding N, Hu W, Zhang X, Wang B, et al. Exosome-mediated microRNA transfer plays a role in radiation-induced bystander effect. RNA Biol. 2016. Available: https://www.tandfonline.com/doi/abs/10.1080/15476286.2015.1100795

44. Azzam EI, Toledo SM de, Little JB. Oxidative metabolism, gap junctions and the ionizing radiation-induced bystander effect. Oncogene. 2003;22: 7050–7057. doi:10.1038/sj.onc.1206961

45. Takanaga H, Chaudhuri B, Frommer WB. GLUT1 and GLUT9 as major contributors to glucose influx in HepG2 cells identified by a high sensitivity intramolecular FRET glucose sensor. Biochim Biophys Acta BBA - Biomembr. 2008;1778: 1091–1099. doi:10.1016/j.bbamem.2007.11.015

46. Kee N, Sivalingam S, Boonstra R, Wojtowicz JM. The utility of Ki-67 and BrdU as proliferative markers of adult neurogenesis. J Neurosci Methods. 2002;115: 97–105. doi:10.1016/S0165-0270(02)00007-9

47. Mandyam CD, Harburg GC, Eisch AJ. Determination of key aspects of precursor cell proliferation, cell cycle length and kinetics in the adult mouse subgranular zone. Neuroscience. 2007;146: 108–122. doi:10.1016/j.neuroscience.2006.12.064

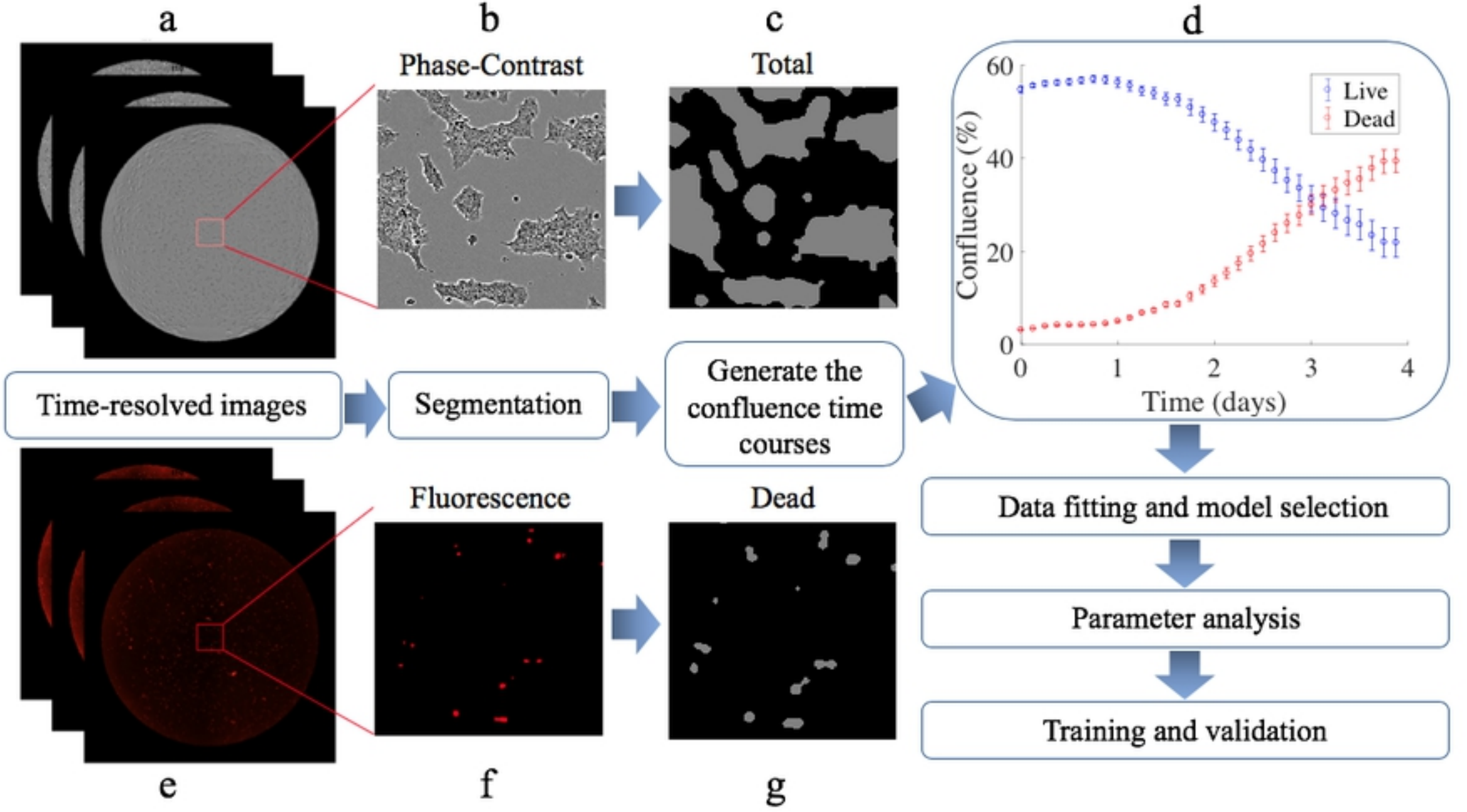
48. Bajar BT, Lam AJ, Badiee RK, Oh Y-H, Chu J, Zhou XX, et al. Fluorescent indicators for simultaneous reporting of all four cell cycle phases. Nat Methods. 2016;13: 993–996. doi:10.1038/nmeth.4045

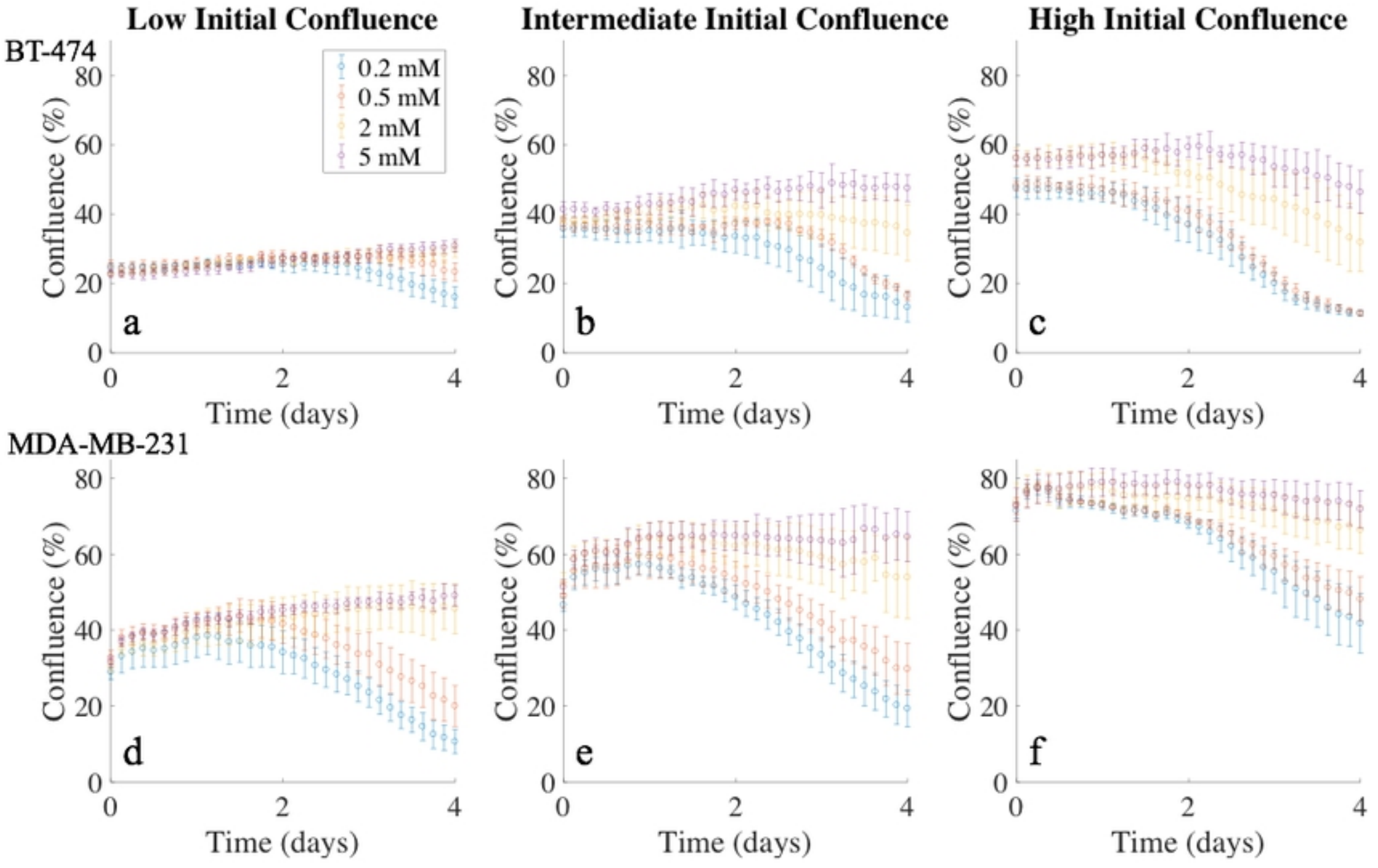
Supporting information

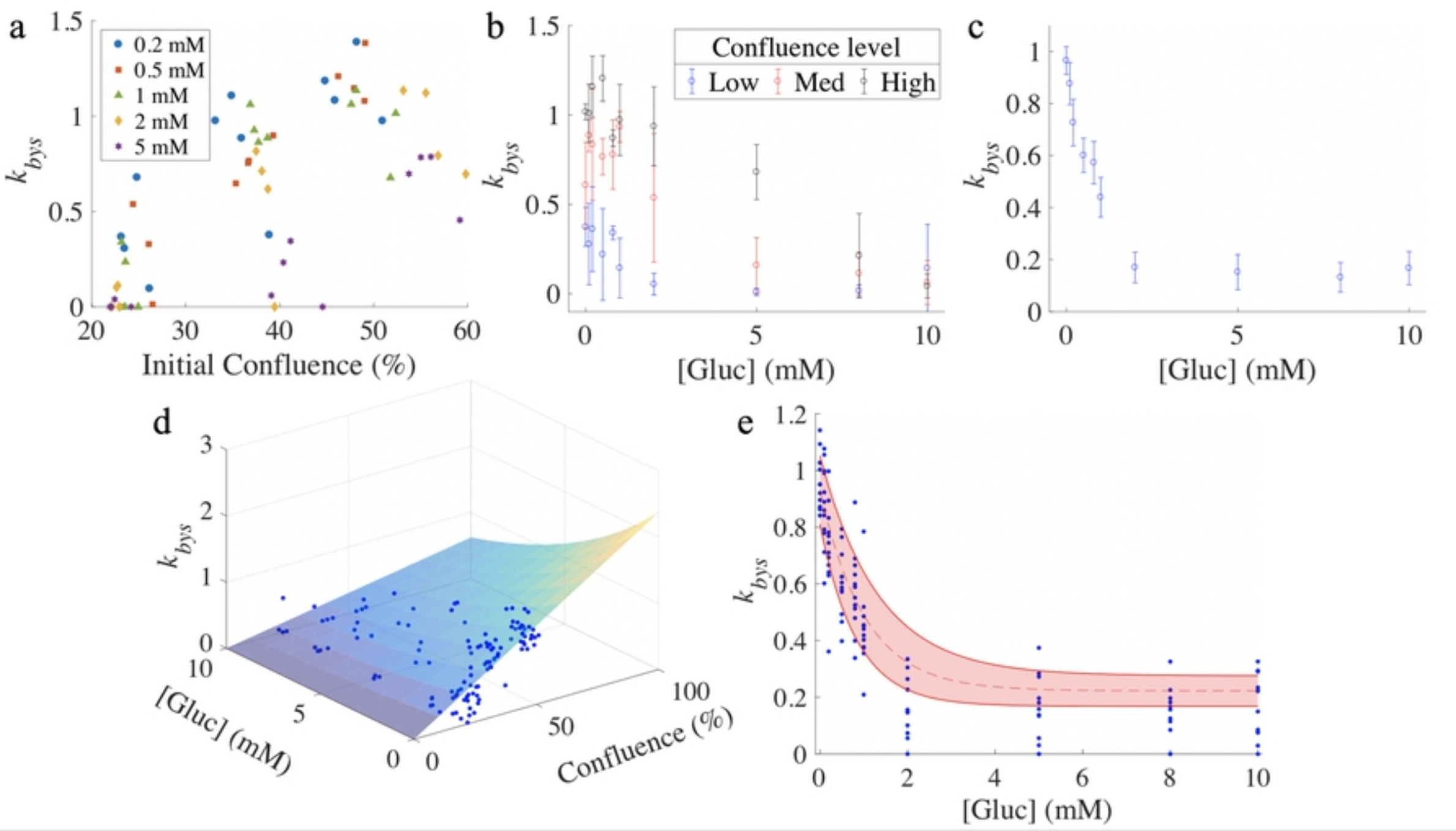
S1 Fig. Steps on cell segmentation for phase-contrast images of BT-474 breast cancer cell lines. The size of the whole well image is 2400 x 2400 pixels. Here we present a window of 400

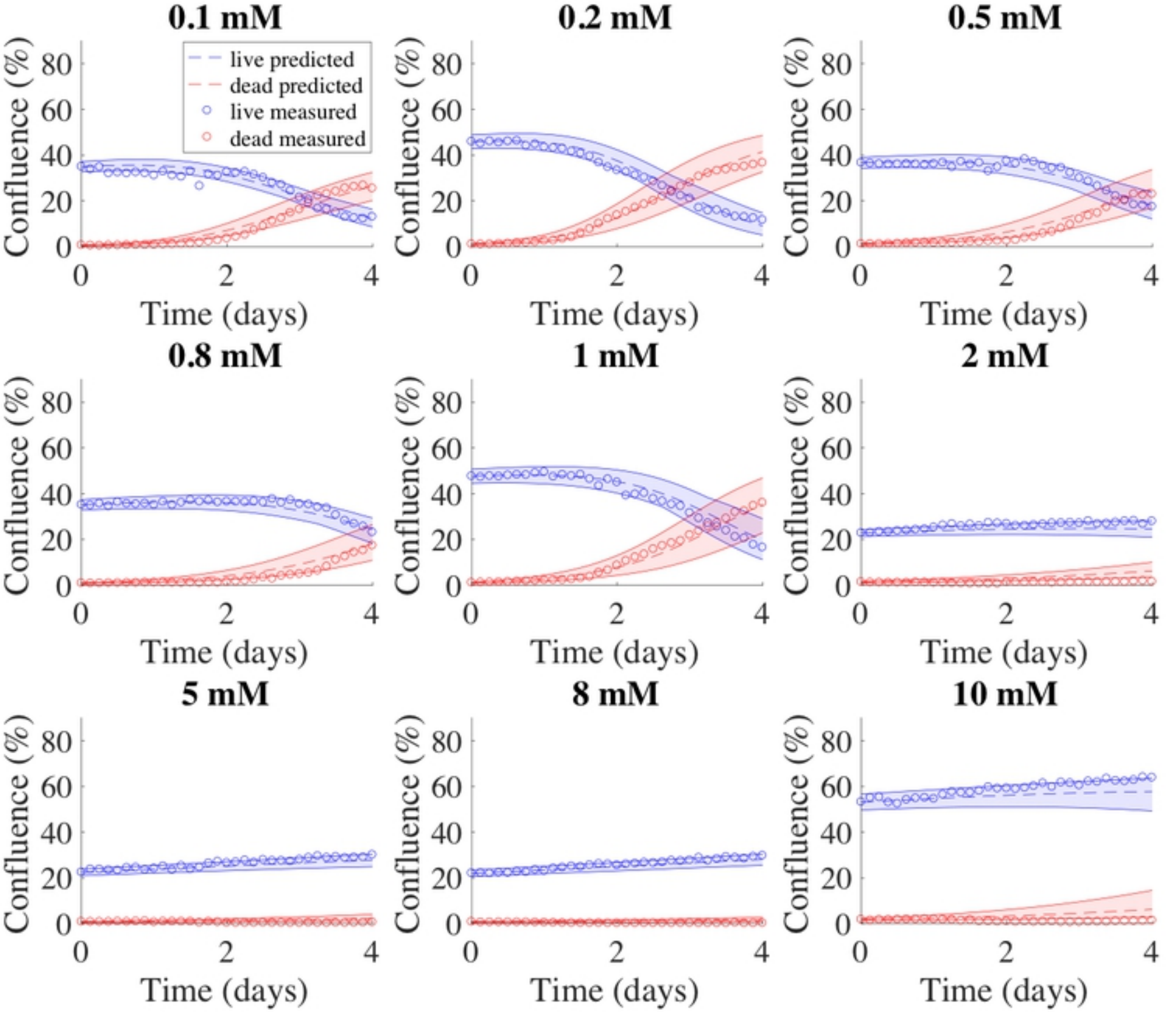
780 x 400 pixels from an example image. Panel A: raw image of BT-474 cells; Panel B: image post
781 'colfilt'; Panel C: image post the Gaussian filter; Panel D: image post 'im2bw'; Panel E: image
782 post 'imerode'; Panel F: image post 'imclose'; Panel G: image post 'imopen'; Panel H: image post
783 'bwareaopen', the final cell mask; Panel I: overlay of raw image and the cell mask for BT-474
784 cells.
785
786 **S2 Fig. Steps on cell segmentation for phase-contrast images of MDA-MB-231 breast cancer**
787 **cell lines.** The size of the whole well image is 2400 x 2400 pixels. Here we present a window of
788 400 x 400 pixels from an example image. Panel A: raw image of MDA-MB-231 cells; Panel B:
789 image post binarization; Panel C: image post 'imclose'; Panel D: image post 'bwareaopen', the
790 final cell mask; Panel E: overlay of raw images and the cell mask for MDA-MB-231 cells.
791
792 **S3 Fig. Time courses of tumor cell confluence in media with 0 mM glucose, grouped by initial**
793 **confluence.** Tumor cells may keep proliferating for some time even in a glucose free medium,
794 even up to 24 hours for MDA-MB-231 (Panel A). The proliferation in glucose free medium is not
795 observed for BT-474 (Panel B).
796
797 **S4 Fig. Estimates of the death rate due to the bystander effect, k_{bys}, as a function of different**
798 **initial confluence, for BT-474 cells.** Each subtitle indicates the initial glucose concentration. For
799 a given initial glucose level, k_{bys} increases with initial confluence. For 8 of 10 initial glucose levels
800 tested (0, 0.1, 0.2, 0.5, 0.8, 1, 2, and 5 mM, the bystander effect death rate is positively correlated
801 with initial confluence, with correlation coefficients all > 0.74 ($p < 0.01$). For the highest two initial

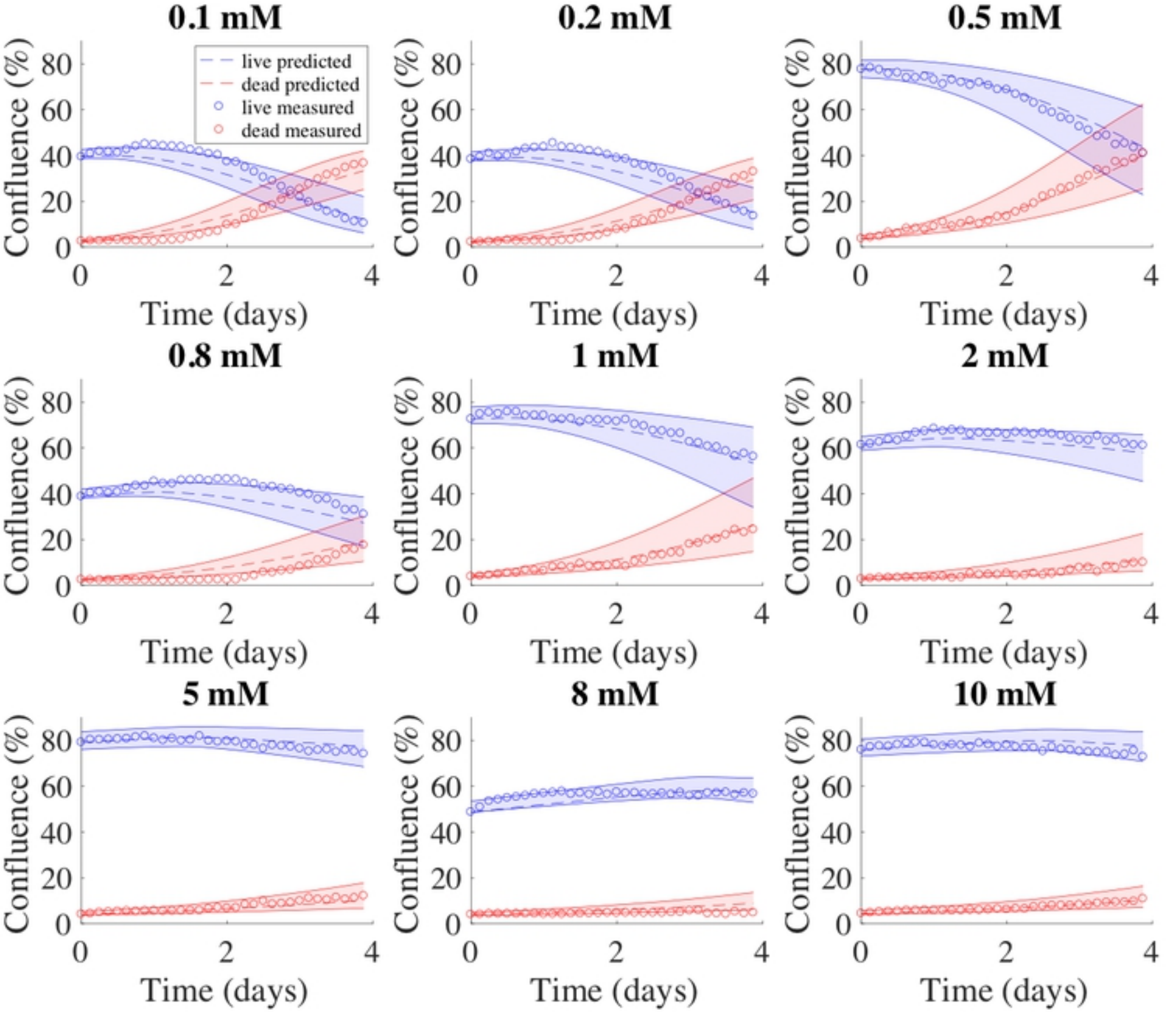
802 glucose levels (8 and 10 mM), there is no significant correlation between the bystander effect death
803 rate and the initial confluence ($p > 0.1$).
804
805 **S5 Fig. Estimates of the death rate due to the bystander effect, k_{bys}, as a function of different**
806 **initial confluence, for MDA-MB-231 cells.** Each subtitle indicates the initial glucose
807 concentration. There is no significant correlation between the bystander effect death rate and the
808 initial confluence ($p > 0.1$).
809
810 **S1 Table. Results of AICc value for model selection.** Model 02 is the complete model described
811 by Eq.[1]-[5]. In Model 01, all the terms involving k_{bys} is removed, while in Model 03, any term
812 involving k_d is removed. In the first calibration, the measured live and dead tumor cell time courses
813 are independently fit to the model to produce separate estimates for each model parameter. In the
814 second calibration, all the parameters are treated as global parameters. In the third calibration, k_{bys}
815 is considered as a local parameter while the other parameters (k_p, k_d, and v) are treated as global
816 parameters.
817
818 **S2 Table. Explicit matrices of variables used to evaluate the model's performance.** $X_{model,ij}$ is
819 the number of live or dead cells of well j at timepoint i calculated from the model, $X_{data,ij}$ is the
820 number of live or dead cells of well j at timepoint i from the measured data, t is the total number
821 of timepoints, w is the total number of wells, and t_{end} is the last timepoint at the end of experiment
822 (EoE).
823
824


a
b
c
d
Phase-Contrast
Total
Live
Dead
Confluence (%)
Time (days)
Time-resolved images
Segmentation
Generate the confluence time courses
Data fitting and model selection
Parameter analysis
Training and validation
Fluorescence
Dead
e
f
g


Low Initial Confluence
Intermediate Initial Confluence
High Initial Confluence
BT-474
MDA-MB-231
0.2 mM
0.5 mM
2 mM
5 mM
Confluence (%)
Time (days)
a
b
c
d
e
f


a
b
c
d
e
0.2 mM
0.5 mM
1 mM
2 mM
5 mM
Initial Confluence (%)
Confluence level
Low
Med
High
[Gluc] (mM)
Confluence (%)
k_{bys}


0.1 mM
0.2 mM
0.5 mM
0.8 mM
1 mM
2 mM
5 mM
8 mM
10 mM
live predicted
dead predicted
live measured
dead measured
Confluence (%)
Time (days)
0
20
40
60
80
0
2
4


0.1 mM
0.2 mM
0.5 mM
0.8 mM
1 mM
2 mM
5 mM
8 mM
10 mM
live predicted
dead predicted
live measured
dead measured
Confluence (%)
Time (days)
0
20
40
60
80
0
2
4