

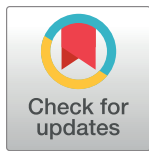
## RESEARCH ARTICLE

# Towards integration of time-resolved confocal microscopy of a 3D *in vitro* microfluidic platform with a hybrid multiscale model of tumor angiogenesis

Caleb M. Phillips<sup>1</sup>, Ernesto A. B. F. Lima<sup>1,2\*</sup>, Manasa Gadde<sup>1</sup>, Angela M. Jarrett<sup>1,7</sup>, Marissa Nichole Rylander<sup>1,3,4</sup>, Thomas E. Yankeelov<sup>1,3,5,6,7,8</sup>

**1** Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, Texas, United States of America, **2** Texas Advanced Computing Center, The University of Texas at Austin, Austin, Texas, United States of America, **3** Department of Biomedical Engineering, The University of Texas at Austin, Austin, Texas, United States of America, **4** Department of Mechanical Engineering, The University of Texas at Austin, Austin, Texas, United States of America, **5** Department of Diagnostic Medicine, The University of Texas at Austin, Austin, Texas, United States of America, **6** Department of Oncology, The University of Texas at Austin, Austin, Texas, United States of America, **7** Livestrong Cancer Institutes, The University of Texas at Austin, Austin, Texas, United States of America, **8** Department of Imaging Physics, The University of Texas at Austin, MD Anderson Cancer Center, Houston, Texas, United States of America

\* [lima@ices.utexas.edu](mailto:lima@ices.utexas.edu)



## OPEN ACCESS

**Citation:** Phillips CM, Lima EABF, Gadde M, Jarrett AM, Rylander MN, Yankeelov TE (2023) Towards integration of time-resolved confocal microscopy of a 3D *in vitro* microfluidic platform with a hybrid multiscale model of tumor angiogenesis. PLoS Comput Biol 19(1): e1009499. <https://doi.org/10.1371/journal.pcbi.1009499>

**Editor:** Philip K. Maini, Oxford, UNITED KINGDOM

**Received:** September 22, 2021

**Accepted:** December 13, 2022

**Published:** January 18, 2023

**Copyright:** © 2023 Phillips et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** We have made the data and code for analysis available for this project at '[https://github.com/CalebPhillips5/Vascular\\_calibration](https://github.com/CalebPhillips5/Vascular_calibration)'.

**Funding:** T.E.Y. acknowledges support from the National Cancer Institute via U01CA174706, R01CA186193, U01CA253540, and U24 CA226110. T.E.Y. also acknowledges support from the Cancer Prevention and Research Institute of Texas (CPRIT) via RR160005. T.E.Y. is a CPRIT Scholar in Cancer Research. M.N.R acknowledges

## Abstract

The goal of this study is to calibrate a multiscale model of tumor angiogenesis with time-resolved data to allow for systematic testing of mathematical predictions of vascular sprouting. The multi-scale model consists of an agent-based description of tumor and endothelial cell dynamics coupled to a continuum model of vascular endothelial growth factor concentration. First, we calibrate ordinary differential equation models to time-resolved protein concentration data to estimate the rates of secretion and consumption of vascular endothelial growth factor by endothelial and tumor cells, respectively. These parameters are then input into the multiscale tumor angiogenesis model, and the remaining model parameters are then calibrated to time resolved confocal microscopy images obtained within a 3D vascularized microfluidic platform. The microfluidic platform mimics a functional blood vessel with a surrounding collagen matrix seeded with inflammatory breast cancer cells, which induce tumor angiogenesis. Once the multi-scale model is fully parameterized, we forecast the spatiotemporal distribution of vascular sprouts at future time points and directly compare the predictions to experimentally measured data. We assess the ability of our model to globally recapitulate angiogenic vasculature density, resulting in an average relative calibration error of  $17.7\% \pm 6.3\%$  and an average prediction error of  $20.2\% \pm 4\%$  and  $21.7\% \pm 3.6\%$  using one and four calibrated parameters, respectively. We then assess the model's ability to predict local vessel morphology (individualized vessel structure as opposed to global vascular density), initialized with the first time point and calibrated with two intermediate time points. In this study, we have rigorously calibrated a mechanism-based, multiscale, mathematical model of angiogenic sprouting to multimodal experimental data to make specific, testable predictions.

support from the VPR Research and Creative Grant and the Walker Department of Mechanical Engineering at The University of Texas at Austin. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Author summary

The integration of experimentally obtained data to complex mathematical models, specifically in relation to tumor angiogenesis, is a rich field that is of great interest to the community. This integration, typically through calibrating model parameters that govern the mathematical model, is paramount to making specific, testable predictions of the system under investigation. In this work, we have used several forms of quantitative experimental data to inform a complex mathematical model to predict tumor angiogenesis. We do this through a multi-step calibration method by separating individual processes described in the model and utilizing data to inform key model parameters involved in that process. At the final step in our calibration method, we use longitudinal studies of tumor sprouting in our 3D microfluidic platform to fully parameterize our multiscale model of tumor angiogenesis. We utilize early time points to inform model parameters and then test our models ability to predict both local vascular features and vascular summary statistics at later time points. We have shown, with uncertainty, the prediction ability of our model and have shown we can reliably predict *in vitro* experiments of tumor-induced angiogenesis.

## 1 Introduction

There is now a mature literature on the mathematical modeling of tumor initiation, expansion, angiogenesis, and invasion [1–12]. These models (which may be discrete, continuous, or hybrid) seek to provide new strategies for understanding the underlying biology, and then use this knowledge to make predictions of the spatiotemporal evolution of the disease as well as its response to therapy [13–20]. A fundamental barrier limiting progress in the field, though, is the paucity of studies linking mechanism-based mathematical models with the appropriate experimental data [21–25]. One fundamental reason for this impasse is that most mathematical models require specific and quantitative data sets to calibrate model parameters; data types that are not frequently available. This is especially true in the field of tumor angiogenesis where the mathematical models demand data that are acquired at both high temporal and spatial resolution to resolve vascular dynamics [22, 26–30]. Furthermore, the models themselves may have a myriad of parameters yielding an ill-posed parameter calibration problem [31–34]. Still, these models, even if uncalibrated against experimental data, can serve as useful tools to guide the experimental study of tumor-induced angiogenesis. For example, Vilanova *et al.* developed a phase-field model of capillary progression with a discrete tip endothelial cell that guides capillaries based on tumor angiogenic factors and neighboring capillaries [35–37]. This model accounts for capillary regression due to anti-angiogenic therapies and qualitatively matches *in vivo* experimental measurements. Their approach was extended by Xu *et al.* who incorporated blood flow and initialized the resulting model with photo-acoustic imaging data characterizing the tumor and surrounding vasculature [38]. Travasso *et al.* developed a phase-field approach describing endothelial, stalk, and tip cells that compares differences in vessel morphology due to tip cell migration and stalk cell proliferation [39]. Their model's predictions were qualitatively compared to experimental data of vascular patterns in response to the production level of angiogenic factors. These efforts are to be commended as some of the first examples of linking mechanism-based, mathematical models of tumor angiogenesis with experimental data. The next step is to calibrate such a model with early time point data, and then make a prediction to directly compare model forecasts to data acquired experimentally at later time points.

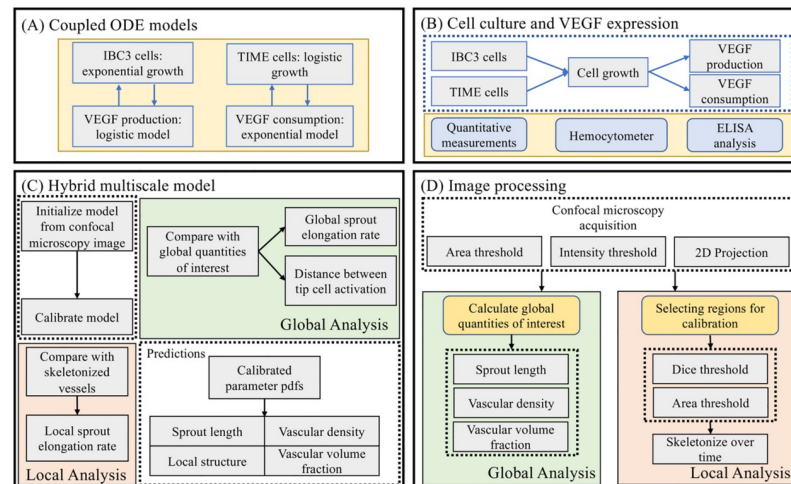
Calibrating mathematical models with experimental data allows for the ability to both predict future model outcomes and more rigorously test and understand the biological phenomena in question [23, 24, 40, 41]. Specifically, for tumor angiogenesis, the ability to calibrate parameters that dictate the development of tumor vasculature in a well-controlled, experimental platform would allow the systematic testing of various models of angiogenesis [21, 22]. However, obtaining the appropriate data for this task remains a difficult challenge. A substantial focus of experimental angiogenesis research is performed *in vitro* on 2D monolayers as well as *in vivo* xenograft models in animals. However, 2D systems do not mimic the complex behaviors observed *in vivo*, while xenograft systems pose substantial barriers to acquiring the data with adequate spatial and temporal resolution to resolve tumor induced angiogenesis [22, 42]. We do note other *in vitro* works in experimental angiogenesis [43–47]. In the present effort, we address the data demands *via* a 3D *in vitro* microfluidic platform [48]. This system has the advantage of emulating the complex architecture of the tumor microenvironment observed *in vivo*, as well as capturing the spatiotemporal behavior of the cells. The 3D *in vitro* microfluidic platform we utilized in this study consists of a collagen matrix seeded with inflammatory breast cancer (IBC) cells around a functional blood vessel to represent the IBC tumor-vascular interface. Over time, the IBC cells secrete pro-angiogenic factors that diffuse and bind to the functional blood vessel, initiating angiogenic sprouting. Spatiotemporal measurements of cell proliferation and apoptosis, branch lengths and number, anastomosis, and lumen formation in the platform will provide key parameters that drive the mathematical model.

In this contribution, we aim to calibrate an agent-based model of tumor angiogenesis [49] with experimental data obtained from the *in vitro* vascularized tumor platform [48] described above to predict vasculature sprouting. First, we perform experiments that isolate key components in the process of angiogenesis and use these data to calibrate important parameters in mathematical models of increasing complexity. This sequential approach allows a systematic, stepwise, approach for linking mathematical models with experimental data across spatial scales. The steps (referred to as scenarios in the remainder of this work) we follow are: 1) calibration of tumor and endothelial cell number to time-resolved hemocytometry, 2) calibration of the secretion and consumption of vascular endothelial growth factor (VEGF) by tumor and endothelial cells, respectively, to time-resolved protein concentration measurements, 3) calibration of stalk cell growth rate against time-resolved measurements of angiogenic sprout length, performing a 4) global and 5) local spatiotemporal calibration of the hybrid model of tumor angiogenesis to confocal microscopy data acquired from the *in vitro* microfluidic platform, and 6) testing the predictive capabilities of the calibrated mathematical model at the global and local scales.

## 2 Methods

### 2.1 Overview of methods

Fig 1 summarizes the computational and experimental methodologies utilized in this study. In Fig 1A, we show the coupled ordinary differential equation (ODE) models to describe tumor (IBC3) and endothelial (TIME) cell number, as well as the production of VEGF by tumor cells and the consumption of VEGF by endothelial cells. These models are calibrated using cell number over time (Scenario 1) counted *via* a hemocytometer, and VEGF concentration (Scenario 2) measured *via* ELISA analysis, shown in Fig 1B. Fig 1C depicts the hybrid multiscale model, initialized from confocal microscopy images observed in our 3D *in vitro* microfluidic platform, shown in Fig 1D, and calibrated by subsequent time points. In our global analysis, shown in green in Fig 1C and 1D, we calculate global quantities of interest from the confocal microscopy images, after thresholding for area and intensity and projecting the data into a 2D



**Fig 1. Schematic of computational and experimental methods.** Fig 1A and 1B show the calibration of tumor and endothelial cell number and VEGF concentration to hemocytometer and ELISA measurements over time. Given the calibrated VEGF production and consumption rates, we inform our hybrid multiscale model, shown to Fig 1C, to confocal microscopy images of angiogenic sprouts, depicted in Fig 1D. We analyze the model globally, calibrating model parameters to summary statistics of the data, namely the sprout length and vascular density, and locally, calibrating the local stalk cell divide time to segmented vascular structures. This sequential approach, starting with protein concentration and cell number experiments to inform the hybrid model prior to integrating confocal microscopy images, allows us to utilize experimental data at multiple scales to inform the multiscale nature of the tumor angiogenesis model.

<https://doi.org/10.1371/journal.pcbi.1009499.g001>

plane. From this 2D image, we calculate the sprout length, vascular density, and vascular volume fraction, which are readily comparable to the angiogenesis model to calibrate and predict the stalk cell divide time (Scenario 3) and the distance between tip cells for activation (Scenario 4). After analyzing the model globally, we assess the ability of the model to recapitulate local vascular features. In our local analysis, shown in peach in Fig 1C and 1D, we select individual sprouts for calibration through a Dice and area threshold over time, and skeletonize the observed vessels (and observe adjacent tumor cells). These skeletonizations, along with the tumor cells, are used to initialize and calibrate the local stalk cell divide time. We then assess the ability of the angiogenesis model to locally predict the final time point (Scenario 5). Finally, we use the calibrated parameter distributions to predict sprout length, vascular density, and vascular volume fraction (globally) and the local structure of the angiogenic sprouts.

## 2.2 Experimental methods

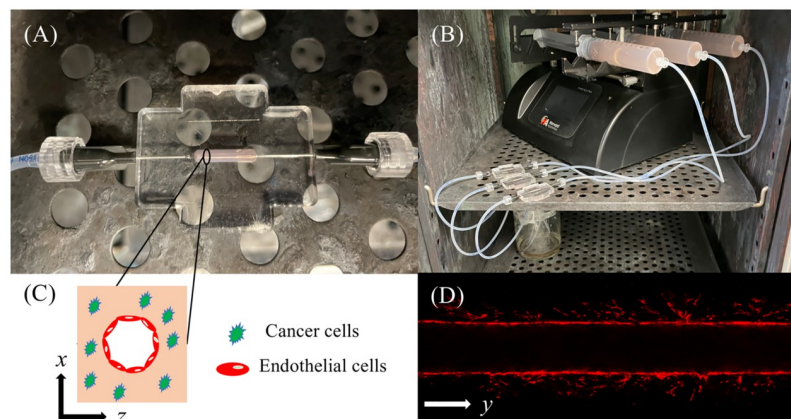
**2.2.1 Cell culture and VEGF concentration.** Human inflammatory breast cancer (IBC) MDA-IBC3 cells and telomerase-immortalized human microvascular endothelial (TIME) cells were used in this study. Green fluorescence protein labelled MDA-IBC3 (a HER2+ IBC cell line) were kindly provided by Dr. Wendy Woodward (MD Anderson Cancer Center, Houston, TX) and mKate labeled TIME cells were a generous gift from Dr. Shay Soker (Wake Forest Institute for Regenerative Medicine, Winston-Salem, NC). MDA-IBC3 cells were cultured in Ham's F-12 media supplemented with 10% fetal bovine serum, 1% antibiotic-antimycotic, 1  $\mu\text{g}/\text{ml}$  hydrocortisone, and 5  $\mu\text{g}/\text{ml}$  insulin. TIME cells were cultured in Endothelial Cell Growth Medium-2 BulletKit (EGM-2, Lonza). All cell cultures utilized in this study were maintained at 5% CO<sub>2</sub> atmosphere and 37°C.

To determine the production of VEGF by MDA-IBC3 cells and consumption of VEGF by TIME cells, cells were seeded with a density of 30,000 cells/cm<sup>2</sup> and 10,000 cells/cm<sup>2</sup>,

respectively, in separate 12 well plates supplemented with EGM-2 media. Media was collected and replaced from the same 4 wells daily for measurement of VEGF concentration and cell growth was imaged with an IncuCyte Zoom (Essen Bioscience, Ann Arbor, MI) every 8 hours for 7 days. VEGF concentration was measured using a Human VEGF Quantikine ELISA Kit (R&D Systems, Minneapolis, MN). The cell number in each well was determined by trypsinization and counting of cells with a hemocytometer on days 1, 2, 3, 5, and 7. For each timepoint, four wells were sacrificed for cell counting and these were separate from the media collection wells for VEGF measurements.

**2.2.2 Vascularized 3D *in vitro* microfluidic platforms.** The *in vitro* 3D tumor microfluidic platforms utilized in this study were composed of extracellular matrix comprised of collagen type I seeded with green fluorescence protein labeled MDA-IBC3 around a hollow vessel lined with mKate labeled TIME cells housed in a polydimethylsiloxane scaffold as described in our published work. 7 mg/ml collagen solution, which has a similar stiffness to that of *in vivo* breast tumors, was used in fabricating extracellular matrix of the platforms [50–52]. MDA-IBC3 cells were seeded at a density of  $1 \times 10^6$  cells/ml in the collagen solution and polymerized around a 22 G needle at 37°C for 25 minutes. After polymerization, the needle was removed, and the resulting hollow vessel was filled with a solution of  $2 \times 10^5$  TIME cells to form an endothelialized vessel lumen. Flow was introduced using a syringe pump system and a graded flow protocol was used to establish a confluent endothelium as we have previously published [48, 51, 53–55]. The microfluidic platform is summarized in Fig 2. Fig 2A shows an example microfluidic platform with an axial cross-section depicting the functional parent blood vessel and cancer cells seeded throughout the microenvironment in Fig 2C. The platforms are connected to a syringe pump system, shown in Fig 2B, that allows for the continuous perfusion of media into the collagen microenvironment. Over time, the cancer cells release pro-angiogenic proteins, which diffuse through the microenvironment and cause tumor-induced angiogenesis. An example of these sprouts is shown in Fig 2D.

**2.2.3 Image acquisition and processing.** Platforms were imaged with a Leica TCS SP8 Confocal Microscope (Leica Microsystems, Germany) on days 3, 5, 7, 9, 11, 15, and 19. The 3D volume of angiogenic sprouts was imaged with z-slices acquired at intervals of 4.28 microns



**Fig 2. Schematic of microfluidic platform.** Fig 2A depicts the housing of the 3D microfluidic platform. The black oval denotes the boundary of the platform containing collagen matrix (shown in light pink) within the housing chamber. The chamber is connected to a syringe pump flow system, shown in Fig 2B, to maintain continuous flow of media through the parent vessel, within the collagen matrix. The axial cross-section of the vessel (the x-z plane) is shown in Fig 2C and an example z-slice image (i.e., longitudinal cross-section along the length of the vessel (the x-y plane)), is shown in Fig 2D.

<https://doi.org/10.1371/journal.pcbi.1009499.g002>



and an in-plane resolution of 2.254 microns  $\times$  2.254 microns in the  $x$ - $y$  plane (shown in Fig 2D). The confocal microscopy images were processed at each time point to ensure the same local structures of the microfluidic platform could be compared over the duration of the experiment. This was done *via* a reference mark. It was assumed that the diameter of the parent vessel did not deviate significantly over the course of the time points used for calibration. The parent vessel periphery was used as a reference to align the segmented region of sprouting in  $x$ - $y$  space, depicted in Fig 2D. We assume that the vessels grow primarily in  $x$ - $y$  space and that little growth is in the  $z$  direction (we comment on this assumption in the Discussion section). Vessel lengths were estimated at multiple time points using the built-in Leica measurement. We average signal intensity across 10 slices in the  $z$  direction (approximately 43 microns in total) and thresholded using area and intensity (to filter out migratory endothelial cells that are not a part of maturing vessels) to create a binarized vessel mask. This mask is used to calculate both the global and local vascular quantities for calibration. The global quantities of interest are the vascular density, calculated by summing the vascular mask along the length of the vessel, and the vascular fraction, calculated by dividing the total vascular mask area by the total area in the experimental computational domain. For local calibrations, we calculate the local vessel morphology (i.e., the structure of specific vasculature) over time.

To select specific vessels for local calibration, we developed an algorithm that calculates the Dice and percentage of overlap across time points. This allows for selecting regions that, when aligned, correspond to angiogenic sprouts growing over time and minimizes the effects of random endothelial cell migration. For the selected regions, we utilized a Zhang-Suen thinning algorithm [56] (available at <https://github.com/linbojin/Skeletonization-by-Zhang-Suen-Thinning-Algorithm>) for skeletonization after dilating the vessels using the 'strel' function in MATLAB (Mathworks, Natick, MA). Given the skeletonization for the model and the data at time  $t$ , we calculate the average distance between the model to the data centerline and from the data to the model centerline to use as the objective function for model calibration by using Python's SciPy Euclidean distance transform [57].

## 2.3 Computational methods

**2.3.1 Overview of computational methods.** We calibrate three mathematical models for the purpose of predicting vascular response to the secretion of VEGF by tumor cells. (We first describe them qualitatively before introducing their mathematical representations.) The first and second are systems of ordinary differential equations that describe VEGF consumption by TIME cells and production of VEGF by MDA-IBC3 cells. These models describe the concentration of VEGF over time and the evolving cell type associated in production or consumption. The first model describes VEGF consumption of endothelial cells and has three model parameters: the consumption rate of VEGF by endothelial cells, and the growth rate and carrying capacity of endothelial cells. All of these parameters may be determined from the VEGF concentration and cell culture data. The second model describes the growth of MDA-IBC3 cells and their production of VEGF through three parameters: the VEGF production rate and carrying capacity, and the growth rate of the tumor cells. These parameters may also be determined from the VEGF concentration and cell culture data. We utilize the consumption rate of endothelial cells and the production rate of endothelial cells as inputs into the third model, a hybrid multiscale model of tumor angiogenesis. An agent-based model describes the tumor and endothelial cell dynamics and a continuum model governs the dispersion of VEGF. This hybrid model computes the spatiotemporal distribution of vascular sprouts and is readily comparable to the confocal microscopy images (and the calculated quantities of interest at the local and global scales) of the 3D *in vitro* microfluidic platform. These three models vary in complexity

and allow us to isolate model parameters in a series of well-controlled experiments, and ultimately to test the predictive capabilities of the multiscale model at the global and local scales.

**2.3.2 Model of VEGF production and consumption.** The following two systems of mathematical models describe the consumption and production of VEGF by TIME and MDA-IBC3 cells, respectively. The first system of ODEs is:

$$\frac{dN_e}{dt} = \lambda_e^s N_e(t) \left(1 - \frac{N_e(t)}{\theta_e}\right), \quad (1)$$

$$\frac{d[VEGF]}{dt} = -\lambda_e^c N_e(t) [VEGF](t), \quad (2)$$

where  $\lambda_e^s$  is the growth rate of endothelial cells,  $N_e(t)$  is the number of endothelial cells at time  $t$ ,  $\theta_e$  is the carrying capacity of endothelial cells,  $[VEGF]$  is the concentration of VEGF, and  $\lambda_e^c$  is the consumption rate of endothelial cells. This system assumes that each endothelial cell consumes VEGF at a constant rate, VEGF is not produced by endothelial cells,  $N_e(t)$  follows logistic growth with carrying capacity  $\theta_e$ , and VEGF decay is negligible. To model the VEGF secretion by tumor cells we write:

$$\frac{dN_T}{dt} = \lambda_T^s N_T(t), \quad (3)$$

$$\frac{d[VEGF]}{dt} = \lambda_T^p N_T(t) \left(1 - \frac{[VEGF](t)}{\theta_v}\right), \quad (4)$$

where  $\lambda_T^s$  is the growth rate of tumor cells,  $N_T(t)$  is the number of tumor cells,  $\lambda_T^p$  is the production rate of VEGF by tumor cells, and  $\theta_v$  is the carrying capacity of VEGF. Since both  $N_e(t)$  and  $N_T(t)$  are measured daily, we calibrate  $\lambda_e^s$  and  $\lambda_T^s$  and utilize the model to predict cell numbers between experimental measurements. To directly compare model predictions to the cell culture and protein concentration data, we take  $N_T(t)$  and  $N_e(t)$  to be MDA-IBC3 cells and the generic endothelial cells to be TIME cells, respectively.

**2.3.3 Hybrid multiscale model of tumor angiogenesis.** We utilize the hybrid multiscale model of tumor induced angiogenesis developed in [49] to compare the global and local quantities of interest with the confocal microscopy data in [48]. We now summarize the salient features of this model and refer the interested reader to [49]. The multiscale model describes tumor and endothelial cell dynamics using an agent-based model, and the evolution of nutrients and vascular endothelial growth factor (VEGF) using a continuous partial differential equation (PDE) model. Tumor cells act as individual agents and may reside as a proliferative, quiescent, hypoxic, or necrotic cell. They transition from quiescent to proliferative stochastically with a linear dependence on nutrient concentration. Tumor cells uptake nutrients, grow, and proliferate until the supply of nutrients is depleted (i.e., less than a threshold,  $\sigma_H$ ), at which point they become hypoxic and secrete VEGF. The VEGF diffuses across the domain and eventually contacts nearby vasculature. The blood vessels, comprised of endothelial cells modeled as agents, uptake VEGF and become activated when the concentration of VEGF is greater than a threshold,  $\alpha_A$ , and the distance between other tip cells is greater than  $d_{tip}$ . The activated endothelial cells then transition to tip cells, which are responsible for the migration of angiogenic sprouts up the concentration gradient of VEGF due to a chemotactic force,

$$\mathbf{F}_{VEGF}^{ij} = -f_v \nabla \psi(\mathbf{d}^{ij}; R_N^i + R_N^j, R^i + R^j, [VEGF]), \quad (5)$$

where  $f_v$  is the VEGF force coefficient,  $d^{ij}$  the distance between tip cell  $i$  and the neighboring

endothelial cell  $j$ ,  $R_N^z$  is the nuclear radius of cell  $\alpha$ ,  $R^\alpha$  is the cytoplasmic radius of cell  $\alpha$ , and the VEGF potential function  $\nabla\psi$  is given as

$$\nabla\psi(\mathbf{d}; R_N, R, [\text{VEGF}]) = \begin{cases} -\left(\frac{R_N|\mathbf{d}|}{R^2} - \frac{2|\mathbf{d}|}{R} + 1\right) \frac{\nabla[\text{VEGF}]}{|\nabla[\text{VEGF}]||\mathbf{d}|}, & 0 \leq |\mathbf{d}| \leq R_N, \\ -\left(\frac{|\mathbf{d}|^2}{R^2} - \frac{2|\mathbf{d}|}{R} + 1\right) \frac{\nabla[\text{VEGF}]}{|\nabla[\text{VEGF}]||\mathbf{d}|}, & R_N \leq |\mathbf{d}| \leq R, \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

The proximal endothelial cells, termed stalk cells, proliferate after time  $d_{sc}$  and allow the elongation of the sprouts. The effective sprout elongation rate can be calculated by summing the stalk cell divide time,  $d_{sc}$ , and the stalk cell growth time,  $g_{sc}$ , divided by the diameter of the stalk cell,  $2R^{sc}$ . Mechanical adhesion and repulsion forces maintain the structural integrity of the vessels and allow the formation of the lumen. Importantly, this formulation of VEGF force properly scales the chemotactic forces acting on the tip cell to maintain vessel structural integrity. Once the sprouts mature, they form complex networks and establish blood flow, enabling the delivery of new nutrients to the tumor. The VEGF and nutrient dynamics are governed by the following reaction-diffusion equations:

$$\left. \begin{aligned} \frac{\partial\sigma}{\partial t} &= \nabla \cdot (D_n \nabla\sigma) - \Lambda_n(\mathbf{x}, t)\sigma + \Gamma_n(\mathbf{x}, t)\sigma(1 - \sigma), \\ \frac{\partial[\text{VEGF}]}{\partial t} &= \nabla \cdot (D_v \nabla[\text{VEGF}]) - \Lambda_v(\mathbf{x}, t)[\text{VEGF}] \\ &\quad + \Gamma_v(\mathbf{x}, t)[\text{VEGF}](1 - [\text{VEGF}]), \end{aligned} \right\} \text{ in } \Omega \times (0, T_{\text{tissue}}), \quad (7)$$

where  $D_n$  and  $D_v$  are the nutrient and VEGF diffusion coefficients respectively,  $\Lambda_n$  is the nutrient uptake rate of tumor cells,  $\Lambda_v$  is the VEGF consumption rate of endothelial cells,  $\Gamma_n$  is the nutrient production rate from endothelial cells that are part of anastomotic loops,  $\Gamma_v$  is the VEGF production rate of hypoxic tumor cells. These source and sink terms, which couple the cellular agent-based model to the continuous PDE model at the tissue scale, are defined as follows:

$$\left. \begin{aligned} \Lambda_n(\mathbf{x}, t) &= \lambda_{pq}^c \phi_{pq}(\mathbf{x}, t) + \lambda_h^c \phi_h(\mathbf{x}, t), \\ \Gamma_n(\mathbf{x}, t) &= \gamma_e \phi_e(\mathbf{x}, t), \\ \Lambda_v(\mathbf{x}, t) &= \lambda_i^c \phi_i(\mathbf{x}, t) + \lambda_s^c \phi_s(\mathbf{x}, t) + \lambda_p^c \phi_p(\mathbf{x}, t), \\ \Gamma_v(\mathbf{x}, t) &= \gamma_h \phi_h(\mathbf{x}, t), \end{aligned} \right\} \quad (8)$$

where the subscripts pq, h, l, t, s, and e specify proliferative and quiescent tumor cells, hypoxic tumor cells, looped endothelial cells, tip endothelial cells, stalk endothelial cells, and endothelial cells, respectively. Also,  $\phi_\alpha(\mathbf{x}, t)$  is the volume fraction of cell type  $\alpha \in \{pq, h, e, t, s, p\}$ , at position  $\mathbf{x}$  and time  $t$ ,  $\gamma_\alpha^c$  is the consumption rate of the cell type  $\alpha$ , and  $\gamma_\alpha$  is the production rate of the cell type  $\alpha$ . The volume fraction of each cell type is calculated by computing the total area of each cell type within each element mesh.

In the analysis that follows, we assume that all tumor cells produce VEGF equal to the value calibrated from the cell culture and VEGF concentration computational-experimental scenario



(see sections 2.4.1 and 2.4.2) but stop producing VEGF whenever they are within the threshold distance from the sprouting vessel. Thus, we are using this distance as a surrogate for nutrient delivery and can perform our computational-experimental studies independent of nutrient/hypoxia status. For the calibration of the stalk cell divide time in the global analysis, we use a simplified model to calculate the length of the angiogenic sprout to compare with the experimental data and assume the growth of the sprout is independent of VEGF concentration. This simplified model is identical to the full model but with a significantly reduced domain and only one tip cell that grows perpendicular to the parent vessel, allowing removal of the VEGF field, thereby dramatically reducing computational cost while maintaining the angiogenic growth dynamics.

**2.3.4 Bayesian calibration.** Bayesian inference provides a statistical framework for calibrating model parameters by accounting for both uncertainties in the mathematical model and the experimental data. Thus, it is an excellent methodology for determining uncertainties in model predictions of experimental outcomes [58–61]. We now summarize the salient features of Bayesian parameter calibration.

Beginning with Bayes' Rule,

$$\pi_{\text{posterior}}(m|y) = \frac{\pi_{\text{prior}}(m)\pi_{\text{likelihood}}(y|m)}{\pi_{\text{evidence}}(y)}, \quad (9)$$

where  $m$  denotes the model parameters,  $y$  is the experimental data,  $\pi_{\text{posterior}}(m|y)$  is the distribution of model parameters given the experimental data,  $\pi_{\text{prior}}(m)$  is the prior distribution of the model parameters,  $\pi_{\text{likelihood}}(y|m)$  is the conditional probability of the experimental data given the model parameters, and  $\pi_{\text{evidence}}(y)$  is a normalizing factor. These probability density functions (PDFs) provide a holistic description of the system (and therefore) the parameter uncertainty. Since  $\pi_{\text{evidence}}(y)$  is a normalizing factor, we can rewrite the posterior distribution as,

$$\pi_{\text{posterior}}(m|y) \propto \pi_{\text{prior}}(m)\pi_{\text{likelihood}}(y|m). \quad (10)$$

The prior information of the parameters,  $\pi_{\text{prior}}(m)$ , is chosen on a problem specific basis. If only the bounds  $a$  and  $b$  of the parameter are known, then the prior is selected as a uniform distribution,  $\pi_{\text{prior}}(m_i) \sim U(a, b)$ . Since only the bounds are known in this work, we exclusively use uniform priors. This means that  $\pi_{\text{prior}}(m_i)$  is constant and we can further simplify Bayes' Rule to

$$\pi_{\text{posterior}}(m|y) \propto \pi_{\text{likelihood}}(y|m). \quad (11)$$

Assuming independent and identically distributed data,

$$\pi_{\text{likelihood}}(y|m) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^N (y_{\text{data}}^i - y_{\text{model}}^i(m))^2\right), \quad (12)$$

where  $\sigma$  is the standard deviation,  $N$  is the total (combination of spatial and temporal) number of data points,  $y_{\text{data}}^i$  is the experimental data, and  $y_{\text{model}}^i$  is the model prediction given parameters  $m$ . This can also be written as,

$$\pi_{\text{likelihood}}(y|m) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \phi(m)\right), \quad (13)$$

where  $\phi(m)$  is the misfit to the data. With a given data misfit, Eq (13) is straightforward to calculate. The hyperparameter,  $\sigma$ , characterizes the uncertainty of the experimental data and the

inadequacy of the model [62, 63]. The posterior distribution for the parameters can be learned by utilizing sampling or quadrature methods to compute  $\pi_{\text{likelihood}}(y|m)$  at specific values of the parameters. Therefore, this framework allows us to use experimental data to learn the updated distribution of parameters. These parameter distributions can then be used to propagate uncertainties in the model predictions for a robust description of prediction uncertainty. We also specifically note that the maximum likelihood parameter values,  $m^*$ , can be obtained from  $\pi_{\text{likelihood}}(y|m)$  by taking the maximum of  $\pi_{\text{likelihood}}(y|m)$  with respect to  $m$ . This can be particularly useful in multilevel calibration schemes, as some parameters from one scenario are useful in later scenarios. For example, we calibrate the growth rate of endothelial cells,  $\lambda_e^g$  in Eq (1), to use as an input for solving VEGF concentration in Eq (2) where we calibrate  $\lambda_e^c$  without considering the uncertainty in  $\lambda_e^g$  (the consumption of VEGF by endothelial cells), which will be used in subsequent scenarios).

## 2.4 Model calibration and prediction scenarios

We now define the parameter estimation problems in the ODE models (i.e., Eqs (1)–(4)) and the hybrid multiscale model (i.e., Eqs (5)–(8)); the parameters to be calibrated are listed in Table 1. The calibration scenarios results and analyzes are provided at [https://github.com/CalebPhillips5/Vascular\\_calibration](https://github.com/CalebPhillips5/Vascular_calibration).

**2.4.1 Scenario 1: Calibration of tumor and endothelial cell growth.** First, we utilize the hemocytometry data to calibrate the parameters from Eqs (1) and (3). We assume tumor and endothelial cells growth is independent of the VEGF concentration and can therefore separate the coupled systems into two separate calibrations. The parameters calibrated in this scenario are the growth rate of tumor and endothelial cells ( $\lambda_T^g$  and  $\lambda_e^g$ , respectively), and the carrying capacity of endothelial cells ( $\theta_e$ ). We then define the parameter estimation problem for parameters  $m$  in both models as: Find  $m = [\lambda_T^g, \lambda_e^g, \theta_e]$  such that

$$\phi(m) = \sum_{i=1}^{N_{\text{days}}} (N_{\text{data}}(t_i) - N_{\text{model}}(t_i, m))^2, \quad (14)$$

is minimized.  $N_{\text{data}}(t_i)$  and  $N_{\text{model}}(t_i, m)$ , denote the measured and model prediction of tumor or endothelial cells respectively, and  $N_{\text{days}}$  is the number of measurements. We use Eqs (1) and (3) with calibrated parameters as input functions required to infer the VEGF production and consumption rates for tumor and endothelial cells, respectively, in Eqs (2) and (4). We compare the relative error in tumor and endothelial cell number to show the model goodness of fit.

**Table 1. Model parameters to be calibrated.**

Parameter	Interpretation	Source
$\lambda_e^g$	Growth rate of TIME endothelial cells	Scenario 1
$\theta_e$	Carrying capacity of TIME endothelial cells	Scenario 1
$\lambda_T^g$	Growth rate of IBC3 tumor cells	Scenario 1
$\lambda_T^p$	VEGF production rate by IBC3 tumor cells	Scenario 2
$\lambda_e^c$	VEGF consumption rate by TIME endothelial cells	Scenario 2
$\theta_v$	Carrying capacity of VEGF	Scenario 2
$d_{sc}$	Angiogenic stalk cell divide time	Scenario 3
$f_v$	VEGF force coefficient	Scenario 3
$d_{tip}$	Distance between new tip cells	Scenario 4
$d_{sc}^*$	Local stalk cell divide time rate	Scenario 5

<https://doi.org/10.1371/journal.pcbi.1009499.t001>

**2.4.2 Scenario 2: Calibration and prediction of VEGF concentration.** For the VEGF calibration problem, we utilize the ELISA data to calibrate the production and consumption rates of VEGF, as well as the carrying capacity of VEGF in the dish (Eqs (2) and (4)). We then define the parameter estimation problem in both models as: Find  $m = [\lambda_T^p, \lambda_e^c, \theta_v]$  such that

$$\phi(m) = \sum_{i=1}^{N_{days}} (V_{data}(t_i) - V_{model}(t_i, m))^2, \quad (15)$$

is minimized, where  $V_{data}(t_i)$  and  $V_{model}(t_i, m)$ , denote the measured and predicted concentration of VEGF, respectively. In this scenario, we employ a naive Bayesian quadrature (i.e., selecting quadrature points uniformly across the domain) approach to minimize  $\phi(m)$  with respect to the parameters  $m$ . Since the VEGF production and consumption rates are assumed to transcend the experimental scenario (as characteristics of the cells as opposed to characteristics of the environment), we will sample from these distributions moving forward in both the prediction of the VEGF concentration as well as the prediction of the vascular ABM (section 2.4.4). To easily generate samples, we approximate the PDFs as Gaussian distributions. We calculate the relative error in VEGF concentration for tumor and endothelial cells and show the propagated uncertainty through time.

**2.4.3 Scenario 3: Calibration and prediction of global vessel length.** We utilize confocal microscopy images to measure the length of angiogenic sprouts over time. This provides a 1D length measurement that can be used to calibrate parameters that govern the growth and development of the sprouts. We selected vessels that can be observed on day 3 and subsequently tracked during the remainder of the experiment to ensure consistent (i.e., the growth of the same sprouts) sprout growth. This is done via a reference mark on the microfluidic platform to align the slices across time. The key parameters are the VEGF force coefficient,  $f_v$ , and the stalk cell divide time,  $d_{sc}$ . We then define the parameter estimation problem as: Find  $m = [f_v, d_{sc}]$  such that

$$\phi(m) = \sum_{i=1}^{N_{days}} \sum_{j=1}^{N_{sprouts}} (L_{data}(t_i, s_j) - L_{model}(t_i, m))^2, \quad (16)$$

is minimized, where  $L_{data}(t_i, s_j)$  and  $L_{model}(t_i, m)$  are the observed and predicted length of sprout  $s_j$  at time point  $t_i$  with parameters  $m$  at time point  $t_i$ . The number of sprouts,  $N_{sprouts}$ , was chosen such that the same sprouts were tracked throughout the duration of the experiments. The calibration is solved using a Metropolis-Hastings algorithm available in the QUESO (Quantification of Uncertainty for Estimation, Simulation, and Optimization) library [64]. The prediction of  $L_{data}(t_i, s_j)$  is solved by running the forward model with parameters selected from specific samples used in the Metropolis-Hastings algorithm for calibration and calculating the mean and standard deviation of the resulting quantities of interest,  $L_{model}(t_i, m)$ . We report the relative error between the observed and predicted vessel length over time.

**2.4.4 Scenario 4: Calibration and prediction of vascular density.** To conclude our global analysis, we calculate the vascular density moving away from the parent vessel across time. This provides a global description of the total vascular sprouting to infer the probability distribution function of the distance between new tip cells. We then define the parameter estimation problem as: Find  $m = [d_{tip}]$  such that

$$\phi(m) = \sum_{i=1}^{N_{days}-1} \sum_{j=1}^{N_x} (D_{data}(t_i, s_j) - D_{model}(t_i, m))^2, \quad (17)$$

is minimized, where  $D_{data}(t_i, s_j)$  and  $D_{model}(t_i, m)$  are the experimentally estimated (from the confocal microscopy data) and model predicted vascular density, respectively, at time  $t_i$  and voxel  $s_j$ , with  $N_x$  being the number of voxels  $N_x$ . As the parent vessel in the microfluidic platform is not perfectly straight, and some endothelial cells may migrate short distances from the parent vessel without forming mature vessels, we only utilize voxels  $s_j$  that are beyond 15 voxels ( $> 35 \mu\text{m}$ ) away from the parent vessel. This mitigates the effects of endothelial cells that do not form mature vessels which is the focus of our model. This calibration scenario is solved by using a naive quadrature approach as the parameter space is only one dimensional and the computational complexity of the model is prohibitive for sampling. A quadrature approach allows us to utilize parallel resources to run thousands of simulations concurrently, while sampling requires several thousand model runs in serial. For the calibration, we use the day 3, 5, and 7 data and the calibrated probability distribution for distance between new tip cells to predict the vascular density at day 9. We note that the computational and experimental domains used in this scenario are identical ( $676.2 \times 3155.6 \mu\text{m}^2$ ).

We assess the predictive capabilities of the hybrid multiscale model of tumor angiogenesis at the global and local scales by comparing the model prediction, using calibrated parameter distributions, to the observed data at the final time point in each experiment. The uncertainty in the predictions of vascular density and vascular volume fraction is quantified globally in two cases: the 1-parameter case, using only the parameter distribution for distance between new tip cells ( $d_{tip}$ ; values sampled from a Gaussian distribution calibrated in Scenario 4), and the 4-parameter case, using the parameter distributions for VEGF production and consumption ( $\lambda_T^p$  and  $\lambda_c^c$ ), stalk cell divide time ( $d_{sc}$ ), and the distance between new tip cells ( $d_{tip}$ ; values are sampled from Gaussian distributions calibrated from Scenarios 2–4). This is accomplished by sampling from the Gaussian fits of the parameter distributions and running the model with the sampled model parameters to calculate the quantities of interest. The 4-parameter case accounts for uncertainty (and therefore the error) of the calibrations performed in Scenarios 2–4, as well as, the propagation of the uncertainty. In the global analysis, we predict the vascular density and vascular volume fraction, while in the local analysis we predict the centerlines of the angiogenic sprouts.

**2.4.5 Scenario 5: Calibration of local stalk cell divide time to predict vessel morphology.** We employ the confocal microscopy images (see 2.2.3 Image acquisition and processing) to calibrate the hybrid multiscale model to local vascular structures over time. To capture the spatial differences between the model and the data, the optimization problem involves comparing the average distance between the centerlines of the vessels observed in the data and computed by the model. Since the local analysis focuses on specific vascular sprouts, we initialize the mathematical model from the data at the first time point (day 3 of the experiment), calibrate the growth of the stalk cells from days 5 and 7, and predict the vascular structure at day 9. Since the model and data are not of the same type (i.e., the data is a binary field and the model is a list of cells interacting *via* forces), initialization is quite complex. We initialize the calibration by setting the direction of VEGF gradient to closely resemble the true vessel skeletonization at day 3. We then define the parameter estimation problem as: Find  $m = [d_{sc}^*]$  such that

$$\phi(m) = \sum_{i=1}^{N_{days}} \sum_{j=1}^{N_{vox}} (D_{MD}(t_i, x_j, m) - D_{DM}(t_i, x_j, m))^2, \quad (18)$$

is minimized, where  $D_{MD}(t_i, x_j, m)$  is the average distance between the voxels (with number of voxels  $N_x$ ) on the centerline of the model to the data and  $D_{DM}(t_i, x_j, m)$  is the average distance between the voxels on the centerline of the data to the model. Across days 5–7 and days 7–9,

only two regions had a Dice score above 0.5 and a percentage of overlap above the average of all regions analyzed.

To assess the model prediction of local vessel morphology, we implemented the following steps: 1) take 1000 samples from the Gaussian distribution calibrated from Scenario 5 (as well as the VEGF consumption and production rate calibrated in Scenario 2), 2) run the forward model with those 1000 parameter realizations, 3) calculate the centerlines of the forward model solutions, 4) sum over the 1000 centerlines to create a vessel centerline heat map, and 5) threshold the heat map at 10, 50, and 100 to provide the prediction envelope for 1%, 5%, and 10% of the simulations (i.e., voxels covered by 10, 50, and 100 predicted vessel centerlines). The quartiles of the centerline prediction of each region are calculated from MATLAB's 'quantiles'.

### 3 Results

#### 3.1 Calibration Scenarios 1 and 2: Cell number and VEGF concentration

Fig 3 provides an overview of how cell number and ELISA analysis are integrated with the ODE models for the calibration of VEGF production and consumption rates. Cell number was calculated *via* hemocytometer at 24, 48, 72, 120, and 168 hours post seeding. The first time point initialized the TIME (Fig 3B) and IBC3 (Fig 3C) growth models and the subsequent four time points were used to calibrate the models for cell number over time (i.e., Eqs (1) and (3)). The number of cells are independent of VEGF concentration allowing the tumor and endothelial cell number calibration to be carried out first with  $N_e(t)$  and  $N_T(t)$  used as input functions into the production and consumption models of VEGF, respectively. The fit of the growth models is shown in Fig 3B and 3C with growth rates of  $8.9 \times 10^{-3} h^{-1}$  and  $3.7 \times 10^{-2} h^{-1}$  for the IBC3 and TIME cells, respectively, and a carrying capacity of  $2.67 \times 10^5$  for the TIME cells. In the ELISA experiments, every 24 hours for 7 days, VEGF concentration is measured, and media is replaced; note that this action causes the total concentration of VEGF in the plate to return to baseline (i.e., 1100 pg/mL) after each measurement. The VEGF experiments for TIME and IBC3 cells are carried out in different plates, so the calibration of each cell line is not coupled. We utilize naive Gaussian quadrature to calculate the probability distribution of VEGF production by TIME cells, as well as the VEGF consumption and carrying capacity of the IBC3 cells. The maximum likelihood values were obtained using MATLAB's 'lsqnonlin' function, and we then utilized a parameter sweep using 100 values for each of the three parameters (total of a million points) taken about the maximum likelihood values. The resulting posterior distributions are shown in Fig 3D. The fit Gaussian distributions are given by

$$\lambda_T^p \sim 10^{-3} \times N(1.7, 2.3 \times 10^{-1}) \frac{\text{pg}}{\text{mL} \cdot \text{h} \cdot \text{cell}}, \quad (19)$$

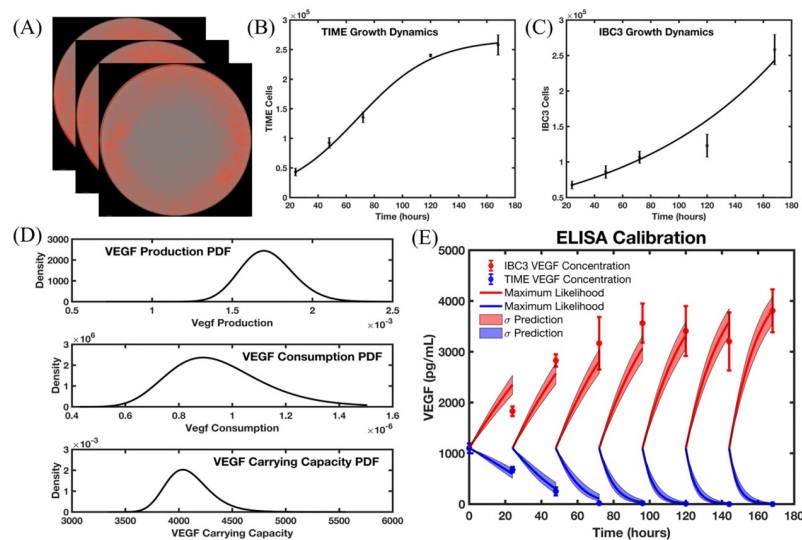
$$\lambda_e^c \sim 10^{-7} \times N(9.1, 2.3) \frac{\text{pg}}{\text{mL} \cdot \text{h} \cdot \text{cell}}, \quad (20)$$

and

$$\theta_e \sim N(4035, 164.6) \text{ pg}, \quad (21)$$

where  $N(a, b)$  denotes a Gaussian distribution with mean  $a$  and standard deviation  $b$ . In Fig 3E, the IBC3 and TIME VEGF concentrations (with standard deviations) are shown as the red and blue dots, respectively. The maximum likelihood value for each parameter (see





**Fig 3. Calibration and prediction scenario for the cell culture and VEGF concentration data.** Fig 3A shows an example of temporally resolved IncuCyte images used for determining the tumor and endothelial cell number (cell fluorescence shown in red). Fig 3B and 3C show the observed number of TIME and IBC3 cells over time with the calibrated logistic growth model for TIME cells and exponential growth for IBC3 cells. Fig 3D shows the posterior distribution function of parameters inferred using the IncuCyte and ELISA data. The calibrated parameters from Fig 3B and 3C are utilized in the models of VEGF secretion and consumption by IBC3 and TIME cells, respectively, to calibrate VEGF production, consumption, and carrying capacity. These posterior distributions are sampled to propagate uncertainty in the ODE models of VEGF secretion and consumption, with the one standard deviation confidence intervals shown in Fig 3E. The VEGF concentration models for tumor and endothelial cell have an average relative error of 11% and 16.7%, respectively. The calibrated parameter distributions for VEGF production and consumption are integrated back into the hybrid multiscale angiogenesis model for subsequent analysis.

<https://doi.org/10.1371/journal.pcbi.1009499.g003>

Section 2.3.4 is

$$\lambda_T^p = 1.7 \times 10^{-3} \frac{\text{pg}}{\text{mL} \cdot \text{h} \cdot \text{cell}}, \quad (22)$$

$$\lambda_e^c = 8.9 \times 10^{-7} \frac{\text{pg}}{\text{mL} \cdot \text{h} \cdot \text{cell}}, \quad (23)$$

and

$$\theta_e = 4035 \text{ pg}. \quad (24)$$

The means (and associated  $\sigma$  prediction envelope) of the best fit of TIME and IBC3 VEGF models are shown in blue and red, respectively. The calculated distributions and maximum likelihood values of VEGF production and consumption rates will be used in subsequent calibration and prediction scenarios described below in Sections 3.2–3.4. The relative errors between the models and experimental measurements are summarized in Table 2.

### 3.2 Calibration Scenario 3: Calibration and prediction of global stalk cell divide time

Fig 4 shows the calibration of length measurements of angiogenic sprouts (imaged *via* confocal microscopy) to the sprout lengths predicted by the vascular ABM. Sprout length measurements are taken on days 3, 7, 11, 15, and 19 after the 72-hour flow protocol (see Section 2.2.2 Vascularized 3D *in vitro* microfluidic platforms). Fig 4A depicts the sprout length

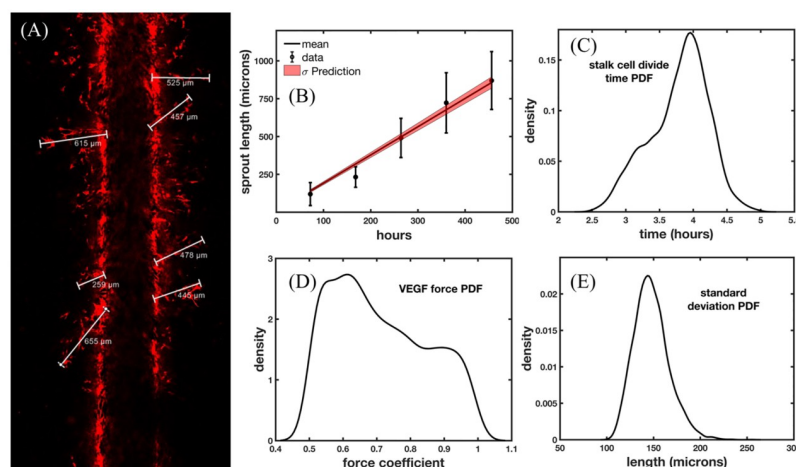
**Table 2. Error between the ODE model output and cell culture and VEGF concentration data.** All errors are relative errors between the model output and experimental data. The dashes (-) denote where the data is not available to calculate the quantity, and the asterisk (\*) denotes when the data records a value of less than 2% of the baseline concentration of 1100 pg/mL, thereby precluding computing the relative error (i.e., it would be near infinite).

Days	Error IBC	Error TIME	Error IBC3 VEGF	Error TIME VEGF
1	2.7%	7.4%	28.5%	12.3%
2	2.9%	5.6%	9.1%	21%
3	-	-	11.2%	*
4	29%	2.6%	13.9%	*
5	-	-	2.3%	*
6	5.8%	1.3%	12%	*
7	-	-	0.5%	*
Average	10.1%	16.9%	11%	16.7%

<https://doi.org/10.1371/journal.pcbi.1009499.t002>

measurements of the microfluidic platform at day 11 using the Leica measurement tool. Fig 4B shows the mean and standard deviation of the data in black and the standard deviation of the prediction of the vascular ABM in light red using the calibrated parameters, along with the mean given by the black line. The posterior distributions of stalk cell divide time, VEGF force, and a hyperparameter standard deviation (which encompasses both the experimental and model uncertainty) are shown in Fig 4C, 4D and 4E, respectively. The error in the calibration is shown in Table 3.

The maximum likelihood values for stalk cell divide time ( $d_{sc}$ ), VEGF force coefficient ( $f_v$ ), and standard deviation ( $\sigma_s$ ) are 3.96 h, 0.61, and 143.5  $\mu\text{m}$ , respectively. The maximum likelihood values (see Section 2.3.4) of the stalk cell divide time and VEGF force values are utilized



**Fig 4. Calibration and prediction scenario for stalk cell divide time and sprout length measurements.** Fig 4A displays the Day 11 length measurements from the confocal microscopy images of the angiogenic vasculature in the 3D microfluidic platform outlined in the experimental methods of Scenario 3 (Section 2.2.2). These specific vessels were selected since they were observable at Day 3 and tracked throughout the remaining time points. The length measurements over time are depicted in Fig 4B, where the dots show the mean of the measurements, and the error bars show the standard deviation. The red region depicts the standard deviation of the prediction of the calibrated mathematical model with calibrated posterior distributions of each parameter shown in Fig 4C and 4D and 4E. Fig 4C depicts the posterior distribution of the stalk cell divide time, Fig 4D depicts the posterior distribution of the VEGF Force, and Fig 4E depicts the posterior distribution of the standard deviation which has been calibrated as a hyperparameter in this scenario. The average error of predicted sprout length over time is 15.3%. The calibrated stalk cell divide time posterior distribution is utilized in the subsequent global analysis.

<https://doi.org/10.1371/journal.pcbi.1009499.g004>

Table 3. Error in calibration and prediction of sprout length using the ABM.

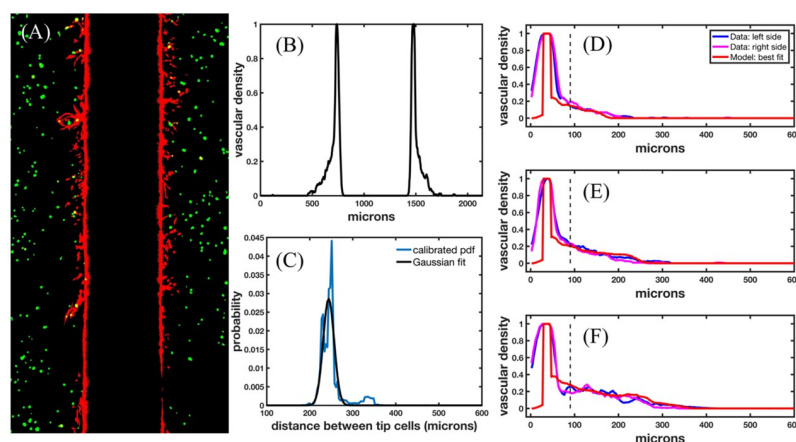
Days	Observed sprout length (microns)	Predicted sprout length (microns)	Predicted error sprout length (%)
3	119.8 ± 75.8	140.2 ± 8.9	17.1
7	232.4 ± 67.7	318.6 ± 14.6	37.1
11	490.6 ± 129.5	497.6 ± 29.5	14.6%
15	722.7 ± 198.7	676.0 ± 29.5	6.3%
19	869.9 ± 191	856.8 ± 37.6	1.5%
Average	-	-	15.3%

<https://doi.org/10.1371/journal.pcbi.1009499.t003>

in the calibration of distance between new tip cells. Approximating these PDFs as Gaussians, the parameters are  $d_{sc} \sim N(3.96, 0.44)$ ,  $f_v \sim N(0.71, 0.14)$ , and  $\sigma_s \sim N(147.8, 18.6)$ . The Gaussian fit of  $d_{sc}$  is utilized moving forward (specifically, in Scenario 3.4) to assess the uncertainty in the model prediction of vascular density and volume fraction. As the increase in vessel length is only due to the division of stalk cells, we can ignore the uncertainty contribution of the VEGF force. This constant acts to maintain physiological sprouting that does not break the vessel but balances the forces to allow vessel elongation; i.e., if this force is higher than the adhesion between endothelial cells, the tips cells would disconnect from the sprout. Therefore this uncertainty contribution does not affect the resulting vascular network.

### 3.3 Scenario 4: Calibration and prediction of vascular density

Fig 5 presents the calibration of the distance between new tip cells by utilizing the vascular density of angiogenic sprouts imaged in the microfluidic platform. Fig 5A depicts a representative binary image of the angiogenic sprouts in the microfluidic platform, and Fig 5B shows the calculated density of the angiogenic sprouts. The vascular density is calculated by summing



**Fig 5. Calibration of distance between tip cells.** Fig 5A displays a representative binarized RGB image (as each channel has been binarized *via* thresholding and can only take values of 0 or 1) from the confocal microfluidic platform used to calculate the vascular density, shown in Fig 5B. Fig 5C depicts the calibrated PDF and Gaussian fit of the distance between new tip cells. Fig 5D-F show the density calculated from the left and right side of the microfluidic platform shown in blue and magenta, respectively, and the best fit of the model shown in red for days 3, 5, and 7, respectively. Specifically, the model best fit is calibrated using days 3 and 5 and then temporally evolved to compare day 7 with the observed experimental data. The vertical dashed lines represent the location away from the parent vessel that we begin to use for calibration (i.e., we utilize the voxels to the right of the dashed vertical line, see Scenario 4, Section 2.4.4). The relative error of the best fit in the calibration for days 3, 5, and 7 are 23.5%, 11.1%, and 18.5% respectively.

<https://doi.org/10.1371/journal.pcbi.1009499.g005>

**Table 4. Error in calibration and prediction of vascular density.** We compare the relative error in the calibration best fit, the 1-parameter, and 4-parameter prediction cases.

Days	Calibration best fit % error	1-Parameter prediction mean % error	4-Parameter prediction mean % error
3	23.6	20.9	21.9
5	11.1	15.9	18
7	18.5	23.8	25.2
Average	17.7	20.2	21.7

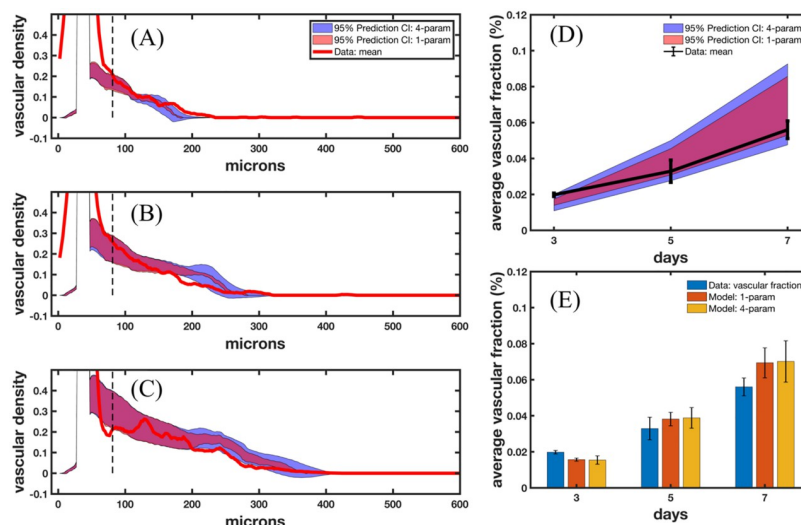
<https://doi.org/10.1371/journal.pcbi.1009499.t004>

the vessel mask along the length of the vessel. Since the left- and right-hand side are not able to be compared directly with the computational domain, we align the vascular density calculated from the image and the model initial conditions (see the blue and magenta lines in Fig 5D–5F for an example alignment). In Fig 5C, we show the calibrated PDF of the distance between new tip cells in blue and the Gaussian fit of the PDF, which may be readily sampled for analyzing the uncertainty in the prediction of the model. In Fig 5D–5F, we show days 3, 5, and 7 of the calculated density of the left and right sides of the microfluidic platform (blue and magenta, respectively), and the model best fit obtained during calibration over days 3 and 5 (red). The day 7 model best fit is the temporally evolved solution of the model best fit calibrated from days 3 and 5. The vertical dashed line represents the location from which we begin to make use of the vascular density observed in the microfluidic platform for the calibration. That location is chosen to minimize the effects of migratory endothelial cells close to the channel wall that do not form mature angiogenic sprouts (see Section 2.4.4. The Gaussian fit of the PDF is  $d_{tip} \sim N(243.9, 18.1)$  microns. The error in the calibration and prediction is shown in Table 4.

We utilize 1000 samples from the probability distribution of the distance between new tip cells, as well as the probability distributions obtained in scenarios 2 and 3, to rigorously quantify the uncertainty in the predictability of the model. The prediction results are shown in Fig 6. In Fig 6A–6C, we show the 95% prediction confidence intervals propagating uncertainty using one parameter distribution for distance between new tip cells shaded in light red and using four parameter distributions (distance between new tip cells, VEGF production and consumption rates, and stalk cell growth rate) shown in light blue. The vascular density calculated from the left and the right sides of the microfluidic platform are shown in blue and magenta, respectively, and the model best fit is shown in red. The vertical black dotted line denotes the distance away from the vessel that we begin to use for calibration. The total uncertainty (summed over space) was 97.4%, 66.2% and 45.2% higher in the four-parameter case than the 1-parameter case on days 3, 5, and 7, respectively. Fig 6D shows the calculated average vascular fraction from the data in black, and the predicted average vascular fraction for the one parameter (blue) and four parameter (red) uncertainty analyses. In Fig 6E, the vascular fraction of the data, 1-parameter, and 4-parameter scenarios are shown in blue, orange, and yellow, respectively.

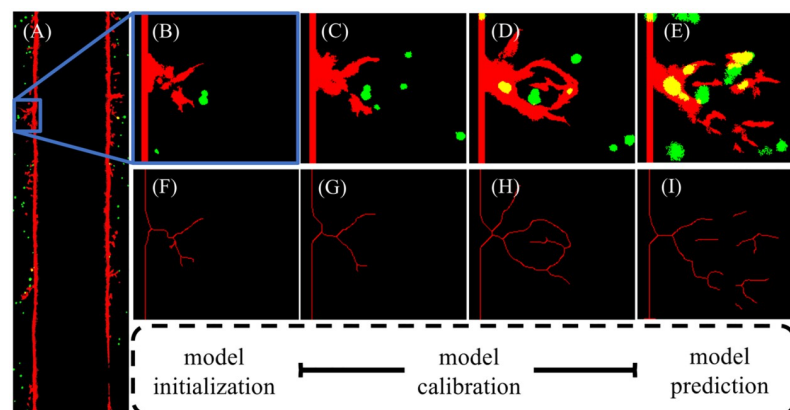
### 3.4 Scenario 5: Calibration of local stalk cell divide time to predict vessel morphology

While sprout length, vascular density, and volume fraction over time are reasonable metrics for analyzing global features of vascular structure, additional local analysis is needed to assess the model's ability to recapitulate local features in the vascular network. Fig 7 shows the local region selected from the microfluidic platform for the calibration of stalk cell growth rate.



**Fig 6. Prediction of vascular density.** Fig 6A–C show the uncertainty in the model prediction assuming uncertainty in one parameter calibrated from scenario 4 (shaded in light red, outlined in black) and four parameters (shaded in light blue, outlined in black) calibrated from calibration scenarios 2, 3, and 4, along with the mean of the data. In the 1-parameter case, we consider only the distance between new tip cells with a PDF shown in Fig 5C. In the 4-parameter case, we consider distance between new tip cells, VEGF production and consumption rates (shown in Fig 3D), and stalk cell growth rate (shown in Fig 4B). Fig 6D presents the volume fraction calculated from the data in black, and the predicted volume fraction from the 1-parameter and 4-parameter case in light red and blue, respectively. Fig 6E depicts the prediction of volume fraction compared to the data (blue), and the 1-parameter and 4-parameter cases in orange and yellow, respectively, with the corresponding standard deviations shown in black. The average error in vascular volume fraction is 20.2% and 21.7% of the 1-parameter and 4-parameter case, respectively. We also note the drastic increase in uncertainty moving away from the parent vessel in the 4-parameter case (Fig 6A 120–180 microns, Fig 6B 200–280 microns, and Fig 6C 200–400 microns), highlighting the effects of the VEGF production and consumption rates and the stalk cell divide time on the uncertainty in the prediction.

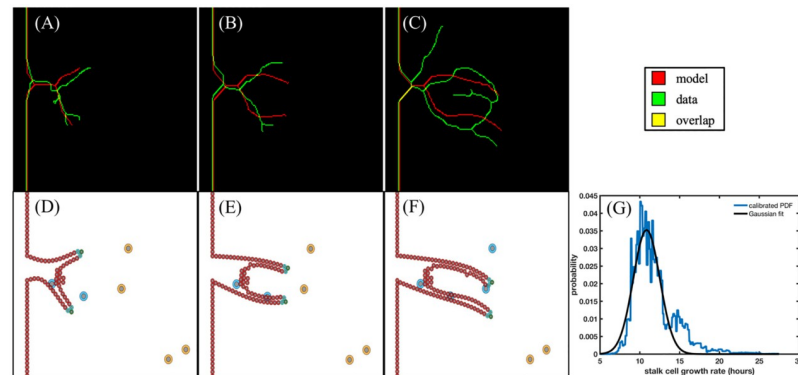
<https://doi.org/10.1371/journal.pcbi.1009499.g006>



**Fig 7. Vascular regions for local calibration.** Fig 7A shows a RGB binarized image of vascular structure of day 3 in the microfluidic platform. Fig 7B–E depict a specific local region segmented over time (blue box in Fig 7A) with vessels shown in red and tumor cells shown in green from days 3, 5, 7, and 9. Fig 7F–I show the corresponding centerline segmentation of Fig 7B–C, computed from a parallel thinning algorithm. The local analysis flowchart is shown by model initialization using the day 3 centerline, model calibration utilizing days 5 and 7, and model prediction of the centerline from day 9.

<https://doi.org/10.1371/journal.pcbi.1009499.g007>



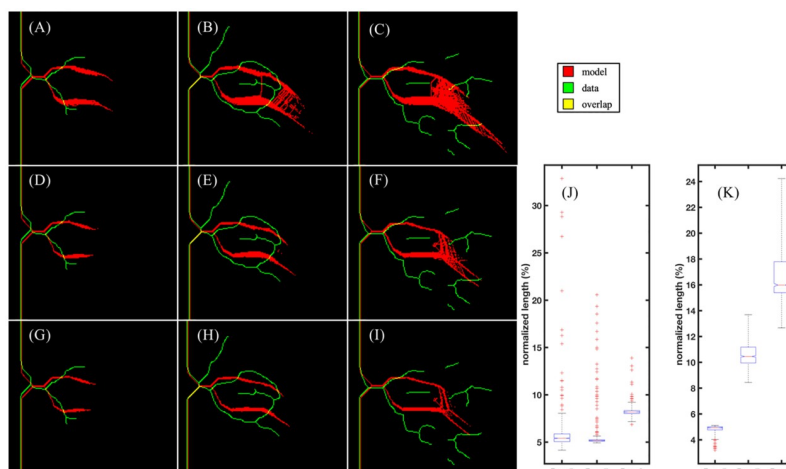


**Fig 8. Local region 1: Calibration of stalk cell growth rate.** Fig 8A–C show the centerlines calculated from the model best fit (red), the data (green), and the overlap (yellow) on day 3 (used to inform the initial conditions), day 5, and day 7 (used to calibrate the stalk cell growth rate). The ABM is shown in Fig 8D–F with tip cells in green, stalk cells in cyan, endothelial cells in red, tumor cells releasing VEGF in orange, and tumor cells not releasing VEGF in blue. In Fig 8G, we show the calibrated PDF and the Gaussian fit of the stalk cell growth rate. The best fit model recapitulates the general structural features of the data without allowing additional sprouts to form. We also note that from day 7 (depicted in Fig 8F) to day 9, only the two tumor cells in the bottom right of the domain continue to release VEGF, guiding the sprout migration to the bottom right of the domain.

<https://doi.org/10.1371/journal.pcbi.1009499.g008>

Confocal images are taken on days 3, 5, 7, and 9, and processed in MATLAB using a Dice score and area threshold to select regions for local calibration (see 2.2.3 Image acquisition and processing). Once selected by our sorting algorithm, we calculate the skeletonization of the local vascular structures using a Zhang-Suen parallel thinning algorithm [56]. We use the day 3 skeletonization, shown in Fig 8A to initialize the model (Fig 8D) by segmenting the tumor cells from the surrounding region and calculating the trajectories of the tip cells shown in the data to formulate the initial conditions of the ABM that match the skeletonization of the data. The stalk cell growth rate is calibrated by running the model forward and comparing the centerline distance between the data at days 5 and 7 (Fig 8B and 8C, respectively) and the vasculature predicted by the model with the maximum likelihood stalk cell growth rate, shown in Fig 8C and 8F. The resulting probability distribution function of stalk cell growth rate (blue) and the Gaussian fit (black) are shown in Fig 8G. The maximum likelihood value for stalk cell growth rate in this scenario is 10.86 hours and the Gaussian fit is  $d_{sc} \sim N(10.85, 2.4)$  hours. We then take 1000 samples from this Gaussian distribution and calculate the resulting vasculature for days 5, 7, and 9 as shown in Fig 8.

In Fig 9, the columns denote day 5, 7, and 9, respectively. The rows depict the 1%, 5%, and 10% prediction envelope (top to bottom, described in 2.2.3 Image acquisition and processing). In Fig 9B, the data and the model calculated centerlines form an anastomosis, predicted in 3.3% of simulations. Anastomosis is also predicted on day 9 in Fig 9C, 9F and 9I in 16.7% of simulations, recapitulating the vessel structure observed in the data at day 7. In Fig 9J, the prediction of the average centerline distance from model to data is calculated from the 1000 samples of the distribution in Fig 8G. Fig 9K presents the normalized average centerline distance from the data to model (normalized to the length of the sprout in day 3). The prediction quartiles of this region and another (shown in S1 Appendix) is shown in Table 5. Finally, Table 6 shows the calibrated model parameters and their Gaussian fit. In particular, we note that  $s_g$  is the global sprout elongation rate which is independent of cell diameter and is calculated from the stalk cell divide time and the growth time of stalk cells,  $g_{sc}$ .



**Fig 9. Local region 1: Prediction.** Each column depicts the data (green), model prediction (red), and overlap of the two (yellow) on days 5, 7, and 9, respectively. Each row shows the predicted centerlines using three different thresholds: 99% prediction (Fig 9A-C: top row), 95% prediction (Fig 9D-F: middle row), and 90% prediction (Fig 9G-I: bottom row) of the simulations. While the data exhibits vessel anastomosis at Day 7, the model prediction only depicts anastomosis in the 1% prediction at Day 7 (Fig 9B) and Day 9 (Fig 9C and 9F and 9I). Fig 9J and 9K shows the prediction of the average centerline distance from model to data and from data to model, respectively. The average centerline distance is normalized by the length of the longest sprout in this region at day 3. This results in a normalized length of less than 10% from model to data and a normalized length of less than 20% from data to model. However, at day 9 the complexity in the vascular network, specifically the vascular remodeling that eliminates the anastomosis, is beyond the capabilities of the model to predict.

<https://doi.org/10.1371/journal.pcbi.1009499.g009>

**Table 5. Error in the prediction of vessel centerlines.** We compare the relative error in the prediction of vessel centerlines calculated from simulations of local vasculature (shown in Fig 9).  $C_{dist}(d, m)$  is the average centerline distance between the skeletonized vessel in the data,  $d$ , and the skeletonized vessel predicted by the model,  $m$ .

Region	Quartiles	$C_{dist}(m, d)$			$C_{dist}(d, m)$		
		25%	50%	75%	25%	50%	75%
1	Day 5	5.1%	5.4%	5.9%	4.8%	4.9%	5.0%
	Day 7	5.1%	5.2%	5.3%	10%	10.5%	11.2%
	Day 9	8.0%	8.2%	8.3%	15.4%	16%	17.8%
2	Day 5	0.7%	0.7%	0.7%	3.6%	3.6%	3.6%
	Day 7	1.1%	1.1%	1.1%	4.7%	4.8%	4.8%
	Day 9	1.2%	1.5%	1.8%	8.9%	9.0%	9.3%

<https://doi.org/10.1371/journal.pcbi.1009499.t005>

**Table 6. Calibrated model parameters.**

Parameter	Interpretation	Source	MLE value	Gaussian fit
$\lambda_e^g$	Growth rate of TIME cells	Scenario 1	$3.7 \times 10^{-2} \text{ h}^{-1}$	-
$\theta_e$	Carrying capacity of TIME cells	Scenario 1	4035 cells	$N(4035, 164.6)$
$\lambda_T^g$	Growth rate of IBC3 cells	Scenario 1	$8.9 \times 10^{-3} \text{ h}^{-1}$	-
$\lambda_T^p$	VEGF production rate by IBC3 cells	Scenario 2	$1.7 \times 10^{-3} \text{ pg h}^{-1}$	$10^{-3} \times N(1.7, 2.3 \times 10^{-1})$
$\lambda_e^c$	VEGF consumption rate by TIME cells	Scenario 2	$8.9 \times 10^{-7} \text{ pg h}^{-1}$	$10^{-7} \times N(9.1, 2.3)$
$\theta_v$	Carrying capacity of VEGF	Scenario 2	$2.67 \times 10^5 \text{ pg}$	-
$d_{sc}$	Angiogenic stalk cell divide time	Scenario 3	3.96 h	$N(3.96, 0.44)$
$s_e$	Sprout elongation rate	$2R^S/(d_{sc} + g_{sc})$	$2.0 \mu\text{m h}^{-1}$	-
$f_v$	VEGF force coefficient	Scenario 3	0.61	$N(0.71, 0.14)$
$d_{tip}$	Distance between new tip cells	Scenario 4	251 $\mu\text{m}$	$N(243.0, 18.1)$
$d_{sc}^*$	Local stalk cell divide time rate	Scenario 5	10.86 h	$N(10.85, 2.4)$

<https://doi.org/10.1371/journal.pcbi.1009499.t006>

## 4 Discussion

We have rigorously calibrated a hybrid, multiscale model of tumor angiogenesis, both with global and local vascular structure, with experimental data obtained from an *in vitro* vascularized tumor platform and tested the ability of the model to make accurate longitudinal predictions. This was done by first utilizing hemocytometry and ELISA to calibrate mathematical models of cell growth and VEGF concentration, respectively, yielding parameter distributions for VEGF production and consumption rates (Fig 3). The maximum likelihood values are used in subsequent calibration and prediction scenarios. Secondly, we analyzed global features such as sprout length, vascular density, and vascular volume fraction over time, to calibrate the global growth rate of angiogenic sprouts and the distance between new tip cells, shown in Figs 4 and 5. We then reported the errors and rigorously accounted for the uncertainty in model predictions, in Table 4 and Fig 6, respectively. The mean error in the prediction of the 1- and 4-parameter cases was 21% and did not increase with time. This is of note because the model predicts from day 0, not from previous time points (i.e., from day 5 to day 7), making prediction significantly more challenging (see Section 2.4.5). Finally, we analyzed the ability of the model to recapitulate local vascular structure by segmenting and skeletonizing specific vessels over time, initializing the model with confocal images at day 3, calibrating with images at day 5 and 7, and testing the predictability of the model against day 9 observations, as summarized by Fig 7. Our calibrated hybrid multiscale model, informed by hemocytometry and VEGF concentration data, can recapitulate local vascular structure longitudinally in our *in vitro* microfluidic platform. The calibration and prediction error, shown in Table 5, was on the order of 10% normalized centerline distance. The normalized length of the centerline distance from data to model increases across time and indicates that the hybrid ABM using the initial conditions of two tip cells (see Fig 9D), while quantitatively and qualitatively matching the description of vasculature at day 7, fails to capture the complexity observed at day 9. This is partially due to the restriction of new tip cells in the local calibrations, as we do not allow for new sprouts. However, the complexity across time points of the local sprouts dramatically increases and requires further study from both a modeling and experimental standpoint.

To the best of our knowledge, this work represents the first study to rigorously calibrate an agent-based model of tumor angiogenesis with multimodal experimental data to forecast vascular growth that is then directly compared to the corresponding data. Importantly, other efforts have pioneered the integration of data in other ways [65, 66]. For example, Perfahl *et al.* employed multiphoton microscopy to image angiogenic vasculature in an *in vivo* dorsal skin fold chamber [7]. These images furnished the initial conditions of a cellular automaton model of vascular tumor growth, with blood flow, and vascular remodeling. However, no model calibration or comparison to experiments was performed. Similarly, in Xu *et al.*, photoacoustic imaging of a murine xenograft model initialized a phase-field model of tumor angiogenesis [38]. While the authors did not utilize time-resolved vascular data, this experimental-computational approach showcases a first step toward longitudinally integrating *in vivo* data into angiogenesis models, through initialization, calibration, and ultimately, prediction. The present study incorporates initialization, calibration, and prediction through utilizing *in vitro* data; however, the experimental and computational approaches in this contribution can be made model-agnostic (and, even, data-agnostic). In particular, all the calibration methods were based on gradient free methods, and all the solvers are independent of the model, data, and the system under investigation.

Many of the calibrated parameter values are readily comparable to previous experimental studies and can be used in future modeling efforts. Specifically, discrete or hybrid mathematical models that include tip cell velocity, which is comparable to stalk cell divide time in our

work, can be informed from the values reported in Table 6 [37–39, 67]. However, in continuous models, specifically in the landmark Anderson 1998 model [5], chemotaxis during angiogenesis depends on the magnitude of the gradient of VEGF. As the magnitude of this value is unknown in our experiments, direct comparisons to the Anderson model (and any model with a purely continuous description of endothelial cells) is not feasible. While our calibrated cell doubling time (stalk cell divide time plus stalk cell growth time) is 2–3× less than what is observed in other experimental protocols, we use a simplified cell geometry with a cell diameter of 10 microns, 2–3× less than what has been observed in endothelial cells. This puts our doubling time within the range measured by others [68, 69]. With regards to endothelial cell migration rates, endothelial cell velocity (in the context of random walks) has been measured to be 10–50 microns per hour in various conditions, while our model calculates an average sprout elongation rate (velocity) of 2.0 microns per hour. This difference is almost certainly because of the very different experimental scenarios considered in the literature which use a 2D framework and neglects the necessity of cells needing to degrade the extracellular matrix for migration. Our system is a 3D framework and explicitly accounts for degradation of the extracellular matrix [48], which necessarily results in a lower average velocity.

A key experimental limitation of the present study is the lack of longitudinal measures of various proteins (*e.g.*, VEGF) and cell markers (*e.g.*, to assign endothelial cells as tip, stalk, phalanx phenotype) in the *in vitro* microfluidic platform. We have addressed this by devising smaller calibration and prediction scenarios (with corresponding experiments) that can isolate the salient phenomena (*e.g.*, VEGF production from tumor cells and growth of stalk cells) that we assume are fundamental properties of the cell and therefore independent of the experimental scenario. Another area for advancement is to determine the generalizability of both the local and global calibration results by applying the approach to multiple microfluidic platforms. We utilized one microfluidic platform over time, yielding two regions for global analysis and two local regions for calibration and prediction. Another fundamental limitation is the approximation that parameter distributions calibrated in simpler experimental/computational scenarios are valid for more complex scenarios. However, due to the complexity in both the 3D microfluidic platform and the tumor angiogenesis model, we are not able to isolate quantities of interest to inform individual model parameters in these complex scenarios. For example, we do not have additional temporal measurements of VEGF concentration to inform model parameters. We believe that by approximating these parameters using readily-available quantitative data from simpler scenarios (*i.e.*, the 2D experiments) we can more accurately approximate the real parameter values in the 3D microfluidic scenario than by attempting to calibrate all the parameters using the limited 3D data. We note that in our global analysis in Scenario 4, as we utilize an ABM for angiogenesis, the model cannot be uniquely initialized from day 3 or day 5 and is thus restarted from day 0 with no angiogenic vasculature for each simulation. This is because the model is not continuous and cannot be populated from the vascular density at days 3 or 5, but the endothelial cells in the ABM must start from a parent vessel.

A key limitation in the modeling is that our calibration was only performed on 2D data. As out of plane effects could play a significant role in vessel sprouting, future efforts will be focused on extending the formalism to 3D. The main limitation to extending to 3D is the characterization of the angiogenic sprouts by endothelial cells that make up both sides (in 2D) of the sprout. This can be overcome by describing the vessels as nodes (which describe the radius of the vessel) connected by edges (which describe the directionality of the vessels), an approach taken by [70, 71]. Another key modeling limitation is the use of the maximum likelihood values from Scenarios 1 and 2, as opposed to Bayesian inference in Scenarios 3–5. If Bayesian inference is used in Scenarios 1 and 2, calculating the likelihood becomes a matter of sampling

from the parameter distributions which increases the computational cost of the subsequent calibrations by the number of samples needed to characterize those distributions. In future work, we aim to utilize more experimental replicates to further calibrate and validate our model and to devise experimental protocols with anti-angiogenic and radiation therapies to model the effects of these treatments on angiogenic sprouts *in vitro* [72].

## 5 Conclusion

We have calibrated and determined the ability of an agent-based model to make accurate predictions of tumor angiogenesis by systematically incorporating data of different scales to inform model parameters. To the best of our knowledge, this represents the first effort to calibrate a mechanism-based mathematical model to spatially and temporally-resolved experimental data of angiogenesis, thereby enabling predictions of future vessel development that could then be directly tested against observation.

## Supporting information

**S1 Appendix. Model invertibility, Gaussian fit of parameters, and time evolved solution of the ABM.** In this appendix, we show the invertibility of the models by calibrating with *in silico* data. We then explore the accuracy of the Gaussian fits of the parameter distributions calibrated in Scenarios 1-5. The longitudinal prediction of the ABM is presented after calibration during Scenarios 1-4 and another region of local vasculature is presented.  
(DOCX)

**S1 Fig. Model invertibility of ODE models.** Panels (A)-(C) depict the parameter distributions from calibrating VEGF production, consumption, and carry capacity, respectively, to 5% Gaussian noised VEGF concentration data. Panel (D) shows the goodness of fit of the VEGF concentration in the tumor (IBC3) and endothelial (TIME) experiments, shown in red and blue, respectively. The relative error between the parameters used to generate the data and the calibrated parameters are 9.3%, 2.0%, and 5.8%.  
(TIF)

**S2 Fig. Model invertibility of ABM to longitudinal density measurements.** This plot shows the calibrated PDF when the model is calibrated against synthetic data.  
(TIF)

**S3 Fig. Solution of the agent-based model in time.** Endothelial cells—red, tip cells—green, deactivated endothelial cells—yellow green, normoxic tumor cells—blue, hypoxic tumor cells—orange. Days 3-13 are predicted using informed parameters calibrated in Scenarios 3 and 4. The x- and y-axis are in microns. The vasculature grows and develops in response to vascular endothelial growth factor, resulting in the final vascular network, with a highlighted region shown in blue. This region has several anastomosing vessels.  
(TIF)

**S4 Fig. Local region 2: Calibration of stalk cell growth rate.** Panels (A)–(C) show the centerlines calculated from the model best fit in red, the data in green, and the overlap in yellow of day 3 (used to inform the initial conditions), Day 5, and day 7 (both used to calibrate the stalk cell growth rate). The agent-based model is shown in Panels (D)–(F) with tip cells in green, stalk cells in cyan, endothelial cells in red, tumor cells releasing VEGF in orange, and tumor cells not releasing VEGF in blue. In Panel (G), we show the calibrated PDF and the Gaussian fit of the stalk cell growth rate. The mean of the local stalk cell divide time ( $\sim 80$  hours) is significantly higher than the global stalk cell divide time calibrated at  $\sim 4$  hours. This highlights



the effect of the hypoxic tumor cells in the top right portion of the domain, as a higher calibrated stalk cell divide time would cause growth toward this region of space, while the data continues toward the bottom right of the domain.

(TIF)

**S5 Fig. Local region 2: Prediction of local vascular structure.** Each column depicts the data (green) and model prediction (red), with overlap in yellow, of days 5, 7, and 9, respectively. Each row shows the voxels predicted by the centerlines using different thresholds, with the voxels in the 99% prediction (top row), 95% prediction (middle row), and 90% prediction (bottom row) of the simulations. In Panel (C), the direction of vessel growth in the model shows the VEGF gradient going away from the direction of data and toward the hypoxic cells shown in Fig 9F. This leads to the overall directionality of the vessel to be misaligned with the data. Panels (J) and (K) shows the prediction of average centerline value from model to data and from data to model with an averaged normalized length difference of less than 2% and 12%, respectively. The average centerline distance is normalized by the length of the longest sprout in this region at day 3.

(TIF)

**S1 Table. Quantiles of calibrated parameter distributions and Gaussian fits.**  $\lambda_t^p$ —production rate of VEGF by tumor cells;  $\lambda_e^c$ —consumption rate of VEGF by endothelial cells;  $d_{SC}$ —stalk cell divide time;  $d_{tip}$ —distance between new tip cells.

(PDF)

## Acknowledgments

The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC resources that have contributed to the research results reported within this paper.

## Author Contributions

**Conceptualization:** Caleb M. Phillips, Ernesto A. B. F. Lima, Thomas E. Yankeelov.

**Data curation:** Manasa Gadde.

**Funding acquisition:** Thomas E. Yankeelov.

**Investigation:** Caleb M. Phillips, Ernesto A. B. F. Lima, Manasa Gadde.

**Methodology:** Caleb M. Phillips, Ernesto A. B. F. Lima, Manasa Gadde, Angela M. Jarrett.

**Software:** Caleb M. Phillips.

**Supervision:** Marissa Nichole Rylander, Thomas E. Yankeelov.

**Writing – original draft:** Caleb M. Phillips, Ernesto A. B. F. Lima, Manasa Gadde, Thomas E. Yankeelov.

**Writing – review & editing:** Caleb M. Phillips, Ernesto A. B. F. Lima, Manasa Gadde, Angela M. Jarrett, Marissa Nichole Rylander, Thomas E. Yankeelov.

## References

1. Lima EABF, Oden JT, Almeida RC. A hybrid ten-species phase-field model of tumor growth. *Mathematical Models and Methods in Applied Sciences*. 2014; 24:2569–2599. <https://doi.org/10.1142/S0218202514500304>

2. Lima EABF, Oden JT, Hormuth DA II, Yankeeelov TE, Almeida RC. Selection, calibration, and validation of models of tumor growth. *Mathematical Models and Methods in Applied Sciences*. 2016; 26. <https://doi.org/10.1142/S021820251650055X> PMID: 28827890
3. Jarrett AM, Hormuth DA, Barnes SL, Feng X, Huang W, Yankeeelov TE. Incorporating drug delivery into an imaging-driven, mechanics-coupled reaction diffusion model for predicting the response of breast cancer to neoadjuvant chemotherapy: theory and preliminary clinical results. *Physics in Medicine & Biology*. 2018; 63. <https://doi.org/10.1088/1361-6560/aac040> PMID: 29697054
4. Hormuth DA II, Weis JA, Barnes SL, Miga MI, Rericha EC, Quaranta V, et al. A mechanically coupled reaction-diffusion model that incorporates intra-tumoural heterogeneity to predict in vivo glioma growth. *Journal of the Royal Society Interface*. 2017; 14(128).
5. Anderson AR. A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion. *Math Med Biol*. 2005; 22:163–186. <https://doi.org/10.1093/imammb/dqi005> PMID: 15781426
6. Macklin P, Edgerton ME, Thompson AM, Cristini V. Patient-calibrated agent-based modelling of ductal carcinoma in situ (DCIS): from microscopic measurements to macroscopic predictions of clinical progression. *Journal of theoretical biology*. 2012; 301:122–140. <https://doi.org/10.1016/j.jtbi.2012.02.002> PMID: 22342935
7. Perfahl H, Byrne HM, Chen T, Estrella V, Alarcón T, Lapin A, et al. Multiscale Modelling of Vascular Tumour Growth in 3D: The Roles of Domain Size and Boundary Conditions. *PLoS One*. 2011. <https://doi.org/10.1371/journal.pone.0014790> PMID: 21533234
8. Olsen MM, Siegelmann HT. Multiscale Agent-based Model of Tumor Angiogenesis. *Procedia Computer Science*. 2013; 18:1016–1025. <https://doi.org/10.1016/j.procs.2013.05.267>
9. Lambert AW, Pattabiraman DR, Weinberg RA. Emerging Biological Principles of Metastasis. *Cell*. 2017; 168:670–691. <https://doi.org/10.1016/j.cell.2016.11.037> PMID: 28187288
10. DeBerardinis RJ, Lum JJ, Hatzivassiliou G, Thompson CB. The Biology of Cancer: Metabolic Reprogramming Fuels Cell Growth and Proliferation. *Cell Metab*. 2008; 7:11–12. <https://doi.org/10.1016/j.cmet.2007.10.002> PMID: 18177721
11. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med*. 2013; 19:1423–1437. <https://doi.org/10.1038/nm.3394> PMID: 24202395
12. Bauer AL, Jackson TL, Jiang Y. A Cell-Based Model Exhibiting Branching and Anastomosis during Tumor-Induced Angiogenesis. *Science*. 2005; 307:58–62.
13. Hawkins-Daarud A, Jonhston SK, Swanson KR. Quantifying Uncertainty and Robustness in a Biomathematical Model-Based Patient-Specific Response Metric for Glioblastoma. *JCO Clinical Cancer Informatics*. 2019. <https://doi.org/10.1200/CCI.18.00066>
14. Jarret AM, Lima EABF, Hormuth DA II, McKenna MT, Feng X, Ekrut DA, et al. Mathematical models of tumor cell proliferation: A review of the literature. *Expert Review of Anticancer Therapy*. 2018; 18:1271–1286. <https://doi.org/10.1080/14737140.2018.1527689>
15. Jain RK, Martin JD, Stylianopoulos T. The role of mechanical forces in tumor growth and therapy. *Annual Review of Biomedical Engineering*. 2014; 16:321–346. <https://doi.org/10.1146/annurev-bioeng-071813-105259> PMID: 25014786
16. Weis JA, Miga MI, Arlinghaus LR, Li X, Abramson V, Chakravarthy AB, et al. Predicting the Response of Breast Cancer to Neoadjuvant Therapy Using a Mechanically Coupled Reaction-Diffusion Model. *Cancer Research*. 2015; 75:4697–4707. <https://doi.org/10.1158/0008-5472.CAN-14-2945> PMID: 26333809
17. Enderling H, Sunassee E, Caudell JJ. Predicting patient-specific radiotherapy responses in head and neck cancer to personalize radiation dose fractionation. *bioRxiv*. 2019; <https://doi.org/10.1101/630806>
18. Nani F, Freedman HI. A mathematical model of cancer treatment by immunotherapy. *Mathematical Biosciences*. 2000; 163:159–199. [https://doi.org/10.1016/S0025-5564\(99\)00058-9](https://doi.org/10.1016/S0025-5564(99)00058-9) PMID: 10701303
19. Martinson WD, Ninomiya K, Byrne HM, Maini PK. Comparative analysis of continuum angiogenesis models. *Journal of Mathematical Biology*. 2021; 82(21). <https://doi.org/10.1007/s00285-021-01570-w> PMID: 33619643
20. Stepanova D, Byrne HM, Maini PK, Alarcón T. A multiscale model of complex endothelial cell dynamics in early angiogenesis. *PLoS Computational Biology*. 2021; 17(1):e1008055. <https://doi.org/10.1371/journal.pcbi.1008055> PMID: 33411727
21. Kazerouni AS, Gadde M, Gardner A, Hormuth DA II, Jarrett AM, Johnson KE, et al. Integrating quantitative assays with biologically-based mathematical modeling for predictive oncology. *Iscience*. 2020; p. 101807. <https://doi.org/10.1016/j.isci.2020.101807> PMID: 33299976
22. Hormuth DA II, Phillips C, Wu C, Lima EABF, Lorenzo G, et al. Biologically-Based Mathematical Modeling of Tumor Vasculature and Angiogenesis via Time-Resolved Imaging Data. *Cancers*. 2021; 13(12):3008. <https://doi.org/10.3390/cancers13123008>

23. Yankeelov TE, Atuegwu N, Hormuth D, Weis JA, Barnes SL, Miga MI, et al. Clinically relevant modeling of tumor growth and treatment response. *Sci Transl Med*. 2013; 5. <https://doi.org/10.1126/scitranslmed.3005686> PMID: 23720579
24. Yankeelov TE, Quaranta V, Evans KJ, Rericha EC. Toward a science of tumor forecasting for clinical oncology. *Cancer Res*. 2015; 75:918–923. <https://doi.org/10.1158/0008-5472.CAN-14-2233> PMID: 25592148
25. Rockne RC, Hawkins-Daarud A, Swanson KR, Sluka JP, Glazier JA, Macklin P, et al. The 2019 mathematical oncology roadmap. *Physical biology*. 2019; 16(4):041005. <https://doi.org/10.1088/1478-3975/ab1a09> PMID: 30991381
26. Jathoul AP, Laufer J, Ogunlade O, Treeby B, Cox B, Zhang E, et al. Deep in vivo photoacoustic imaging of mammalian tissues using a tyrosinase-based genetic reporter. *Nature Photonics*. 2015; 9:239–246. <https://doi.org/10.1038/nphoton.2015.22>
27. Ehling J, Theek B, Gremse F, Baetke S, Möckel D, Maynard J, et al. Micro-CT imaging of tumor angiogenesis: Quantitative measures describing micromorphology and vascularization. Deep in vivo photoacoustic imaging of mammalian tissues using a tyrosinase-based genetic reporter. *Am J Pathol*. 2014; 184(2):431–441. <https://doi.org/10.1016/j.ajpath.2013.10.014> PMID: 24262753
28. Stamatelos SK, Kim E, Pathak AP, Popel AS. A bioimage informatics based reconstruction of breast tumor microvasculature with computational blood flow predictions. *Microvas Res*. 2014; 91:8–21. <https://doi.org/10.1016/j.mvr.2013.12.003> PMID: 24342178
29. McDonald DM, Choyke PL. Imaging of angiogenesis: from microscope to clinic. *Nat Med*. 2003; 9:713–725. <https://doi.org/10.1038/nm0603-713> PMID: 12778170
30. Gadde M, Lima EABF, Phillips CM, Rylander MN, Yankeelov TE. A multi-scale model of tumor growth and angiogenesis calibrated by data from a tissue-mimicking, 3D, microfluidic platform. In *Proceedings of the Cancer Research UK-AACR Joint Conference on Engineering and Physical Sciences in Oncology*, London, England. 15–17 October 2019;9:713–725.
31. Lieberman C, Willcox K, Ghattas O. Parameter and state model reduction for large-scale statistical inverse problems. *SIAM J Sci Comput*. 2010; 32(5):2523–2542. <https://doi.org/10.1137/090775622>
32. Subramanian S, Scheufele K, Mehl M, Biros G. Where did the tumor start? An inverse solver with sparse localization for tumor growth models. *Inverse problems*. 2020; 36(4):045006. <https://doi.org/10.1088/1361-6420/ab649c> PMID: 33746330
33. Subramanian S, Scheufele K, Himthani N, Biros G. Multiatlas calibration of biophysical brain tumor growth models with mass effect. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer; 2020. p. 551–560.
34. Chvetsov AV, Sandison GA, Schwartz JL, Rengan R. Ill-posed problem and regularization in reconstruction of radiobiological parameters from serial tumor imaging data. *Phys Med Biol*. 2015; 60(21):8491–503. <https://doi.org/10.1088/0031-9155/60/21/8491> PMID: 26485348
35. Vilanova G, Colominas I, Gomez H. A mathematical model of tumour angiogenesis: growth, regression and regrowth. *J R Soc Interface*. 2017; 14. <https://doi.org/10.1098/rsif.2016.0918> PMID: 28100829
36. Vilanova G, Colominas I, Gomez H. Capillary networks in tumor angiogenesis: From discrete endothelial cells to phase-field averaged descriptions via isogeometric analysis. *International Journal for Numerical Methods in Biomedical Engineering*. 2013; 29(10). <https://doi.org/10.1002/cnm.2552> PMID: 23653256
37. Vilanova G, Colominas I, Gomez H. Computational Modeling of Tumor-Induced Angiogenesis. *Archives of Computational Methods in Engineering*. 2017; 24:1071–1102. <https://doi.org/10.1007/s11831-016-9199-7>
38. Xu J, Vilanova G, Gomez H. Phase-field model of vascular tumor growth: Three-dimensional geometry of the vascular network and integration with imaging data. *Computer Methods in Applied Mechanics and Engineering*. 2020; 359. <https://doi.org/10.1016/j.cma.2019.112648>
39. Travasso RDM, Poiré EC, Castro M, Rodriguez-Manzanique JC. Tumor Angiogenesis and Vascular Patterning: A Mathematical Model. *PLOS ONE*. 2011. <https://doi.org/10.1371/journal.pone.0019989> PMID: 21637756
40. Willcox KE, Ghattas O, Heimbach P. The imperative of physics-based modeling and inverse theory in computational science. *Nature Computational Science*. 2021; 1(3):166–168. <https://doi.org/10.1038/s43588-021-00040-z>
41. Yankeelov TE, An G, Saut G, Luebeck EG, Popel AS, Ribba B, et al. Multi-scale Modeling in Clinical Oncology: Opportunities and Barriers to Success. *Annals of Biomedical Engineering*. 2016; 44:2626–2641. <https://doi.org/10.1007/s10439-016-1691-6> PMID: 27384942
42. Lim B, Woodward WA, Wang X, Reuben JM, Ueno NT. Inflammatory breast cancer biology: the tumour microenvironment is key. *Nature Reviews Cancer*. 2018; 18:485–499. <https://doi.org/10.1038/s41568-018-0010-y> PMID: 29703913

43. Hachey SJ, Movsesyan S, Nguyen QH, Burton-Sojo G, Tankazyan A, Wu J, et al. An in vitro vascularized micro-tumor model of human colorectal cancer recapitulates in vivo responses to standard-of-care therapy. *Lab Chip*. 2021; 21:1333–1351. <https://doi.org/10.1039/d0lc01216e> PMID: 33605955
44. Hsu YH, Moya ML, Hughes CCW, George SC, Lee AP. A microfluidic platform for generating large-scale nearly identical human microphysiological vascularized tissue arrays. *Lab Chip*. 2013; 13:2990–2998. <https://doi.org/10.1039/c3lc50424g> PMID: 23723013
45. Wang X, Sun Q, Pei J. Microfluidic-Based 3D Engineered Microvascular Networks and Their Applications in Vascularized Microtumor Models. *Micromachines*. 2018; 9(10):493. <https://doi.org/10.3390/mi9100493> PMID: 30424426
46. Zhang W, Zhang YS, Bakht SM, Aleman J, Shin SR, Yue K, et al. Elastomeric free-form blood vessels for interconnecting organs on chip systems. *Lab Chip*. 2016; 16(9):1579–86. <https://doi.org/10.1039/c6lc00001k> PMID: 26999423
47. Khan OF, Sefton MV. Endothelial cell behaviour within a microfluidic mimic of the flow channels of a modular tissue engineered construct. *Biomed Microdevices*. 2012; 13(1):69–87. <https://doi.org/10.1007/s10544-010-9472-8>
48. Gadde M, Phillips C, Ghousifam N, Sorace AG, Wong E, Krishnamurthy S, et al. An In Vitro Vascularized Tumor Platform for Modeling Breast Tumor Stromal Interactions and Characterizing the Subsequent Response. *Biotechnology and Bioengineering*. 2020; 117(11).
49. Phillips CM, Lima EABF, Woodall RT, Brock A, Yankeelov TE. A hybrid model of tumor growth and angiogenesis: In silico experiments. *PLOS ONE*. 2020. <https://doi.org/10.1371/journal.pone.0231137> PMID: 32275674
50. Antoine EE, Vlachos PP, Rylander MN. Tunable Collagen I Hydrogels for Engineered Physiological Tissue Micro-Environments. *Plos One*. 2015. <https://doi.org/10.1371/journal.pone.0122500> PMID: 25822731
51. Buchanan CF, Verbridge SS, Vlachos PP, Rylander MN. Flow shear stress regulates endothelial barrier function and expression of angiogenic factors in a 3D microfluidic tumor vascular model. *Cell Adh Migr*. 2014; 8:517–524. <https://doi.org/10.4161/19336918.2014.970001> PMID: 25482628
52. Paszek MJ, Zahir N, Johnson KR, Lakins JN, Rozenberg GI, Gefen A, et al. Tensional homeostasis and the malignant phenotype. *Cancer Cell*. 2005; 8(3):241–254. <https://doi.org/10.1016/j.ccr.2005.08.010> PMID: 16169468
53. Michna R, Gadde M, Ozkan A, DeWitt M, Rylander M. Vascularized microfluidic platforms to mimic the tumor microenvironment. *Biotechnol Bioeng*. 2018. <https://doi.org/10.1002/bit.26778> PMID: 29940072
54. Gadde M, Marrinan D, Michna RJ, Rylander MN. Three Dimensional In Vitro Tumor Platforms for Cancer Discovery. In: Soker S, Skardal A, editors. *Tumor Organoids*. Springer International Publishing; 2018. p. 71–94.
55. Buchanan CF, Voigt EE, Szot CS, Freeman JW, Vlachos PP, Rylander MN. Three-dimensional microfluidic collagen hydrogels for investigating flow-mediated tumor-endothelial signaling and vascular organization. *Tissue Eng Part C Methods*. 2014; 20:64–75. <https://doi.org/10.1089/ten.tec.2012.0731> PMID: 23730946
56. Chen W, Sui L, Xu Z, Lang Y. Improved Zhang-Suen thinning algorithm in binary line drawing applications. In: 2012 International Conference on Systems and Informatics (ICSAI2012); 2012. p. 1947–1950.
57. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*. 2020; 17:261–272. <https://doi.org/10.1038/s41592-019-0686-2> PMID: 32015543
58. Farrell K, Oden JT, Faghihi D. A Bayesian framework for adaptive selection, calibration, and validation of coarse-grained models of atomistic systems. *Journal of Computational Physics*. 2015; 295:189–208. <https://doi.org/10.1016/j.jcp.2015.03.071>
59. Hawkins-Daarud A, Prudhomme S, van der Zee KG, Oden JT. Bayesian calibration, validation, and uncertainty quantification of diffuse interface models of tumor growth. *Journal of mathematical biology*. 2013; 67(6):1457–1485. <https://doi.org/10.1007/s00285-012-0595-9> PMID: 23053536
60. Jha PK, Cao L, Oden JT. Bayesian-based predictions of COVID-19 evolution in Texas using multispecies mixture-theoretic continuum models. *Computational Mechanics*. 2020; 66(5):1055–1068. <https://doi.org/10.1007/s00466-020-01889-z> PMID: 32836598
61. Lipková J, Angelikopoulos P, Wu S, Alverts E, Wiestler B, Diehl C, et al. Personalized radiotherapy design for glioblastoma: Integrating mathematical tumor models, multimodal scans, and bayesian inference. *IEEE transactions on medical imaging*. 2019; 38(8):1875–1884. <https://doi.org/10.1109/TMI.2019.2902044> PMID: 30835219
62. Lima EABF, Ghousifam N, Ozkan A, Oden JT, Shahmoradi A, Rylander MN, et al. Calibration of Multi-Parameter Models of Avascular Tumor Growth Using Time Resolved Microscopy Data. *Scientific Reports*. 2018; 8:14558. <https://doi.org/10.1038/s41598-018-32347-9> PMID: 30266911

63. Oden JT, Babuška I, Faghihi D. Predictive Computational Science: Computer Predictions in the Presence of Uncertainty. *Encyclopedia of Computational Mechanics*. 2017; 2.
64. Estacio-Hiroms KC, Prudencio EE, Malaya Np, Vohra M, McDougall D. The QUESO Library, User's Manual. *arXiv*. 2016;384. [arXiv:1611.07521](https://arxiv.org/abs/1611.07521).
65. Yang J, Virostko J, Hormuth DA II, Liu J, Brock A, Kowalski J, et al. An experimental-mathematical approach to predict tumor cell growth as a function of glucose availability in breast cancer cell lines. *Plos One*. 2021. <https://doi.org/10.1371/journal.pone.0240765> PMID: 34255770
66. Jarrett AM, Shah A, Bloom MJ, McKenna MT, Hormuth DA II, Yankeelov TE, et al. Experimentally-driven mathematical modeling to improve combination targeted and cytotoxic therapy for HER2+ breast cancer. *Scientific Reports*. 2019; 9:12830. <https://doi.org/10.1038/s41598-019-49073-5> PMID: 31492947
67. Xu J, Vilanova G, Gomez H. A Mathematical Model Coupling Tumor Growth and Angiogenesis. *PLoS One*. 2016; 11(2). <https://doi.org/10.1371/journal.pone.0149422> PMID: 26891163
68. Marin V, Kapllanski G, Grès S, Farnarier C, Bongrand P. Endothelial cell culture: protocol to obtain and cultivate human umbilical endothelial cells. *J Immunol Methods*. 2001; 254(1-2):183–90. [https://doi.org/10.1016/S0022-1759\(01\)00408-2](https://doi.org/10.1016/S0022-1759(01)00408-2) PMID: 11406163
69. Garipcan B, Maenz S, Pham T, Settmacher U, Jandt KD, Zanow J, et al. Image Analysis of Endothelial Microstructure and Endothelial Cell Dimensions of Human Arteries—A Preliminary Study. *Advanced Engineering Materials*. 2010; 13(1-2):B54–57.
70. Fritz M, Jha PK, Köppl T, Oden JT, Wagner A, et al. Modeling and simulation of vascular tumors embedded in evolving capillary networks. *Computer Methods in Applied Mechanics and Engineering*. 2021; 384:113975. <https://doi.org/10.1016/j.cma.2021.113975>
71. Wu C, Hormuth DA II, Easley T, Eijkhout V, Pineda F, Karczmar GS, et al. An *in silico* validation framework for quantitative DCE-MRI techniques based on a dynamic digital phantom. *Medical Image Analysis*. 2021; 73:102186. <https://doi.org/10.1016/j.media.2021.102186> PMID: 34329903
72. Liu J, Hormuth DA II, Davis T, Yang J, McKenna MT, Jarrett AM, et al. A time-resolved experimental-mathematical model for predicting the response of glioma cells to single-dose radiation therapy. *Integrative Biology*. 2021; 13(7):167–183. <https://doi.org/10.1093/intbio/zyab010> PMID: 34060613