Archives of Computational Methods in Engineering
https://doi.org/10.1007/s11831-025-10291-y

REVIEW

An Overview from Physically-Based to Data-Driven Approaches of the Modelling and Simulation of Glioblastoma Progression in Microfluidic Devices

Jacobo Ayensa-Jiménez[1,2,3] · Marina Pérez-Aliacar[2,3] · Mohamed H. Doweidar[3] · Eamonn A. Gaffney[4] · Manuel Doblaré[1,2,3]

Received: 24 July 2024 / Accepted: 29 April 2025

Abstract

In silico models and computational tools are invaluable instruments that complement experiments to improve our understanding of complex phenomena such as cancer evolution. This work offers a perspective on different approaches that can be used for mathematical modeling of glioblastoma, the most common and lethal brain cancer, in microfluidic devices, the most biomimetic in vitro cell culture technique nowadays. These approaches range from purely knowledge-based solutions to data-driven, and hence completely model-free, algorithms. In particular, we focus on hybrid approaches, which combine physically-based and data-driven strategies, demonstrating how this integration can enhance the understanding we get from simulation by revealing the underlying model structure and thus, in turn, the prospective biological mechanism.

1 Introduction

Glioblastoma (GBM), also known as grade IV astrocytoma, has the worst prognosis of any brain tumor [1], with a median survival of 14 months in patients that have undergone the standard treatment (maximal resection by surgery followed by radiotherapy with concomitant and adjuvant chemotherapy) [2]. Moreover, it accounts for 48% of all primary malignant central nervous system tumors [3], being the most common.

GBM progression is characterized by fast cell proliferation around blood vessels, eventually provoking their collapse and leading to local hypoxia. Subsequently, a necrotic core is formed around the occluded vessel and the surviving cells migrate towards more oxygenated regions [4, 5], where they restart the process of proliferation. These waves of migrating tumor cells, which typically surround the necrotic cores, are known as pseudopalisades [4]. The succession of cycles of local hypoxia triggering migration has been proposed as one of the main driving forces contributing to the aggressiveness and poor prognosis of GBM [6]. Indeed, severe hypoxia is a defining feature of the GBM Tumor MicroEnvironment (TME) but, in addition, other factors such as the extracellular matrix, the microvasculature or the immune cell population influence the heterogeneous nature of the TME and hence, in turn, GBM evolution [7].

Jacobo Ayensa-Jiménez and Marina Pérez-Aliacar have contributed equally to this work.

✉ Jacobo Ayensa-Jiménez
jacoboaj@unizar.es

Marina Pérez-Aliacar
marina.perez@unizar.es

Mohamed H. Doweidar
mohamed@unizar.es

Eamonn A. Gaffney
gaffney@maths.ox.ac.uk

Manuel Doblaré
mdoblare@unizar.es

[1] TME Lab, Institute for Health Research Aragón, Avenida San Juan Bosco, 13, 50009 Zaragoza, Spain

[2] TME Lab, Aragon Institute for Engineering Research, Calle Mariano Esquillor, sn, 50018 Zaragoza, Spain

[3] Department of Mechanical Engineering, School of Engineering and Architecture, University of Zaragoza, Calle María de Luna, 3, 50018 Zaragoza, Spain

[4] Wolfson Centre for Mathematical Biology, Wolfson Centre for Mathematical Biology, University of Oxford, Woodstock Road, OX2 6GG Oxford, UK

Published online: 10 July 2025

Understanding the complex interplay between GBM and its TME is crucial to design successful treatments for this tumor. Yet, animal models often fail to mimic human pathologies [8]. Traditional two-dimensional culture techniques also fail to recreate this complexity, which is inherently three-dimensional. Recently, microfluidics has emerged as a powerful 3D tool that enables the reproduction of TME features not previously observed in 2D cultures [9–11], taking promising steps towards the development of preclinical models that can be translated into the clinic to reduce reliance on animal models [8]. One of the most exploited applications of microfluidic devices has been precisely cell culture, since it provides a more realistic platform than the classical Petri dish for the study of processes that occur in the GBM microenvironment. Microfluidics has been applied to the study of many features of tumor dynamics in oncology, such as tumor-induced angiogenesis [12], tumor invasion [13], tumor intravasation [14], tumor extravasation [15] and metastasis [16]. In particular, for GBM, using microfluidic devices, it has been possible to reproduce the main histological features of the tumor in vitro, namely necrotic cores and pseudopalisades [17, 18].

Notwithstanding the advances made by microfluidic in vitro experiments, there are still some challenges that need to be addressed to characterize GBM progression. Aspects such as the isolation of the different phenomena or the quantification of the effect induced by each control parameter remain elusive, even under the most realistic and controlled experimental setups. Indeed, cultures in microfluidic devices can recreate complex and multi-physiological scenarios bringing the experiments closer to the biological reality. However, the presence of numerous complex, interacting behaviors hinders mechanism-based modeling. Furthermore, classifying outcomes under different scenarios or treatments even in vitro becomes prohibitive in terms of cost and labor due to the combinatorial explosion in the number of different conditions to test. To overcome these limitations, in vitro assays must be combined with in silico mathematical models and computational tools.

In the past years, there have been many attempts to build mathematical models to describe how GBM grows and responds to therapies, both in vitro and in vivo [19–24]. These models integrate the knowledge on the mechanisms governing GBM progression and translate them into mathematical formalisms, typically in the form of partial differential equations in the case of continuum models [21] or a set of rules in agent-based models [25]. However, defining a mathematical model requires prior knowledge of the functional relationships for the interactions of the different populations involved, and their responses to stimuli and interventions in the TME. These relations are not known in many cases, so prior assumptions have to be made, which, in most of the cases, become the main subject of the investigation. Indeed, the misspecification of hypotheses may generate flawed conclusions especially in untested scenarios [26], so the combination of proper validation strategies and uncertainty quantification (including parameter correlation) becomes essential.

Antithetical in the computational spectrum in terms of required knowledge, data science and artificial intelligence (AI) tools have arisen to overcome this limitation, enabling predictions for tumor evolution without *a priori* knowledge on the underlying phenomena. In the case of GBM, AI tools have been used in medical image processing for tumor segmentation and diagnosis [27, 28] and for prediction of treatment response [29, 30]. Nevertheless, there are some well-known drawbacks of pure AI approaches, such as their black-box nature and subsequent lack of explainability [31, 32] or the dependence on the dataset [33, 34] that may also lead to poor generalization.

Recently, a new paradigm is rising in Simulation-Based Engineering and Sciences, which aims to judiciously combine AI and in particular Machine Learning (ML) tools with physical information represented by mathematical models. This new discipline, coined as Scientific Machine Learning (SML) [35, 36] or Physics-Informed Machine Learning (PIML) [37], has proven to be successful in numerous frameworks and circumstances over the past decade [38–42], overcoming important challenges of conventional ML such as, interpretability [43], the need for enormous amounts of data [44] or the poor reliability and robustness of some predictions [45]. Indeed, the lack of explanatory capacity of ML may be supplied by centuries of research, which have resulted in high knowledge about our environment, thus whitening the black box character of AI, for instance in the study of GBM progression.

The aim of this review is to summarize several prospective and established approaches that have been exploited to understand, analyze and simulate GBM progression in microfluidic devices. As microfluidic technologies are high throughput screening platforms [46] generating extensive data, as well as supported by physical understanding, the study of GBM progression may range from conventional biophysical modeling, where a parametric model is specified, to pure ML approaches, where the unique information specified is the one provided by the experimental data. This is illustrated in Fig. 1.

The structure of the paper is as follows. Section 2 describes the general mathematical framework developed to simulate GBM progression in microfluidic devices, together with specific models found in the literature. In particular, the different equations, parameters, and interactions between the different functions and variables are presented, with accompanying analytical approximations to the solution of the resulting equations. Section 3 illustrates how the mathematical model can be fitted to data obtained from microfluidic device experiments.

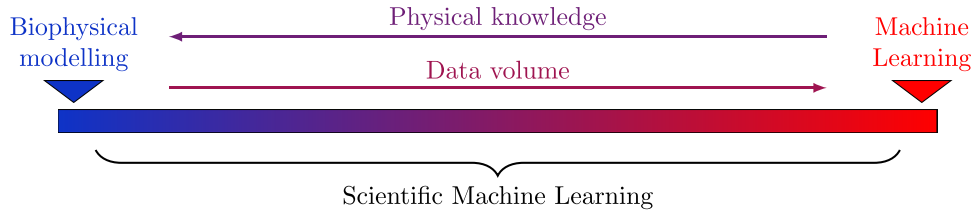


Fig. 1 Scientific machine learning. Depending on the use of data and physical knowledge, we move from conventional biophysical modeling to pure Machine Learning approaches

If the data are abundant enough, and the observations are stochastic, the adjusted model can include probabilistic elements for exploitation in the design of new experiments. Section 4 describes a ML approach, where a Neural Network is able to use microscopy images to infer the degree of aggressiveness of a particular population of GBM cells in in vitro cultures. Finally, in Sect. 5, a novel approach that combines both biophysical equations and elements of ML to identify hidden functional relations is presented. The article ends with a discussion of the different approaches outlined together with the main conclusions that can be extracted.

2 Describing Glioblastoma Evolution Using Partial Differential Equations

2.1 Mathematical Model

2.1.1 General Framework

We present in this section a general framework for modeling cell cultures in microfluidic devices. We define the equations controlling the dynamics of both the different cell phenotype populations and chemical substances present in the device. Dead cells are considered as a particular phenotype without capacity for proliferation, migration, or differentiation, but can eventually be transported by the fluid flow. Variables such as temperature, substrate stiffness and electric potential act as moderators of the different cell processes. We consider that the temperature and the electric potential are known fields, previously determined by solving thermal or electromagnetic problems, whereas stiffness, or in general, the parameters related with mechanical stimuli, are initially known but may be affected by the different cell processes. Therefore, given the known fields $\theta = \theta(\boldsymbol{x}, t)$ and $V = V(\boldsymbol{x}, t)$ representing the temperature and electric potential, our aim is to compute the evolution of $n + m + k$ fields, $C_i = C_i(\boldsymbol{x}, t)$, $i = 1, \dots, n$, representing the different cell phenotypes, $C'_i = C'_i(\boldsymbol{x}, t)$, $i = 1, \dots, m$, representing the chemical species and $p_i = p_i(\boldsymbol{x}, t)$, $i = 1, \dots, k$, representing the mechanical parameters.

Reaction-convection-diffusion equations are common in the modeling of biological population dynamics [47–50]. For a scalar field u, a quite general form of the transport equation is

$$\frac{\partial u}{\partial t} = \boldsymbol{\nabla} \cdot (D(u)\boldsymbol{\nabla} u) - \boldsymbol{\nabla} \cdot (\boldsymbol{v}(u, \boldsymbol{x}, t)u) + F(u), \quad \boldsymbol{x} \in \Omega, \tag{1}$$

where the first term of the righthand-side is the diffusion term, the second is the advection term and the third is the reaction term. The advection term may be due to the flow of a fluid convecting the scalar field u or to active migration mechanisms, such as chemotaxis or mechanotaxis, among others. When many scalar fields are considered, such as different cell populations and different chemical species, the terms D, $\boldsymbol{v}$ and F could, in principle, depend on the rest of the fields and their gradients. In particular, chemotaxis is usually modeled using the Keller–Segel framework [51], in all of its particular forms [52, 53].

Initial conditions and boundary conditions are required for these equations. The initial conditions are often, but not necessarily, that u is constant at the starting time. In general, one may state

$$u(\boldsymbol{x}, 0) = u_0(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega, \tag{2}$$

where $u_0 : \Omega \to \mathbb{R}$ is a known function. Regarding boundary conditions, we can define them in the boundary of the domain, $\partial\Omega = \Gamma_D \cup \Gamma_R$ with u satisfying in the Dirichlet part, Γ_D:

$$u(\boldsymbol{x}) = f(\boldsymbol{x}), \quad \boldsymbol{x} \in \Gamma_D, \tag{3a}$$

where $f : \Gamma_D \to \mathbb{R}$ is a known function, and in the Robin part, Γ_R:

$$\kappa(\boldsymbol{x})u + \frac{\partial u}{\partial \boldsymbol{n}} = g(\boldsymbol{x}), \quad \boldsymbol{x} \in \Gamma_R. \tag{3b}$$

with $\kappa, g : \Gamma_R \to \mathbb{R}$ also known functions.

Compatibility between the initial and boundary conditions is not strictly necessary, although it generates an artifact non-smoothness in the solutions that manifest in practice via extremely fast initial transients that are not physical, as observed in previous work [54]. We now consider how the framework of Eq. (1) can be used to consider both cell population concentrations and chemical species concentrations.

Cell population concentrations Particularizing Eq. (1) for the i-th cellular phenotype C_i we get:

$$\frac{\partial C_i}{\partial t} + (\boldsymbol{v}\cdot\boldsymbol{\nabla})C_i + \boldsymbol{\nabla}\cdot\boldsymbol{q}_i = C_iF_i - \sum_{\substack{j=1\\ j\neq i}}^{n} C_iF_{ij} + \sum_{\substack{j=1\\ j\neq i}}^{n} C_jF_{ji}, \quad \boldsymbol{x}\in\Omega, \tag{4}$$

where $\Omega \subset \mathbb{R}^d$ represents the domain of study (d is the spatial dimension considered), $\boldsymbol{v}$ is the fluid velocity (convection term), $\boldsymbol{q}_i$ is the flux vector associated with the diffusive and tactic migration of phenotype i, F_i is the source (reaction) term corresponding to the population growth, F_{ij} is the source term corresponding to phenotype switching, that is, the number of cells that differentiate from phenotype i to phenotype j per unit cell and time. All biological phenomena driving cell migration (different types of taxis) are modelled by means of expressions affecting the flux vector $\boldsymbol{q}_i$ as discussed below. For each phenotype i, we need to specify the partition $\Omega = \Gamma_{D_i} \cup \Gamma_{R_i}$ as well as the values of the functions $u_{0,i}$, f_i, g_i and κ_i, $i = 1, \dots, n$ for defining the initial and boundary conditions.

Species concentrations Similarly, the transport equation for the i-th chemical species $C'_i, i = 1, \dots, m$, including the reaction-convection-diffusion phenomena, writes

$$\frac{\partial C'_i}{\partial t} + (\boldsymbol{v}\cdot\boldsymbol{\nabla})C'_i + \boldsymbol{\nabla}\cdot\boldsymbol{q}'_i = \sum_{j=1}^{n} C_jF'_{ij}, \quad i = 1, \dots, m, \quad \boldsymbol{x}\in\Omega. \tag{5}$$

Again, $\boldsymbol{v}$ is the fluid velocity, $\boldsymbol{q}'_i$ is the flux vector associated with chemical species i and F'_{ij} the net source-term corresponding to production/consumption of species i per unit cell concentration of phenotype j. Chemical phenomena influencing species transport are again modelled using flux vectors. In an analagous manner, we need to provide the partition $\Omega = \Gamma_{D'_i} \cup \Gamma_{R'_i}$ as well as the values of the functions $u'_{0,i}$, f'_i, g'_i and κ'_i, $i = 1, \dots, m$ for the initial and boundary conditions.

Extracellular matrix remodeling

One last step in this global framework corresponds to regeneration/degradation of the ECM. Mechanical parameters of the substrate $p_1, \dots, p_k$ can be interpreted in general as evolving parameters, and thus their evolution is specified:

$$\dot{p}_i = R_i(\boldsymbol{C}, \boldsymbol{p}, \theta), \quad i = 1, \cdots, k. \tag{6}$$

The PDE system (4)–(5), together with the ODE system (6) and suitable initial and boundary conditions drive the dynamics for models of cell culture evolution in microfluidic devices under quite general conditions. However, to close the system of PDEs, we need to model the specific biological transport phenomena, that is, to make explicit the expressions for the source vectors F_{ij} and F'_{ij} as well as the relations between the flux vector fields $\boldsymbol{q}_i$ and $\boldsymbol{q}'_i$ on the one hand, and the rest of physical variables that promote or inhibit cell migration (and/or substance diffusion) on the other, such as temperature (thermotaxis), electric potential (electrotaxis), substrate stiffness (durotaxis), strain (tensotaxis), chemical species concentration (chemotaxis), or cell concentrations themselves (cell pedesis, proper diffusion by random walk).

Let us denote by $\boldsymbol{C}$, $\boldsymbol{C'}$, $\boldsymbol{p}$ the vectors representing cell phenotypes, chemical species and mechanical parameters respectively, that is, $\boldsymbol{C}(\boldsymbol{x}, t) = (C_1, \cdots, C_n)^\intercal$, $\boldsymbol{C'}(\boldsymbol{x}, t) = (C'_1, \cdots, C'_m)^\intercal$ and $\boldsymbol{p}(\boldsymbol{x}, t) = (p_1, \cdots, p_k)^\intercal$. First, we consider the source-terms in Eqs. (4) and (5):

- **Proliferation**: The proliferation of the i-th phenotype in Eq. (4) can be expressed as:

  $$F_i = F_i(\boldsymbol{C}, \boldsymbol{C'}, \boldsymbol{p}, \theta), \quad i = 1, \dots, n,$$

  where we include the influence of features such as nutrient concentration, mechanical stimuli, growth factors and temperature. Some examples of common growth models used for cancer modeling are exponential or Matlthusian growth [55], logistic or Verhulstian growth [56], Gompertz growth [57] or its generalizations, Weibull and Frechet models [58], Richards model [59], which is a generalization of the logistic model, or the general family of hyperbolastic models [60]. A brief comparison between the general features of some of these models may be found in Ref. [61].
- **Differentiation:** The differentiation of the i-th phenotype to the j-th phenotype in Eq. (4) can be modeled using a similar equation:

  $$F_{ij} = F_{ij}(\boldsymbol{C}, \boldsymbol{C'}, \boldsymbol{p}, \theta), \quad i = 1, \dots, n.$$

  This very general expression is usually simplified to

  $$F_{ij}(\boldsymbol{C}, \boldsymbol{C'}, \boldsymbol{p}, \theta) = \frac{1}{\tau}\Phi(h_{\text{phys}}(\boldsymbol{C}, \boldsymbol{C'}, \boldsymbol{p}, \theta)),$$

  where τ is a characteristic time, Φ an activation function, for example, the Heaviside function ($\Phi(x) = 0$ if $x \leq 0$ and $\Phi(x) = 1$ if $x > 0$), or the sigmoid function ($\Phi(x) = \frac{1}{2}(1 + \tanh(x))$), and h_{phys} defines the domain of physiological behavior of a cell, that is, if $h_{\text{phys}}(\boldsymbol{C}, \boldsymbol{C'}, \boldsymbol{p}, \theta) \leq 0$, the cell is in its physiological state.
- **Consumption and production:** Source-terms in Eq. (5) include generation and uptake of species. The net production/consumption of the i-th chemical species by the j-th cell phenotype can be described as:

  $$F'_{ij} = F'_{ij}(\boldsymbol{C}, \boldsymbol{C'}, \theta), \quad i = 1, \dots, m, \quad j = 1, \dots, n.$$

With respect to migration-terms (biologically related transport) in Eqs. (4) and (5) the flux vector $\boldsymbol{q}_i$ is postulated as a linear decomposition:

$$\boldsymbol{q}_i = \boldsymbol{q}_{\mathrm{D},i} + \boldsymbol{q}_{\mathrm{M},i} + \boldsymbol{q}_{\mathrm{E},i} + \boldsymbol{q}_{\mathrm{T},i} + \boldsymbol{q}_{\mathrm{C},i}, \quad i = 1, \ldots, n.$$

where the flux vectors $\boldsymbol{q}_{\mathrm{D},i}$, $\boldsymbol{q}_{\mathrm{M},i}$, $\boldsymbol{q}_{\mathrm{E},i}$, $\boldsymbol{q}_{T,i}$ and $\boldsymbol{q}_{C,i}$ are associated with diffusion, mechanotaxis (durotaxis and tensotaxis), electrotaxis, thermotaxis and chemotaxis phenomena, respectively. Regarding the species flux $\boldsymbol{q}_i$, the species are considered to be inert, so the only flux term considered is diffusion.

- **Random pedesis:** Modeled by means of a Fickean diffusion model:

  $$\boldsymbol{q}_{\mathrm{D},i} = -\boldsymbol{K}_{\mathrm{D},i}\boldsymbol{\nabla} C_i, \quad i = 1, \ldots, n,$$

  where matrix $\boldsymbol{K}_{D,i} = \boldsymbol{K}_{\mathrm{D},i}(\boldsymbol{C}, \boldsymbol{C}', \boldsymbol{p}, \theta)$ is the diffusion matrix.
- **Mechanotaxis:** This term can be expressed in a similar manner, but using the mechanical parameters as controllers or drivers of cell migration:

  $$\boldsymbol{q}_{\mathrm{M},i} = \sum_{j=1}^{k} \boldsymbol{K}_{\mathrm{M},i,j}\boldsymbol{\nabla} p_j, \quad i = 1, \ldots, n,$$

  where $\boldsymbol{K}_{\mathrm{M},i,j} = \boldsymbol{K}_{\mathrm{M},i,j}(\boldsymbol{C}, \boldsymbol{C}', \boldsymbol{p}, \boldsymbol{\nabla p}, \theta)$ is the mechanotactic motility matrix of cell phenotype i with respect to mechanical parameter p_j. Generalization to consider gradients in mechanical tensors, such as stress, can be considered analogously.
- **Electrotaxis and thermotaxis:** They can be modeled as:

  $$\boldsymbol{q}_{\mathrm{E},i} = \boldsymbol{K}_{\mathrm{E},i}\boldsymbol{\nabla} V, \quad i = 1, \ldots, n,$$
  $$\boldsymbol{q}_{\mathrm{T},i} = \boldsymbol{K}_{\mathrm{T},i}\boldsymbol{\nabla}\theta, \quad i = 1, \ldots, n,$$

  where $\boldsymbol{K}_{\mathrm{E},i} = \boldsymbol{K}_{\mathrm{E},i}(\boldsymbol{C}, V, \boldsymbol{\nabla} V, \boldsymbol{p}, \theta)$ is the electrostatic motility matrix and $\boldsymbol{K}_{T,i} = \boldsymbol{K}_{\mathrm{T},i}(\boldsymbol{C}, \boldsymbol{p})$ is the thermotactic motility matrix.
- **Chemotaxis:** The chemotaxis term is expressed analogously

  $$\boldsymbol{q}_{\mathrm{C},i} = \sum_{j=1}^{m} \boldsymbol{K}_{\mathrm{C},i,j}\boldsymbol{\nabla} C'_j, \quad i = 1, \ldots, n,$$

  where $\boldsymbol{K}_{\mathrm{C},i,j} = \boldsymbol{K}_{\mathrm{C},i,j}(\boldsymbol{C}, \boldsymbol{C}', \boldsymbol{\nabla C}', \boldsymbol{p}, \theta)$ is the chemotactic motility matrix with respect to the chemical species j. It is common to use saturation models [23].
- **Diffusion of species:** The diffusion is expressed as:

  $$\boldsymbol{q}'_i = -\boldsymbol{K}'_{\mathrm{D},i}\boldsymbol{\nabla} C'_i, \quad i = 1, \ldots, m,$$

  with $\boldsymbol{K}'_{\mathrm{D},i} = \boldsymbol{K}'_{\mathrm{D},i}(\boldsymbol{C}, \theta, \boldsymbol{p})$ the diffusion matrix.

2.1.2 Some Particular Models for Cell Cultures in Microfluidic Devices

The following assumptions are usually made for cell cultures in microfluidic devices:

- **Negligible remodeling:** Cell culture experiments typically last some days, which is too short a timescale for tissue remodeling. Therefore, in these situations, it is possible to neglect the variation of mechanical parameters, that is $\dot{p}_i = 0, \forall i = 1, \ldots, k$.
- **Negligible thermal and electrical effects:** If temperature and electrical fields are not control variables in the experiment, $\theta(\boldsymbol{x}, t) = \theta_0$, with θ_0 the room temperature, and $V(\boldsymbol{x}, t) = V_0$. Therefore, $\boldsymbol{\nabla}\theta = \boldsymbol{\nabla} V = \boldsymbol{0}$.
- **Substrate homogeneity:** If the culture hydrogel is homogeneous from the macroscopic point of view, we also have $\boldsymbol{\nabla} p_i = \boldsymbol{0}, \forall i = 1, \ldots, k$.
- **Substrate isotropy:** In most cases, isotropy also holds in the cell culture so all cell motility matrices may be considered as a scaling of the identity, $\boldsymbol{K} = K\boldsymbol{I}$.
- **No fluid flow:** Finally, usually cell cultures are seeded and cultured under static or negligible flow conditions, so that $\boldsymbol{v} = \boldsymbol{0}$.

After these particular hypotheses, Eqs. (4) and (5) reduce to

$$\frac{\partial C_i}{\partial t} = \boldsymbol{\nabla} \cdot \Big(K_{D,i}\boldsymbol{\nabla} C_i - \sum_{j=1}^{m} K_{C,i,j}\boldsymbol{\nabla} C'_j \Big) + C_i F_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} \big(C_i F_{ij} - C_j F_{ji} \big), \tag{7a}$$

$$\frac{\partial C'_i}{\partial t} = \boldsymbol{\nabla} \cdot \Big(K'_{D,i}\boldsymbol{\nabla} C'_i \Big) + \sum_{j=1}^{n} C_j F'_{ij}. \tag{7b}$$

Although, in recent years, some models describing GBM evolution have been developed (see Ref. [62] for a comprehensive review), only a few of them tackle GBM evolution in microfluidic devices, even if this platform is considered one of the most biomimetic culture techniques [63, 64].

Martínez-González et al. [65] developed a model with two phenotypes for live cells: normoxic and hypoxic, as well as dead cells and oxygen as species. Cells become hypoxic when the oxygen concentration drops below a certain threshold and return to the normoxic threshold when it rises above the threshold again. Transitions between phenotypes are mediated by step-like activation functions with a characteristic time. Normoxic cells are assumed to be, based on the go or grow hypothesis [66], less migratory and more proliferative than hypoxic ones. Normoxic cells also consume

more oxygen, while hypoxic cells are the only ones that can become necrotic.

Later, Ayuso et al. [17] added an extra phenotype to the previously presented model, accounting for hypoxic cells that arrive to normoxic areas and differentiate in the so-called malignant phenotype, with a higher proliferation rate (and also a higher oxygen consumption rate) than normoxic cells.

Another possibility is to consider uniquely two cell phenotypes (live and dead cells), together with oxygen as a species, but to incorporate the metabolic switch between proliferative and migratory activity via two activation functions [67, 68]. Since this third model [67] will be the basis for later developments in this review, we further detail its equations and nomenclature. Specifically, the governing equations for normoxic C_1 and necrotic C_2 cell concentrations and oxygen concentration $S = C_1'$ are expressed as:

$$\frac{\partial C_1}{\partial t} = \nabla \cdot \left(K_{D,1}\nabla C_1 - K_{C,1,1}\nabla S\right) + C_1F_1 - C_1F_{12} + C_2F_{21}, \tag{8a}$$

$$\frac{\partial C_2}{\partial t} = \nabla \cdot \left(K_{D,2}\nabla C_2 - K_{C,2,1}\nabla S\right) + C_2F_2 - C_2F_{21} + C_1F_{12}, \tag{8b}$$

$$\frac{\partial S}{\partial t} = \nabla \cdot \left(K'_{D,1}\nabla S\right) + C_1F'_{11} + C_2F'_{21}. \tag{8c}$$

Also, some biologically-based assumptions are made [67]:

- Living cells follow the *go-or-grow* paradigm, that is, when the oxygen level is low (hypoxia) its behavior is mainly migratory and when the oxygen level is high (normoxia) its behavior is mainly proliferative [66].
- Both cell growth and migration are constrained to the space availability.
- The two main mechanisms causing living cell death are apoptosis (programmed death to control cell development and/or growth) and necrosis (uncontrolled death that occurs when cells are exposed to extreme variations in physiological conditions). In our case, necrosis will occur due to extremely low levels of oxygen (anoxia) so oxygen is the only variable mediating the overall cell death.
- Dead cells are considered as an inert population so they do not move, proliferate, consume oxygen or disappear throughout the duration of the experiment.
- Oxygen changes are due only to consumption by the normoxic population (cellular respiration) and oxygen supply in the microdevice, typically via the lateral channels.

With these considerations, we have

$$K_{D,1} = P, \quad K_{C,1,1} = \chi\Pi_{go}(S)F_{go}(C_1, C_2)C_1, \quad F_1 = \frac{1}{\tau_{gr}}\Pi_{gr}(S)F_{gr}(C_1, C_2)C_1,$$
$$F_{12} = \frac{1}{\tau_d}\Pi_d(S), \quad F_{21} = 0, \quad K_{D,2} = 0,$$
$$K_{C,2,1} = 0, \quad F_2 = 0, \quad K'_{D,1} = D,$$
$$F'_{11} = \alpha\Pi_c(S), \quad F'_{21} = 0.$$

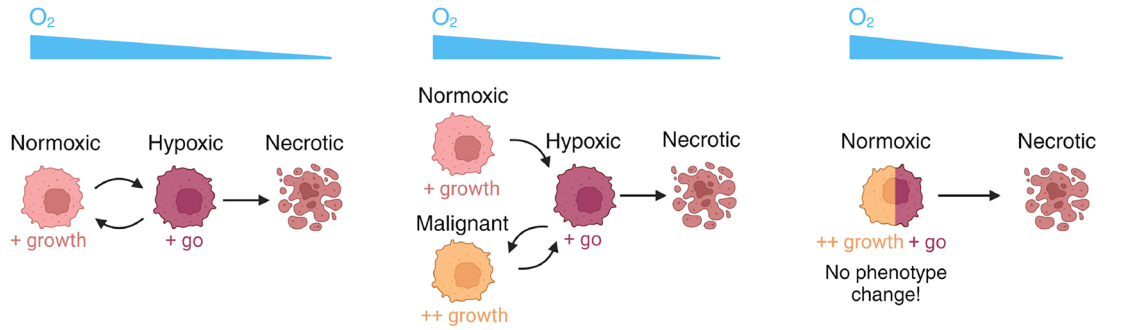
Here, the functions Π_{go}, Π_{gr}, Π_d, and Π_c are functions that model the metabolic cell activity in terms of the oxygen level, while the functions F_{go} and F_{gr} take into account the changes in the behavior of cells due to the availability of space or other nutrients (that is, migration and growth saturation). All must be phenomenologically defined in order to close the model. Observe that GBM cell progression is considered here to be controlled only by the oxygen concentration. This assumption comes from previous experiments that showed that other nutrients such as glucose do not produce relevant changes in cell response on the timescale of the observation [18].

2.1.3 Model Comparison. Advantages and Disadvantages of the Described Models

A schematic representation of the main features of the three presented models, in terms of the phenotypes they consider and the relationships between them, is shown in Fig. 2.

The models described in the previous section and represented in Fig. 2, although describing the same phenomena (GBM evolution and interaction with oxygen within microfluidic devices), present different levels of complexity. In general, increasing the number of phenotypes increases the number of equations and hence the number parameters, yielding a more complex model. The 2-phenotype model has 11 parameters that would need to be calibrated, the 3-phenotype model has 14 parameters, and the 4-phenotype model has 19 parameters, making the latter more difficult to calibrate as well as more prone to identifiability issues and overfitting.

With respect to computational cost, the models have one PDE per phenotype, so increasing the number of phenotypes also entails a higher computer cost, which is negligible in the 1D case but may become a limitation in higher dimensions. However, it has to be pointed out that the 4-phenotype model can capture in some sense cellular adaptation to hypoxia due to the presence of an irreversible transition from normoxic to hypoxic phenotypes. Both the 2 and 3-phenotype models have reversible transitions between hypoxia and normoxia, but in the 3-phenotype model this transition is discrete while the 2-phenotype model allows for a range of intermediate states which better represent the underlying biological reality.



(a) 3-phenotype model [65]. (b) 4-phenotype model [17]. (c) 2-phenotype model [67].

Fig. 2 Schematic representation of the different models of glioblastoma evolution in microfluidic devices. The different subfigures represent the phenotype structure of the models presented in the section, together with the main characteristics of each feature and the transitions between them. Created with BioRender.com

Finally, to assess the influence of the model parameters on the output as well as the models' robustness, a global sensitivity analysis was performed using the Morris Method [69]. The complete methodology and results can be accessed in Appendix A. As the most noteworthy aspect, it can be seen that the 2-phenotype model is the most insensitive to parameter variations, despite being the most nonlinear, which entails that it has the most correlations between its parameters.

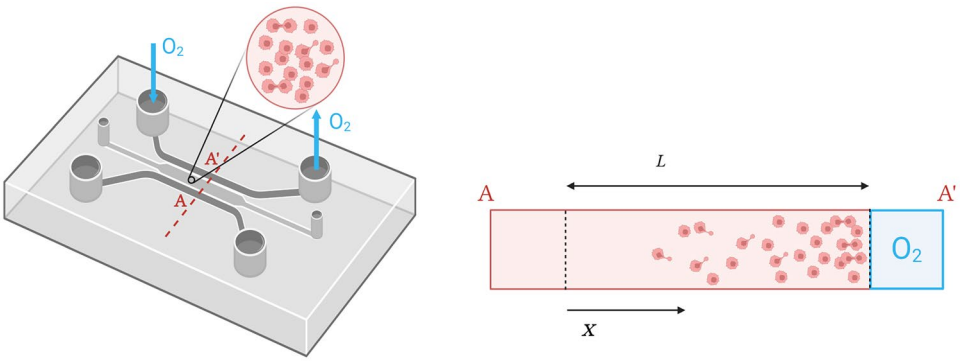
2.2 Analytical Solutions for Some Particular Situations

Solving Eqs. (4), (5) and (6) numerically can be computationally intensive and challenging. Particularly, the task is frequently prohibitive for inverse problems, parameter fitting, model selection, Design of Experiments (DoE), sensitivity analyses, model structural analysis, and Uncertainty Quantification (UQ). Hence, having analytical solutions, even if they are approximate, is invaluable. These solutions serve also as essential tools for testing and validating numerical algorithms, enhancing our understanding of the mechanisms across the parameter space, and enabling preliminary predictions of the behaviors of interest. For instance, they help predicting phenomena like traveling fronts and equilibria, in particular to study GBM progression in certain regimes [70–72].

Let us consider GBM cell progression under hypoxic conditions with oxygen acting as a regulator of cell processes, but in limited supply [67, 68]. In many experimental situations, the geometry of the problem is effectively one-dimensional [17, 54, 67], as illustrated in Fig. 3, with the oxygen gradient between the two lateral channels kept constant and known, $\Delta S = S_{\max} - S_{\min}$. Therefore, Eqs. (8a) and (8c) are expressed as

$$\frac{\partial C_1}{\partial t} = \frac{\partial}{\partial x}\left(P\frac{\partial C_1}{\partial x} - \chi\Pi_{\text{go}}(S)C_1\frac{\partial S}{\partial x}\right) + \frac{1}{\tau_{\text{gr}}}\Pi_{\text{gr}}(S)C_1\left(1 - \frac{C_1}{C_{\text{sat}}}\right) - \frac{1}{\tau_{\text{d}}}\Pi_{\text{d}}(S)C_1, \tag{9a}$$

$$\frac{\partial S}{\partial t} = \frac{\partial}{\partial x}\left(D\frac{\partial S}{\partial x}\right) - \alpha\Pi_{\text{c}}(S)C_1, \tag{9b}$$



(a) Experimental set-up. (b) Detail of A-A' section.

Fig. 3 Schematic representation of an experimental set-up, that can be assumed as one-dimensional. Cells are seeded in the central chamber, while medium with a controlled oxygen concentration is perfused through the two lateral channels. Cells migrate towards oxygenated areas due to the hypoxia-mediated *go-or-grow* behavior. Created with BioRender.com

where we have considered $F_{go}(C_1, C_2) = 1$ and $F_{gr}(C_1, C_2) = 1 - \frac{C_1}{C_{sat}}$, that is, a saturation effect for growth but not for chemotaxis.

When there is no possible cell migration out of the central chamber, the boundary conditions are expressed as

$$f_1(x = 0, t) = 0, \tag{10a}$$

$$f_1(x = L, t) = 0, \tag{10b}$$

$$S(x = 0, t) = S_{min}, \tag{10c}$$

$$S(x = L, t) = S_{max}, \tag{10d}$$

where, without loss of generality we have assumed that the minimal oxygen concentration is at $x = 0$ and the maximal at $x = L$ and $f_1 = P\frac{\partial C_1}{\partial x} - \chi\Pi_{go}(S)C_1\frac{\partial S}{\partial x}$ is the live cell flux.

After appropriate non-dimensionalization [54], Eqs. (9) become

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\pi_1\frac{\partial u}{\partial x} - \pi_2\pi_{go}(v)u\frac{\partial v}{\partial x}\right) + \pi_{gr}(v)u(1-u) - \pi_3\pi_d(v)u, \tag{11a}$$

$$\frac{\partial v}{\partial t} = \frac{\partial}{\partial x}\left(\pi_4\frac{\partial v}{\partial x}\right) + \pi_5\pi_c(v)u, \tag{11b}$$

with

$$\pi_1 = \frac{P\tau_{gr}}{L^2}, \quad \pi_2 = \frac{\chi\Delta S\tau_{gr}}{L^2}, \quad \pi_3 = \frac{\tau_{gr}}{\tau_d},$$
$$\pi_4 = \frac{D\tau_{gr}}{L^2}, \quad \pi_5 = \frac{\alpha\tau_{gr}C_{sat}}{\Delta S}, \quad \pi_*(v) = \Pi_*\big(S_{min} + (S_{max} - S_{min})v\big),$$

where with the symbol $*$ we refer indistinctly to $*=$ go, gr, d, c.

A fundamental question is to explore whether it is possible to find analytical solutions to Eqs. (11) under some specific considerations. This problem has been explored in Ref. [54] for the case in which $\pi_1, \pi_3 \ll 1$ (low cell random motility and negligible cell death) and $\pi_4, \pi_5 \gg 1$ (fast oxygen dynamics relative to cellular processes). Neglecting any very fast initial transients, the analytical expression for the dimensionless cell density, when it exhibits chemotactically-driven wave-like behavior, is given by

$$u(x,t) = \begin{cases} u^*(x^*(t),t)\left(\mathrm{erf}\left(\frac{x-x^*(t)}{2\sqrt{t\pi_1}}\right) + 1\right) & \text{if } x \le x^*(t), \\ u^*(x^*(t),t)\left(\mathrm{erf}\left(\frac{x-x^*(t)}{2\sqrt{t\pi_1}}\right) - 1\right) + u^*(x,t) & \text{if } x \ge x^*(t). \end{cases} \tag{12}$$

with erf the error function, $x^*(t)$ the position of the wave-front moving from the low-oxygenated to the high-oxygenated boundary and $u^*(x,t)$ the expression of the dimensional cell profile far from the transition. The particular expression of the functions $x^* = x^*(t)$ and $u^* = u^*(x,t)$ depends on the structure of the functions $\pi_{go}(v)$, $\pi_{gr}(v)$ and $\psi(u,v) = \pi_c(v)u$, and has been established for some particular cases (see Refs. [54, 68] for the additional details), which are summarized

Table 1 Position of the transition and cell profile far from the transition

Migration activation $\pi_{go}(v)$	Growth activation $\pi_{gr}(v)$	Consumption term $\psi(u,v)$	Position of the transition $x^*(t)$	Cell profile far from the transition $u^*(x,t)$
1	1	0	$\pi_2 t$	$\frac{u_0 e^t}{1+u_0(e^t-1)}$
$1-v$	v	0	$1 - e^{-\pi_2 t}$	$\frac{u_0 e^{a(x,t)}}{1+u_0\left[e^{a(x,t)}-1-\pi_2 b(x,t)\right]}$
1	1	λ	$\frac{2-\lambda}{2\lambda}\left(e^{\pi_2\lambda t} - 1\right)$	$\frac{u_0 e^{(1-\pi_2\lambda)t}}{1+\frac{u_0}{1-\pi_2\lambda}\left(e^{(1-\pi_2\lambda)t}-1\right)}$
1	1	λv	$\frac{1}{\sqrt{\lambda}}\ln\left[\frac{1+\tan\left(\frac{\pi_2\lambda t}{2}\right)}{1-\tan\left(\frac{\pi_2\lambda t}{2}\right)}\right]$	$\frac{u_0 e^{c(x)t}}{1+\frac{u_0}{c(x)}\left(e^{c(x)t}-1\right)}$

The results are presented for different activation functions related to migration, proliferation and consumption

The functions a, b and c are defined as $a(x,t) = (1+\pi_2)t + \frac{1-e^{\pi_2 t}}{\pi_2}(1-x)$,

$b(x,t) = \exp\left(-\frac{e^{\pi_2 t}}{\pi_2}(1-x)\right)\int_0^t \exp\left[(\pi_2+1)u + \frac{e^{\pi_2 t}}{\pi_2}(1-x)e^{-\pi_2 u}\right]\mathrm{d}u$,

and

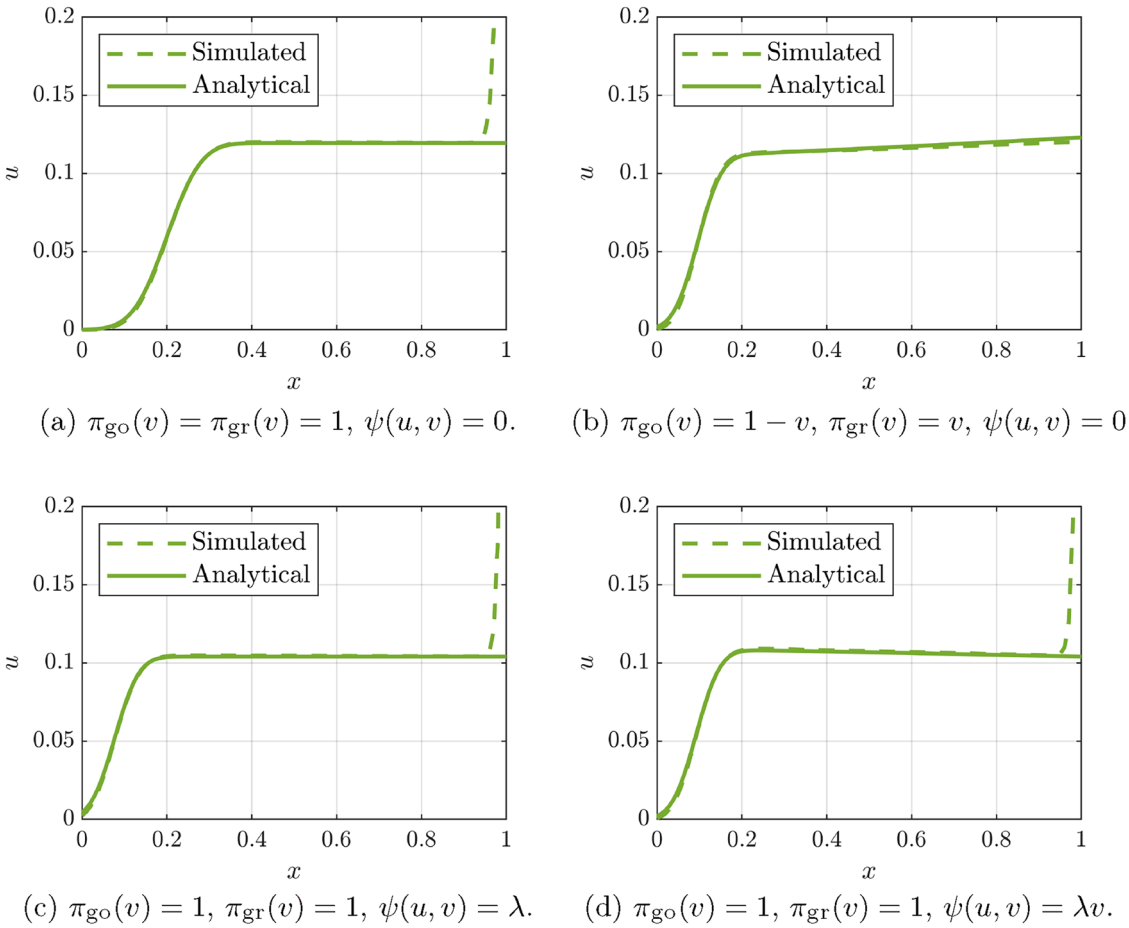
$c(x) = 1 - \frac{\pi_2\sinh(\sqrt{\lambda}x)}{\sinh(\sqrt{\lambda})}$.

Note that for $u_0 \ll 1$, in all the cases the cell profile after the front wave has a (sometimes heterogeneous) exponential growth

in Table 1, for an homogeneous initial cell profile. Some different biological scenarios leading to these expressions are the following:

- There is a clear interest on the scientific community in whether GBM follows a *go-and-grow* or a *go-or-grow* behavior [73]. A simple although enlightening approach for this analysis is to consider $\pi_{\mathrm{go}}(v) = \pi_{\mathrm{go}}(v) = 1$ for the *go-and-grow* behavior and $\pi_{\mathrm{go}}(v) = 1 - v$, $\pi_{\mathrm{go}}(v) = 1 - \pi_{\mathrm{go}}(v) = v$ for the *go-or-grow*.
- Suppose that there are no other cells in the microfluidic device besides the cell culture of our interest and we are at the low cell regime (far from cell saturation capacity), that is $u_0 = u(x, t = 0) \ll 1$. In that case, after rapid initial transients describing the diffusion relaxation of oxygen, and too fast to be on the timescale of cellular motility, in Eq. (11b) we can assume $\pi_{\mathrm{c}}(v)u \sim 0$ so that $\psi(u, v) = 0$.
- Suppose that there are other non-migrating cells within the microfluidic device in addition to the migrating cells, for instance for a metastasis model, with the other cells at constant concentration and in excess of the tumor cells. Additionally, we assume that we have high oxygen supplies. Under these considerations, it is possible to state $\psi(u, v) = \lambda$ [54].
- Suppose again that there are other non-migrating cells at a constant concentration but with low oxygen concentrations. For many consumption kinetic models, such as the Michaelis-Menten model $\pi_{\mathrm{c}}(v) = \frac{v}{v+\kappa}$ or the exponential model $\pi_{\mathrm{c}}(v) = 1 - \mathrm{e}^{-\kappa v}$, it is possible to assume that $\pi_{\mathrm{c}}(v) \propto v$ when $v \sim 0$ so that in that case $\psi(u, v) = \lambda v$ [54].

The accuracy of the different analytical solutions is illustrated in Fig. 4, where they are compared to numerical solutions using Matlab `pdepe` routine, which numerically approximates the solution of initial-boundary value problems for systems of parabolic and elliptic partial differential equations in one space variable and time. Note that the full model is parabolic since $\pi_1 > 0$. This routine uses



(a) $\pi_{\mathrm{go}}(v) = \pi_{\mathrm{gr}}(v) = 1$, $\psi(u, v) = 0$.

(b) $\pi_{\mathrm{go}}(v) = 1 - v$, $\pi_{\mathrm{gr}}(v) = v$, $\psi(u, v) = 0$.

(c) $\pi_{\mathrm{go}}(v) = 1$, $\pi_{\mathrm{gr}}(v) = 1$, $\psi(u, v) = \lambda$.

(d) $\pi_{\mathrm{go}}(v) = 1$, $\pi_{\mathrm{gr}}(v) = 1$, $\psi(u, v) = \lambda v$.

Fig. 4 Analytical and numerical solutions obtained for the field $u = u(x, t)$ for four different cases. The different activation functions considered are the ones presented in Table 1 and the values selected for the parameters are $\pi_1 = \pi_3 = 0.01$, $\pi_2 = 1$ and $\pi_4 = \pi_5 = 100$. An initial (dimensionless) cell profile of $u(x, 0) = u_0 = 0.1$ is considered and the (dimensionless) cell profile is shown for $t = 0.1$

a time-space integrator based on a piece-wise nonlinear Galerkin approach which is second-order accurate in space [74]. For illustration purposes, we consider $\pi_1 = \pi_3 = 0.01$, $\pi_2 = 1$, $\pi_4 = \pi_5 = 100$ and $\lambda = 0.5$, and the initial profile is set to $u(x, t = 0) = u_0 = 0.1$. The degree of approximation is evaluated at $t = 0.1$ and good agreement can be observed, except in the area close to $x = 1$ where a transition region (boundary layer) develops. This is because the term u_{xx} cannot be neglected, despite the fact that $\pi_2 \sim 0$, due to the sharp variations of the cell concentration that are induced by the boundary conditions at $x = 1$ [54, 75]. The analytical solutions can be derived for situations with heterogeneous initial cell profiles, $u_0 = u_0(x)$ without too much effort [54], although here we present homogeneous initial cell profiles for illustration purposes.

2.3 Numerical Results

Even with the results derived in the previous section, in most problems it is not possible to obtain analytical solutions. Hence, we must resort to numerical solutions, via suitable discretization and resolution of the mathematical models. A standard approach for solving time and spatially dependent PDEs systems used here is the combination of the Finite Element Method (FEM) for the spatial discretization and explicit Runge–Kutta time integrators [68].

To illustrate the simulations that can be performed, we present here the results of a 3D "toy" model of cell culture evolution in a microfluidic device [68]. The aim of such in vitro experiments is to reproduce, within the device, the in vivo TME. This includes mechanical properties, nutrient and oxygen supply.

A simulation in a 3D chip-like geometry is performed for the simplified model with two phenotypes presented in Sect. 2.1.2, Eqs. (8), with the specific functions

$$K_{D,1} = P, \qquad K_{C,1,1} = \chi C_1, \qquad F_1 = \frac{1}{\tau_{\mathrm{gr}}}\left(1 - \frac{C_1}{C_{\mathrm{sat}}}\right),$$
$$F_{12} = \frac{1}{\tau_{\mathrm{d}}}\Pi_{\mathrm{d}}(S), \qquad F_{21} = 0, \qquad K_{D,2} = 0,$$
$$K_{D,2,1} = 0, \qquad F_2 = 0, \qquad K'_{D,1} = D,$$
$$F'_{11} = \Pi_{\mathrm{c}}(S), \qquad F'_{21} = 0.$$

where $\Pi_{\mathrm{d}}(S) = 1 - H(S - S_A)$, with H the Heaviside step function, and $\Pi_{\mathrm{c}}(S) = S/(S + k_{\mathrm{m}})$. The problem equations were solved using an in-house Matlab FEM code. An exhaustive code verification as well as the particular values of the model parameters P, χ, τ_{gr}, C_{sat}, τ_{d}, S_A and k_{m}, even if they are not relevant for the purposes of this illustrative simulation, can be found in Ref. [68]. The meshed geometry of the central chamber of the device, where cells are seeded, is shown in Fig. 5a. The geometry presented has a respective maximum width, length, and height of 600 μm, 600 μm, and 70 μm, respectively. It can be seen that in the lateral walls of the microfluidic devices there are some intrusions that correspond to solid pillars. Figure 5b shows the boundary conditions imposed for the simulations. Dirichlet boundary conditions are imposed in the areas where the central chamber is in direct contact with the lateral channels, while Neumann boundary conditions are imposed in the rest of the surface. Illustrative results are presented in Fig. 5c and d where the cell concentration fields in the device can be observed. Indeed, it can be seen that the concentration of live cells is higher close to the lateral channels whereas cell death has occurred predominantly in the central area of the chamber (necrotic core).

3 Calibrating Parametric Models from Cell Culture Data

We focus in this section on one of the parametric mathematical models presented above, which will be used as a benchmark to evaluate different calibration approaches. Also, we focus on data obtained from microfluidic devices whose geometry can be identified with the one illustrated in Fig. 3, which allows a restriction to one dimensional problems.

3.1 Data and Model Description

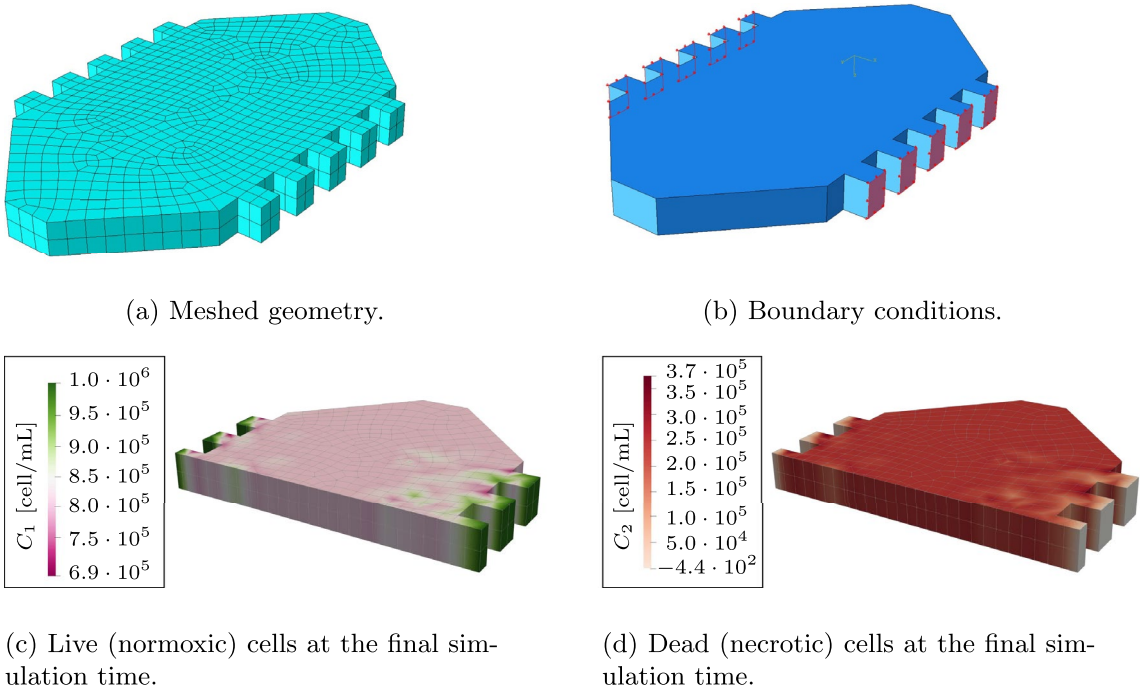
3.1.1 Experimental Data

To reproduce the hypoxic TME of GBM, oxygen impermeable microfluidic devices consisting of a central chamber and two lateral microchannels have been used [17, 18].

Laser confocal and fluorescence images were acquired at different focal planes using a Nikon Eclipse Ti-E C1 confocal microscope. Images were analyzed using Fiji software (`http://fiji.sc/Fiji`). Fluorescence intensity across the central microchamber of the microdevice was quantified, in accordance with the software instructions, by selecting a rectangular region across the central microchamber after creating the projection image. An assumption of proportionality is used to transform fluorescence intensity into cell concentration. The constant of proportionality was calculated assuming that the integral of the initial cell concentration along the chamber equals the initial total number of cells.

Three different experimental configurations have been explored:

- For a necrotic core formation, a high density of cells (40×10^6 cells/mL) was embedded in the collagen hydrogel and injected within the central microchamber.



(a) Meshed geometry.

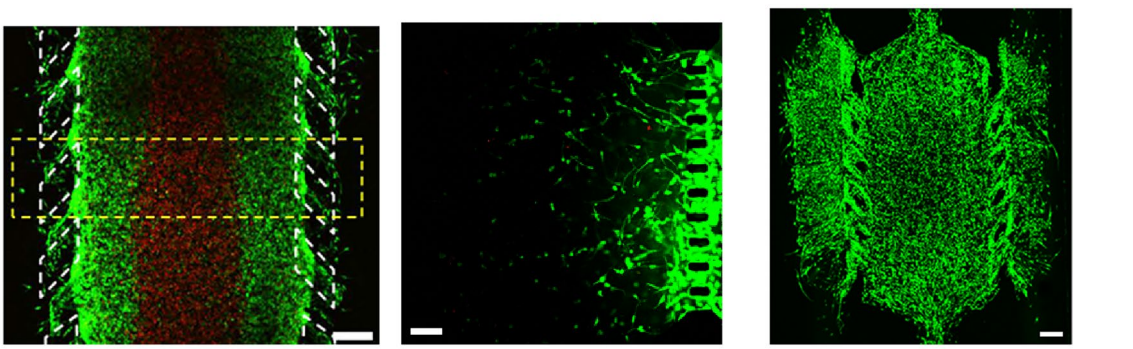
(b) Boundary conditions.

(c) Live (normoxic) cells at the final simulation time.

(d) Dead (necrotic) cells at the final simulation time.

Fig. 5 Central chamber of the three-dimensional microfluidic device used for illustrative simulations. Subfigure **a** depicts the geometry and mesh of the chamber. We have considered a hexahedral mesh with two elements along the device thickness. Subfigure **b** depicts the geometry with the corresponding boundary condition for each surface. Surfaces in red indicate Dirichlet boundary conditions while surfaces in blue represent Neumann boundary conditions. Subfigures **c** and **d** depict the simulation results for live and dead cells respectively

Growth medium was refreshed every day and the culture was maintained for 6 days. Nutrients and oxygen are not able to reach the central part of the device due to cell consumption close to the microchannels, thus causing cell death in the central region and the appearance of an autoinduced necrotic core (Fig. 6a), mimicking the parts of the tumor far from functional blood vessels [18].



(a) Necrotic core (Day 6, scale bar = 400 μm).

(b) Pseudopalisade (Day 6, scale bar = 200 μm).

(c) Double pseudopalisade (Day 21, scale bar = 400 μm).

Fig. 6 In vitro recreation of glioblastoma evolution in microfluidic devices using U251 cells. Viable cells were stained green with calcein AM and dead cells were labeled red with propidium iodide. Subfigure **a** necrotic core formation, with cells seeded at the concentration of 40×10^6 cells/mL within the central microchamber and growth medium perfused every day through the lateral channels. Subfigure **b** pseudopalisade formation, with cells seeded at a concentration of 4×10^6 cells/mL within the microdevice, and medium flow enabled to flow only through the right channel. Subfigure **c** double pseudopalisade formation, with cells seeded at the concentration of 4×10^6 cells/mL at the central chamber and growth medium changed every day

- To promote pseudopalisade formation, cells were seeded at a low density (4×10^6 cells/mL) within the central microchamber and one lateral channel was blocked, while constant medium flow was perfused through the other lateral channel. As the region next to the sealed channel became hypoxic, cells migrated towards the perfused channel, rich in nutrients and oxygen (Fig. 6b) [17].
- Finally, in the case of double pseudopalisade formation, cells were seeded again at low density (4×10^6 cells/mL) within the central microchamber. In this case, the medium was perfused through both lateral channels and refreshed every day during 21 days [67]. Hypoxic conditions in the center of the microchamber induced cell migration towards the perfused channels and invasion of both of them (Fig. 6c).

3.1.2 Model Parameters

In order to particularize the general Eqs. (8) presented for modeling the population and oxygen evolution in the in vitro experiments for GBM cells, it is necessary to specify the nonlinear corrections of the model together with the values of the model parameters: P, the random pedesis coefficient of the normoxic cell population, χ, the chemotaxis coefficient, τ_{gr}, the proliferation characteristic time, τ_d, the death characteristic time, D, the oxygen diffusion coefficient and α, the oxygen consumption rate. Regarding the nonlinear activation functions, F_{gr} and F_{go} are corrections for cell growth and migration due to space constraints, Π_c is related to oxygen consumption kinetics, Π_{go} and Π_{gr} model the *go-or-grow* behavior while Π_d considers the role of oxygen as moderator of cell death [67, 68].

Since cell populations adapt their behavior to oxygen supply and space availability, it has been assumed for the migration term that cellular motility is only possible when the surrounding tissue is not cell saturated [76] and that migration following the oxygen gradient happens only when the oxygen supply is below a critical threshold, activating the cell motility mechanism [77, 78]. According to these two major assumptions, a piecewise linear activation function was used to take into account each of these two phenomena [67], so the chemotaxis corrections may be written as:

$$F_{go}(C_1, C_2) = \phi(C_1; C_{sat}), \tag{13}$$

$$\Pi_{go}(S) = \phi(S; S_H), \tag{14}$$

with

$$\phi(x;\theta) = \begin{cases} 1 & \text{if} \quad x \leq 0 \\ 1 - \frac{x}{\theta} & \text{if} \ 0 \leq x \leq \theta, \\ 0 & \text{if} \quad x > \theta \end{cases}$$

Here S_H is the hypoxia-induced migration activation threshold, representing the oxygen level below which cell migration is activated and C_{sat} is the cell saturation concentration. The proposed model is in line with the *go-or-grow* dichotomy established in the GBM literature [66]: cell energetic resources are spent either in cell migration or in cell proliferation.

Cell proliferation also depends on other needs such as nutrient supply or availability of space to grow and split. Thus, a model was proposed combining logistic growth and the *go-or-grow* paradigm based on oxygen supply [67]:

$$F_{gr}(C_1, C_2; C_{sat}) = 1 - \frac{C_1 + C_2}{C_{sat}}, \tag{15}$$

$$\Pi_{gr}(S; S_H) = 1 - \phi(S; S_H). \tag{16}$$

The function Π_{gr} is responsible for the *go-or-grow* dichotomy and F_{gr} is the logistic model for cell population growth.

Cell death is a natural process depending on many factors and agents and has an inherent stochastic nature [79]. Anoxia is one fundamental cause of cell death [80]. A two-parameter sigmoid has been proposed [67], able to capture necrosis and apoptosis phenomena:

$$\Pi_d(S; S_A, \Delta S_A) = \frac{1}{2}\left(1 - \tanh\left(\frac{S - S_A}{\Delta S_A}\right)\right), \tag{17}$$

with S_A and ΔS_A the location and spread parameters associated with the oxygen concentration inducing cell death.

Finally, oxygen consumption is a complex phenomenon related to the oxidative phosphorylation that occurs in the membrane of cellular mitochondria [81]. Many authors have considered a zero-order consumption function, i.e. a constant consumption rate independent of oxygen concentration [82–84]. A more realistic assumption is describing the consumption function using the Michaelis-Menten model for enzyme kinetics [85, 86]. This type of equation was observed for the oxygen consumption rate in the late 1920s and early 1930s [87] and describes more accurately the consumption at low oxygen concentrations. Thus, it has been proposed [67] to use as the activation function

$$\Pi_c(S; k_m) = \frac{S}{S + k_m}, \tag{18}$$

where k_m is the oxygen concentration at which the reaction rate is half of the rate in a fully oxygenated medium, therefore related to the oxidative phosphorylation kinetics, the cell structure and morphology (size and number of mitochondria, etc.) and diffusion in the cytoplasm.

The parametric mechanism-based model for GBM progression under hypoxic conditions is closed by the above

Table 2 Summary of the different model parameters

Symbol	Meaning	Dimensions
P	Cell random pedesis coefficient	$L^2 \cdot T^{-1}$
χ	Chemotaxis coefficient	$L^2 \cdot Y^{-1} T^{-1}$
τ_{gr}	Growth characteristic time	T
τ_d	Dead characteristic time	T
D	Oxygen diffusion	$L^2 \cdot T^{-1}$
α	Oxygen consumption rate per unit cell concentration	$Y \cdot L^3 \cdot T^{-1}$
C_{sat}	Saturation capacity	L^{-3}
S_H	Hypoxia threshold	Y
S_A	Anoxia threshold	Y
ΔS_A	Anoxia sensibility	Y
k_m	Michaelis–Menten constant for oxygen consumption kinetics	Y

L is space, T is time and Y is the physical magnitude representing oxygen level (for instance oxygen mass concentration, oxygen molar concentration, oxygen partial pressure...). In the following, the oxygen level is measured using pressure units (mmHg)

definitions together with the model parameters, with the latter summarized in Table 2 [67, 68].

3.1.3 Boundary and Initial Conditions

In the microfluidic device, the culture chamber is connected to the oxygen supplying channels by means of small cavities. The volume and the number of these cavities depend on the microfluidic device design and they are directly related to potential cell losses during the experiment. Actually, when cell populations arrive to the interface between these cavities, some of them may reach the channel and leave the culture. To take this into account, Robin boundary conditions are usually considered [67]:

$$C_1(x,t) + J\left(P\frac{\partial C_1}{\partial x} - \chi F_{go}(C_1,C_2)\Pi_{go}(S)\frac{\partial S}{\partial x}C\right)\Big|_{x=0} = 0, \tag{19a}$$

$$C_1(x,t) - J\left(P\frac{\partial C_1}{\partial x} - \chi F_{go}(C_1,C_2)\Pi_{go}(S)\frac{\partial S}{\partial x}C\right)\Big|_{x=L} = 0. \tag{19b}$$

As dead cell population is slave to live cells, no boundary conditions are needed for C_2. Regarding oxygen supply, two different scenarios are considered, depending on the operating conditions [67]:

- When oxygen is supplied normally, Dirichlet boundary conditions are considered (the oxygen concentration at the channels remains constant and known throughout the experiment):

$$S(x=0,t) = S_{left}, \quad S(x=L,t) = S_{right}, \tag{20}$$

  where S_{left} and S_{right} are known values.
- On the other hand, when a channel is sealed, Neumann boundary conditions are considered:

$$D\frac{\partial S}{\partial x}\Big|_{x=0} = 0, \quad D\frac{\partial S}{\partial x}\Big|_{x=L} = 0. \tag{21}$$

Finally, we assume that, at time $t = 0$, all cells are alive and the cell population concentration is known throughout the whole culture chamber, giving the initial conditions:

$$C_1(x,t=0) = C_1^0(x), \quad C_2(x,t=0) = 0. \tag{22}$$

Moreover, the initial oxygen profile is assumed to be constant along the chamber and equal to the higher concentration in the channels:

$$S(x,t=0) = \max\{S_{left}, S_{right}\}. \tag{23}$$

3.1.4 Numerical details

The PDEs (8) with boundary conditions (19) and (20) or (21), and initial conditions (22) and (23) were solved following the approach presented in Sect. 2.2. The domain length (associated with the microfluidic device) and mesh size used for the simulation of each experiment are summarized in Table 3.

Table 3 Domain and mesh size for the different simulations

Experiment	Chamber length $L\,[\mu m]$	Mesh size $\Delta x\,[\mu m]$
Necrotic core formation	2000	3.0
Pseudopalisade formation	916	4.8
Double pseudopalisade formation	2897	12.0

The mesh is adapted to the domain size, the measurement resolution, and also the cell size so that it can resolve even sharp cell density gradients and the associated gradients in chemical species that these may induce

3.2 Probabilistic Parameter Fitting

The model for GBM evolution presented in Sect. 2.1.2 (and, in general, any other physical model) can be described in terms of measurable variables $\boldsymbol{u}$ (cell and oxygen concentrations), controlled variables λ (boundary and initial conditions of cell and oxygen concentrations), and a mathematical

model $\boldsymbol{F}$ which, in case of parametric models as ours, is given in terms of a set of parameters $\boldsymbol{\theta}$. That is:

$$\boldsymbol{u} = \boldsymbol{F}(\lambda, \boldsymbol{\theta}). \tag{24}$$

The parameter vector $\boldsymbol{\theta}$ is usually obtained following a deterministic approach from some experimental data $\boldsymbol{u}_{\text{exp}}$, solving the minimization of the residual $\boldsymbol{R}$, defined as:

$$\boldsymbol{R}(\boldsymbol{\theta}) = \sum_{i=1}^{N} \|\boldsymbol{u}^i_{\text{exp}} - \boldsymbol{F}(\lambda^i, \boldsymbol{\theta})\|^2, \tag{25}$$

with N the number of available experiments for the parameter fitting.

However, in general for complex problems, the experimental data are noisy and the model is not perfect. Besides, in biological problems, it is difficult to define universal values for the parameters and there are highly correlated phenomena. Indeed, in the model previously presented, correlation between parameters was identified [67].

Until relatively recently, traditional computational biology research has often only considered pointwise estimates of parameters, overlooking the correlations that might exist [88]. However, there can often be numerous correlations for instance due to two main principal reasons, namely samples variability and model complexity. Sample variability refers to the inherent correlation present among physical phenomena. We can also refer to this type of correlation as *prior* correlation. On the other hand, model complexity refers to the nonseparability of the model, which prevents separating phenomena. In this case, the correlation is not inherent to the problem, but caused by the lack of information, and we may also call it *posterior* correlation.

As a consequence of all this, it is more appropriate to move towards a stochastic approach for Eq. (24):

$$\boldsymbol{U} = \boldsymbol{F}(\lambda, \boldsymbol{\Theta}), \tag{26}$$

where $\boldsymbol{U}$ and $\boldsymbol{\Theta}$ are now random variables. This approach is suitable when:

- The model $\boldsymbol{F}$ has many parameters and/or is non-separable. That is, there are many coupled phenomena (model complexity condition).
- The dimension of the measurement space and/or the sample size are large. That is, a sufficiently big number of experiments may be performed (data availability condition).

In what follows, we particularize the presented stochastic approach for the model of GBM evolution presented above. Due to scarcity of data, we consider that we have only one experimental configuration, corresponding to the formation of the necrotic core experiment [68]. The experimental configuration is determined via three control parameters $\lambda = [S_{\text{left}}, S_{\text{right}}, C_1^0]$. In this experiment, a high cell concentration is initially seeded within the device, $C_1^0 = 4 \cdot 10^7$ cell/mL and oxygen is perfused through both lateral channels $S_{\text{left}} = S_{\text{right}} = 7$ mmHg. For a more general case in which the variability induced by the experimental configuration is taken into account, the reader is referred to [89].

We study the most relevant parameters $\boldsymbol{\theta} = [P, \chi, 1/\tau_{\text{gr}}, \alpha, S_{\text{H}}]$, which are those related to cell migration, proliferation and oxygen uptake. We assume that the values of the remaining parameters are fixed and dependent on the particular experimental configuration (the cell concentration saturation C_{sat} and the parameter accounting for cell escape in Robin boundary conditions J, taken from Ref. [67]) or known from literature (the oxygen diffusion coefficient D [90], the parameters related to cell death, τ_{d}, ΔS_{A} [65] and S_{A} [67], as well as the Michaelis-Menten constant k_{m} [91]). The values of these parameters are listed in Table 4.

To characterize the model from a probabilistic point of view, we use copulae [92, 93]. This mathematical tool is able to cope with non-linear structural dependencies while keeping the parameters within a physically-convenient range. A copula is a n-multivariate probability distribution whose marginals are uniform distributions and describes the structure of the dependence between random variables. *Copulae* are nowadays applied to a wide variety of areas, ranging from economic sciences [94, 95] to climate-agents time series [96]. In this work, we apply the concept of *copulae* to biological problems, since they are useful to analyze highly-correlated multiparametric models such as those appearing in biology and cancer evolution.

The methodology followed to obtain the probabilistic model is:

Table 4 Fixed parameter values

Parameter	Value	Units	References
C_{sat}	5.0×10^7	cell/mL	[67]
J	1.0×10^6	s/cm	[67]
D	1.0×10^{-5}	cm^2/s	[90]
τ_{d}	1.7×10^5	s	[65]
S_{A}	1.6	mmHg	[67]
ΔS_{A}	0.1	mmHg	[65]
k_{m}	2.5	mmHg	[91]

The value of these parameters are fixed since they are directly related to the experimental configuration or they are well characterized in literature

Table 5 Reference values for the copula parameters

Parameter	Fitted value	Units
P	6.6×10^{-10}	cm^2/s
χ	7.5×10^{-9}	$cm^2/mmHg \cdot s$
τ_{gr}	7.2×10^5	s
α	1.0×10^{-9}	$mmHg \cdot mL/cell \cdot s$
S_H	7	mmHg

Taken from Ref. [67]

1. Dataset generation. A sample of $M = 6000$ simulations was created by perturbing each parameter in $\boldsymbol{\theta}$, yielding a sample $\Theta \sim \mathcal{N}(\theta^*, 0.3\,\theta^*)$ with θ^* a reference value for each parameter taken from the deterministic fitting in Ref. [67], reproduced in Table 5, with a 30% of relative uncertainty.

   We define the quality of a simulation via the objective cost function T:

$$T = \frac{1}{C_{sat}} \sqrt{\frac{1}{n(n_d - 1)L} \sum_{j=1}^{n} \sum_{i=2}^{n_d} \int_0^L \left(u_j(x, t_i) - u_j^e(x, t_i) \right)^2 dx}, \tag{27}$$

   where u_j^e corresponds to the experimental measurement of the live ($j = 1$) or dead ($j = 2$) cell concentration, u_j is the simulated result, L the chip length, n the number of cell populations and n_d the number of days where there are experimental data available. The vector parameters that provided simulated solutions close enough to the experimental necrotic core results ($T < 0.18$) were kept.
2. Fitting of marginal distributions via kernel density estimations (with Gaussian kernels with variable bandwidth), using the parameters described in Ref. [89].
3. Transformation of the marginals into uniform distributions.
4. Copula fitting to capture the joint dependence. In this case, we chose a t-Student copula, following the criteria specified in Ref. [89].

Then, following a typical train-test approach (dividing the dataset in two separate subsets), we test the statistical fitting in terms of both the marginal fittings and the whole joint structural dependence. The technical details of both train and test procedures can be found in Ref. [89]. Figure 7 illustrates the generated dataset in terms of the marginal distributions, the scatter plot of the fitted parameter values and the linear correlation coefficient for each pair of parameters. These aspects of the dataset used to generate the copula (labelled as training dataset) are compared to those obtained with a sample of parameters generated with the copula, to test if it is able to reproduce the variability existing in the data.

Once characterized, the probabilistic model can be used to estimate the output variables, such as the cell concentration. That is, the distribution of the random vector $\boldsymbol{U}$. Figure 8 shows the 90% confidence band for the simulated profiles, together with the experimental results and the concentration profile resulting from a deterministic fitting (from Ref. [67], that is, taking the parameter values from Tables 4 and 5), for the three different experimental configurations described in Sect. 3.1.1: the necrotic core (Fig. 8a), the pseudopalisade (Fig. 8b) and the double pseudopalisade formation (Fig. 8c). All three experiments share the same parametric structure since they were carried out with the same cell line, but they have different domains (Table 3) and boundary conditions. In particular, in the double pseudopalisade and necrotic core formation, Dirichlet boundary conditions were chosen, so that oxygen partial pressure in the channels was assumed to remain constant throughout the whole experiment, since medium flow provision through the channels was sufficiently frequent to keep a fixed partial pressure without important variations despite oxygen diffusion and cell uptake. Therefore, $S_{left} = S_{left} = 7.0\,mmHg$. In the pseudopalisade formation however, an impermeability condition (no flux) was imposed at the sealed channel, while, again, a Dirichlet condition was imposed at the right channel with a value of $S_{right} = 2\,mmHg$, as in this experiment the medium was not renewed as in the previous cases, so the oxygenation was lower [67]. Live cells can escape from the chamber and so Robin boundary conditions are imposed, with the J parameter dependent on the particular experiment. For the necrotic core, $J = 1 \cdot 10^6\,s/cm$; for the pseudopalisade $J = 1 \cdot 10^9\,s/cm$; and finally in the double pseudopalisade case $J = 1.2 \cdot 10^7\,s/cm$.

As can be seen in Fig. 8 the experimental results mostly lay within the confidence interval, while the areas where they lay outside highlight the model limitations. It is also important to note that only the necrotic core experiment was used in the copula model fitting, so the results obtained for the pseudopalisade and double pseudopalisade experiments show the ability of the model to generalize. Compared to the deterministic fitting, it can be seen that the stochastic approach takes into account the inherent parametric variability and the structural dependence, and the width of the band reflects the influence of this variability on model predictions.

One of the main problems in mathematical modeling is the lack of reliable values for the many parameters involved that often forces estimates to rely on values fitted from diverse situations, leading sometimes to unreliable conclusions. The presented probabilistic model can be applied to parameter estimation. We use Highest Density Regions (HDR) [97], due to their easy interpretation and direct implementation. An example of HDR is displayed in Fig. 9 for the motility parameters: the pedesis coefficient P and the chemotaxis coefficient χ, and levels of confidence of 90

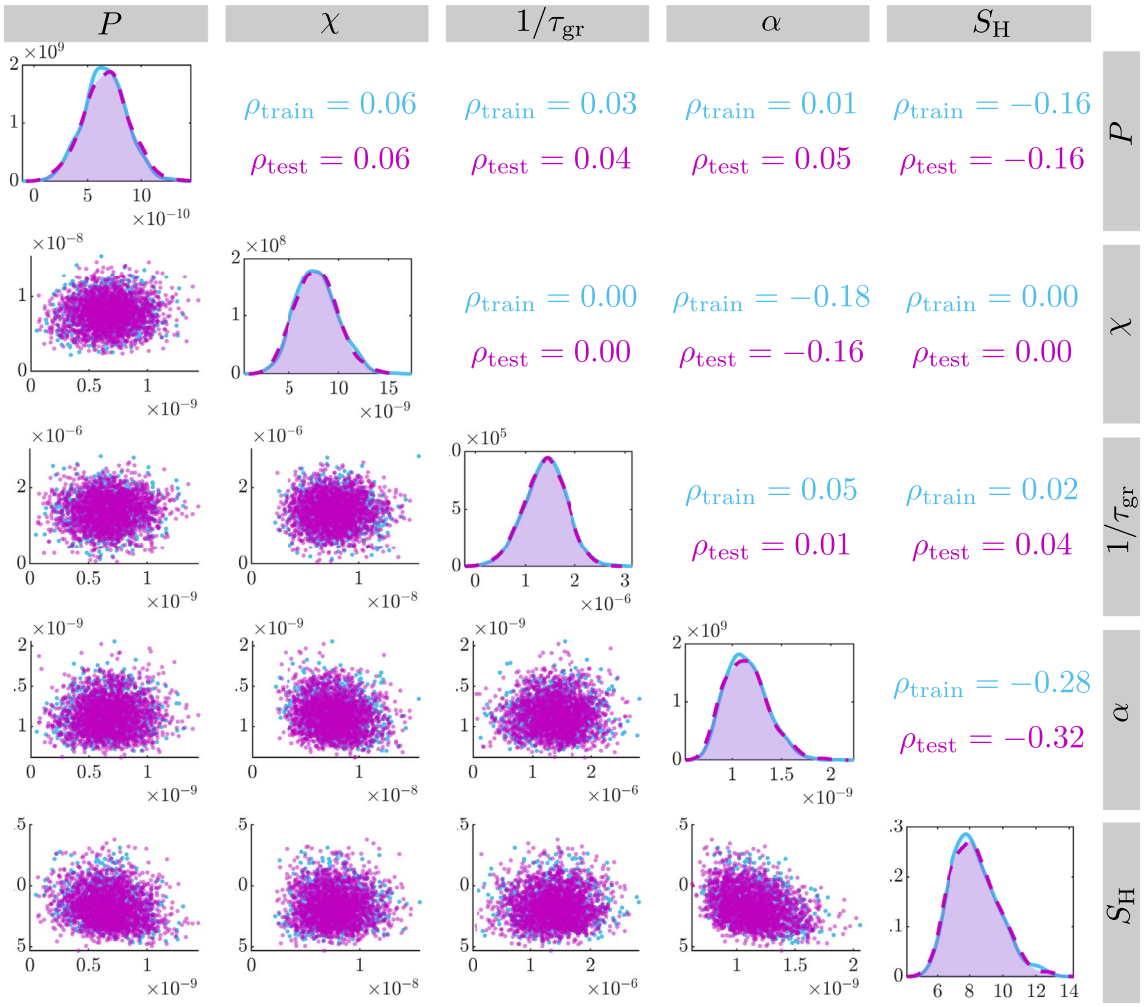
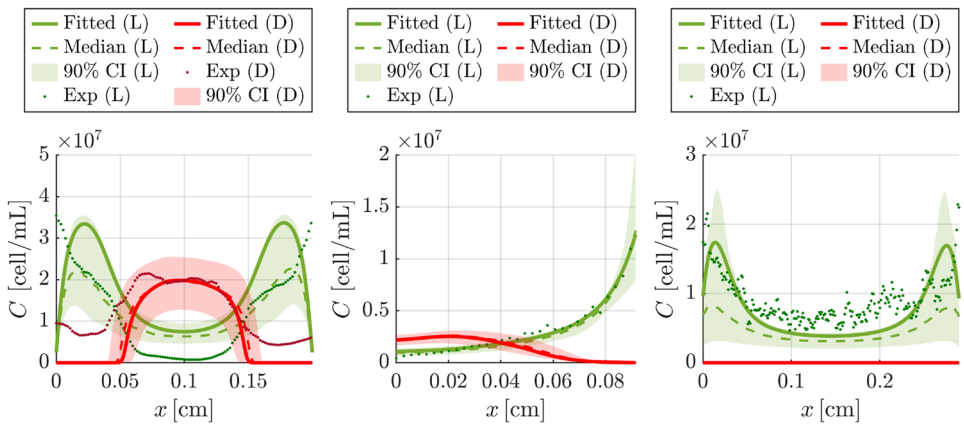


Fig. 7 Dataset validation. Marginal distributions, the scatter plot of the fitted parameter values and the linear correlation coefficient for each pair of parameters of the generated dataset (in blue color) compared to the marginal distributions, scatter plots and correlation coefficients of a sample generated with the copula (in purple color)



(a) Necrotic core. (b) Pseudopalisade. (c) Double pseudopalisade.

Fig. 8 Stochastic prediction of cell concentration with the probabilistic model. The 90% confidence interval (CI) and median for live (L) and dead (D) cells obtained from the probabilistic model (using 100 parameter samples) is depicted together with the results of a deterministic fitting (named "Fitted" in the legend) and compared to the experimental results ("Exp") for the three experimental configurations available. In all the cases, the results are shown for the final experimental time

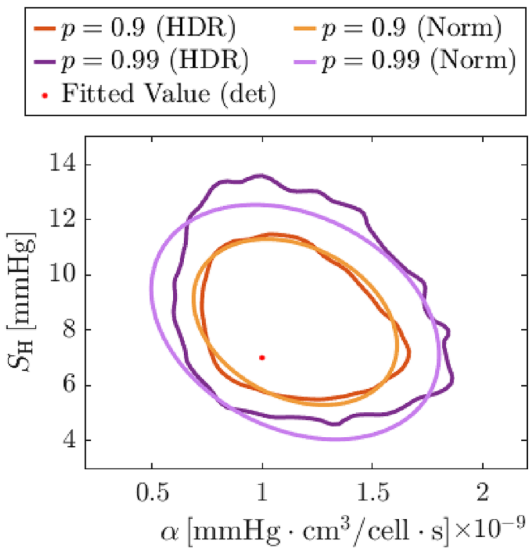


Fig. 9 Parameter estimation of the motility parameters. Highest Density Regions and ellipsoid estimation assuming normality are plotted for two confidence levels: $p = 0.90$ and $p = 0.99$.

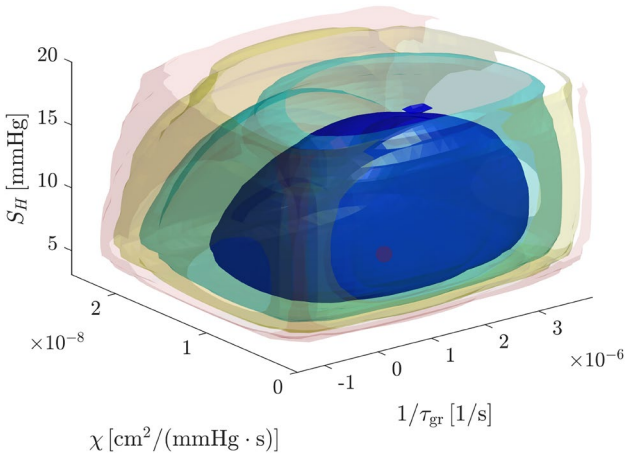


Fig. 10 Isosurfaces of the 3D probability density function. The isosurfaces have been represented log-spaced between the maximal and minimal value of the density at the represented region. The red dot is the deterministic fitted value. Figure extracted from Ref. [68]

and 99%. It can be observed that there are differences in the estimated region assuming normality, justifying the use of distribution-free techniques like HDR.

The stochastic methodology presented has also been used for characterizing the *go-or-grow* behavior, of particular interest in GBM evolution, characterized in the parametric model by the chemotaxis coefficient χ, the growth rate $1/\tau_{gr}$ and the hypoxic threshold S_H. A kernel estimation of the marginal distribution of these three parameters has been explored, using a Gaussian kernel with variable bandwidths for the marginal distributions and a *t*-Student *copula* for representing the structural dependence [68].

In Fig. 10 some isosurfaces of the three-dimensional probability density function at the space $(1/\tau_{gr}, \chi, S_H)$ are shown. The results are compared with the deterministic value associated with the values obtained in Ref. [67].

3.3 Design of an Experimental Campaign

On many occasions, the design of experimental campaigns is a cumbersome and challenging task due to the combinatorial explosion in the number of possibilities that arises when the experimental configuration is determined by different parameters or control variables λ. A paradigmatic example of this situation is experimentation with a microfluidic device, where we can control the oxygen and nutrient flow in both channels, as well as the cell concentration in the central chamber and the properties of the seeded hydrogel, together with other design parameters such as the chamber length. Furthermore, generally, it is not straightforward to determine the configurations that will work well or produce interesting results for our purpose. In such cases, the Design of Experiments (DOE) techniques can be useful to reduce the number of experiments required [98], thus reducing the time and resources required. The purpose of DOE is to maximize the information obtained from each performed experiment. Different methods within DOE have been used in systems biology (see Ref. [99] for a review). There are several optimality criteria, being the most widespread A-, D- and E-optimal designs [100], that have been used in parameter estimation in mathematical models of different biological problems, such as microbiology [101], signaling pathways [100] or animal cell models [102, 103].

In particular, the use of probabilistic models such as the one herein presents enables the application of techniques within Bayesian Experimental Design (BED), based on the Bayesian interpretation of probability, in a natural way. According to the BED theory, the aim is to choose the experiment that provides the highest utility, and the optimal experiment depends on the selected utility criterion [104, 105].

BED is particularly useful to define sequential campaigns, in which the information available to determine the next optimal experiment is based on all the previous experiments conducted [106]. In particular, in this work we present a methodology to design an experimental campaign based on the probabilistic model defined in Sect. 3.2 and using Shannon information theory to define the utility function [107]. According to this theory, the utility of an experimental configuration $U(\lambda)$ to characterize some model parameters θ of a parametric space Ω is defined as the prior-posterior gain in Shannon entropy or information:

$$U(\lambda) = \int_{\mathcal{U}} \int_{\Omega} f(\boldsymbol{\theta}, \boldsymbol{u}|\lambda) \log f(\boldsymbol{u}|\boldsymbol{\theta}, \lambda) \, \mathrm{d}\boldsymbol{\theta} \, \mathrm{d}\boldsymbol{u} - \int_{\mathcal{U}} f(\boldsymbol{u}|\lambda) \log f(\boldsymbol{u}|\lambda) \, \mathrm{d}\boldsymbol{u}, \tag{28}$$

with $\boldsymbol{u} \in \mathcal{U}$ is the experimental measurement, being $\mathcal{U}$ the measurement space. The probability density function of obtaining $\boldsymbol{u}$ with the configuration λ and the underlying model parameters $\boldsymbol{\theta}$ is $f(\boldsymbol{u}|\boldsymbol{\theta}, \lambda)$ and:

$$f(\boldsymbol{\theta}, \boldsymbol{u}|\lambda) = f(\boldsymbol{\theta}) f(\boldsymbol{u}|\boldsymbol{\theta}, \lambda). \tag{29}$$

As mentioned before, the experimental campaign is designed to define the most useful experiments to characterize the *a priori* unknown values of the parameters defining the *go-or-grow* behavior, that is, $\boldsymbol{\theta} = [\chi, \tau_{\mathrm{gr}}, S_{\mathrm{H}}]$. We assume that, initially, there is one experiment available, in this case the necrotic core (that is, $\lambda = [4 \cdot 10^7 \, \mathrm{cell/mL}, 7 \, \mathrm{mmHg}, 7 \, \mathrm{mmHg}]$), and proceed iteratively to obtain the next experiments to perform. For illustration purposes, we generate the experimental data synthetically using the previously described mathematical model and some correlations that we define as ground truth to generate the data but these are unknown from the point of view of the experimental design. In particular, we assume normal distributions for $[\chi, \tau_{\mathrm{gr}}, S_{\mathrm{H}}]$ with mean and standard deviations detailed in Table 6, while the rest of parameters are considered fixed at the value fitted in Ref. [67].

Moreover, we assume that there exist certain correlations between that three parameters, with a Pearson correlation matrix:

$$P = \begin{bmatrix} 1 & 0 & 0.6 \\ 0 & 1 & 0 \\ 0.6 & 0 & 1 \end{bmatrix} \tag{30}$$

In further detail, the process followed is:

1. Data generation. We use the ground truth model to generate 10 replicates of the available experiments. The data are uniformly perturbed additively with uniform noise of 20% of the measured value.
2. The optimization problem is solved to obtain the model parameters' marginal distributions from the available data.
3. The transformation from the marginal distribution to a uniform distribution is constructed.
4. The Copula fit is implemented.
5. A utility computation is performed to determine the experiment with the highest utility.
6. Finally, we return to step 1 adding the new experiment to the available data.

Table 6 Parameters of the distributions used for data generation

Parameter	χ [mmHg · cm^3/(cell · s)]	τ_{gr} [h]	S_{H} [mmHg]
Mean μ	$7.5 \cdot 10^{-9}$	200	7
Standard deviation σ	$1.5 \cdot 10^{-9}$	40	1.4

For simplicity, normal distributions are assumed for the *go-or-grow* parameters, using as mean the reference value fitted in Ref. [67]

The scheme of the whole process is represented in Fig. 11. For illustrative purposes, we perform 8 iterations of the procedure, adding the optimal experiment at each iteration to the available data. The list of optimal experiments obtained at each iteration (that is, the experimental configuration with the maximum utility value) is presented in Table 7.

As expected, it is shown in Fig. 12 how the fitted correlation coefficients approach the ground truth with increasing iteration number. In a similar fashion, the set of marginal distributions approach the ground truth marginals (Fig. 13). Even if the distribution of S_{H} (Fig. 12c) seems to look better at the first iteration, this may be a fortuitous event (in fact, in iteration 4 it has worsen and then it improves in iteration 8) and the set of the three marginals improves along iterations.

Finally, we evaluated how the prediction of the experimental outcome, in terms of the cell concentration profiles, evolves throughout the experimental campaign. Figure 14 shows how the predicted 90% confidence band sharpens around the ground truth as more experiments are performed.

4 Pure Data-Driven Approaches

Notwithstanding the significant advances achieved by the combination of experiments and mathematical modeling of GBM, there are still some fundamental aspects to tackle, such as the heterogeneity between patients [108]. Patient-specific models become crucial to characterize the defining features of each individual tumor, thus capturing the heterogeneity and allowing personalized prediction [33]. As such, we need to calibrate the model for each patient, obtaining the parameters governing tumor evolution for that specific GBM. Since this task is computationally very expensive using traditional optimization techniques, artificial intelligence and, in particular deep learning tools, which are evaluated in real time, can prove very useful.

Indeed, artificial intelligence has become recently very popular and permeated almost every research field, including medicine and, in particular, oncology [109]. Regarding the aforementioned problem of parameter identification to create individualized models, artificial neural networks (ANN) have been used to characterize biological soft tissues [110–112] or to design customized intraocular lenses [113]. Machine learning can also be used for cell culture monitoring, for cell counting and viability analysis [114] as well as

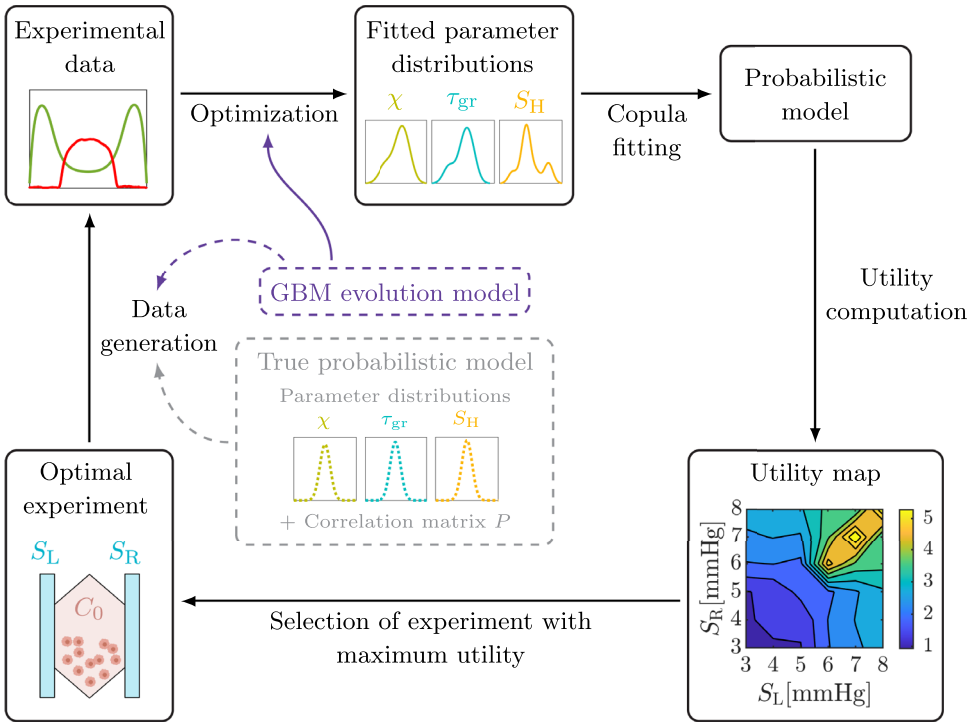
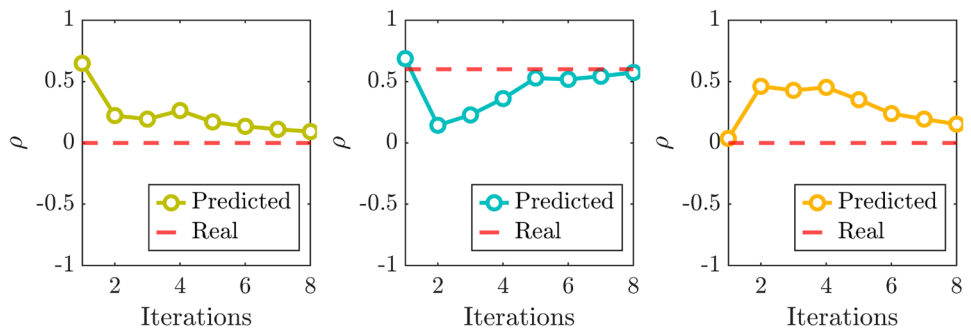


Fig. 11 Schematic representation of the procedure of design of an experimental campaign. The iterative procedure starts with some available experimental data, which allow the fitting of a probabilistic model and, via utility maximization, the determination of the optimal experiments to perform at the next step

Table 7 Optimal experimental configurations

Iteration	1	2	3	4	5	6	7	8
C_0 [cell/mL]	$4 \cdot 10^7$	$4 \cdot 10^7$	$4 \cdot 10^7$	$4 \cdot 10^7$	$4 \cdot 10^7$	$4 \cdot 10^7$	$4 \cdot 10^6$	$4 \cdot 10^7$
S_L [mmHg]	7	7	8	8	8	8	7	8
S_R [mmHg]	7	7	8	8	8	8	7	8

The experiment with the maximum utility obtained at the end of each iteration is represented in terms of the initial cell concentration C_0 and the oxygen concentration at the left S_L and right S_R channels respectively



(a) P_{12} ($\chi - \tau_{gr}$) (b) P_{13} ($\chi - S_H$) (c) P_{23} ($\tau_{gr} - S_H$)

Fig. 12 Evolution of the linear correlation coefficients. The fitted value of the linear correlation coefficients between each pair of parameters is depicted throughout the iterations of the experimental campaign, in which new experiments are progressively being conducted to improve the fitting of the probabilistic model

for cell segmentation and classification [115, 116]. Indeed, the emergence of microfluidics enables the acquisition of massive amounts of data, opening the door to the use of machine learning in microfluidic devices, for data screening and automated culture [117] or for device design and control [118]. This paradigm has been recently coined as *Intelligent Microfluidics* [119, 120]. Within this framework, we present a deep learning approach for parameter identification in our GBM evolution model, which allows the characterization of model parameters from a cell culture image from a microfluidic device [121].

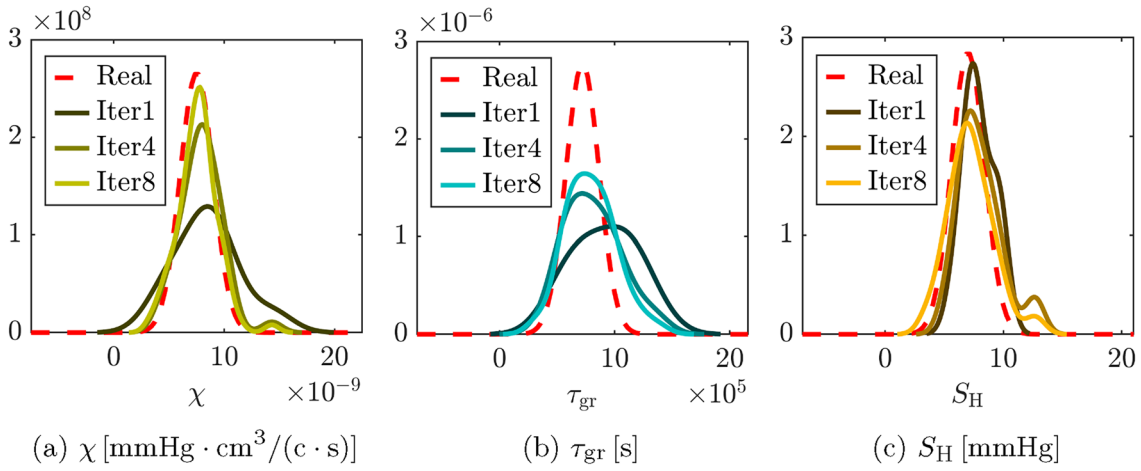


Fig. 13 Evolution of the parameter distributions. The fitted value of the parameter distributions is depicted throughout the iterations of the experimental campaign, in which new experiments are progressively being conducted with the aim of improving the fit of the probabilistic model

We use a Convolutional Neural Network (CNN) which, from fluorescence images of GBM culture in microfluidic devices, is able to predict the value of the *go-or-grow* parameters of the mathematical model presented in Sect. 3 [121]. These parameters characterize the evolution of the particular GBM of the input image in terms of the proliferation and migration activity, which are, as discussed below, key in the prognosis of the tumor. Subsequently, the parameters can be used to predict GBM evolution in different conditions.

The methodology followed as well as the main results are summarized in Fig. 15. More detailed information can be found at [121]. Firstly, Fig. 15a depicts the network training procedure. The dataset used is generated with a non-dimensional version[1] of the previously described mathematical model, recreating the experiment of the formation of a necrotic core in a microfluidic device. As commented before, the aim is to predict the value of the parameters involved in the *go-or-grow* behavior (suitably non-dimensionalized), τ_{gr}^*, χ^*, and S_H^*. These parameters are sampled from uniform distributions while the remaining parameters is considered fixed at the fitted value (Tables 4 and 5). Consequently, the dataset is composed of $n = 12000$ live and dead cell concentration profiles and the corresponding set of parameters $\{\tau_{gr}^*, \chi^*, S_H^*\}$ used to generate them. For full details on the generation of the dataset, the reader is referred to [121].

We used a Convolutional Neural Network (CNN) since they are suitable for image processing and inferring hierarchical patterns. These features are useful for the problem of parameter identification since the information contained in cell concentration profiles is likely to be locally distributed. The network receives a two-channel input, corresponding to live and dead concentration cell profiles at the end of the simulation time each discretized with 51 points. The CNN has four convolutional layers, with Rectified Linear Units (ReLUs) as activation functions, each followed by batch normalization and average pooling layers. After the last convolutional layer, there is a fully connected layer, followed by a regression layer, which returns the three predicted parameters.

The size of the convolutional layers is:

1. 12 convolutions with kernel size [30, 1], stride [1, 1] and same padding.
2. 32 convolutions with kernel size [30, 1], stride [1, 1] and same padding.
3. 64 convolutions with kernel size [30, 1], stride [1, 1] and same padding.
4. 64 convolutions with kernel size [50, 1], stride [1, 1] and same padding.

From the total amount of results, $n_{train} = 8000$ were used for training the network, $n_{val} = 2000$ for validation during the training, and the remaining $n_{test} = 2000$ for testing the network once trained. The training consisted on 50 epochs, with batches of 30 virtual experiments, using the Adam optimizer with a learning rate of $\beta = 1 \cdot 10^{-3}$. Further details on the training procedure can be found in Ref. [121]. The CNN was implemented in the Deep Learning Toolbox of Matlab [122].

The Root Mean Squared Error (RMSE) together with the linear correlation coefficient (ρ) obtained on the test dataset after training for each parameter can be found in the Table in Fig. 15a. It can be seen that for the three parameters

[1] We use non-dimensionalization to avoid differences in the magnitude of the parameters predicted by the CNN.

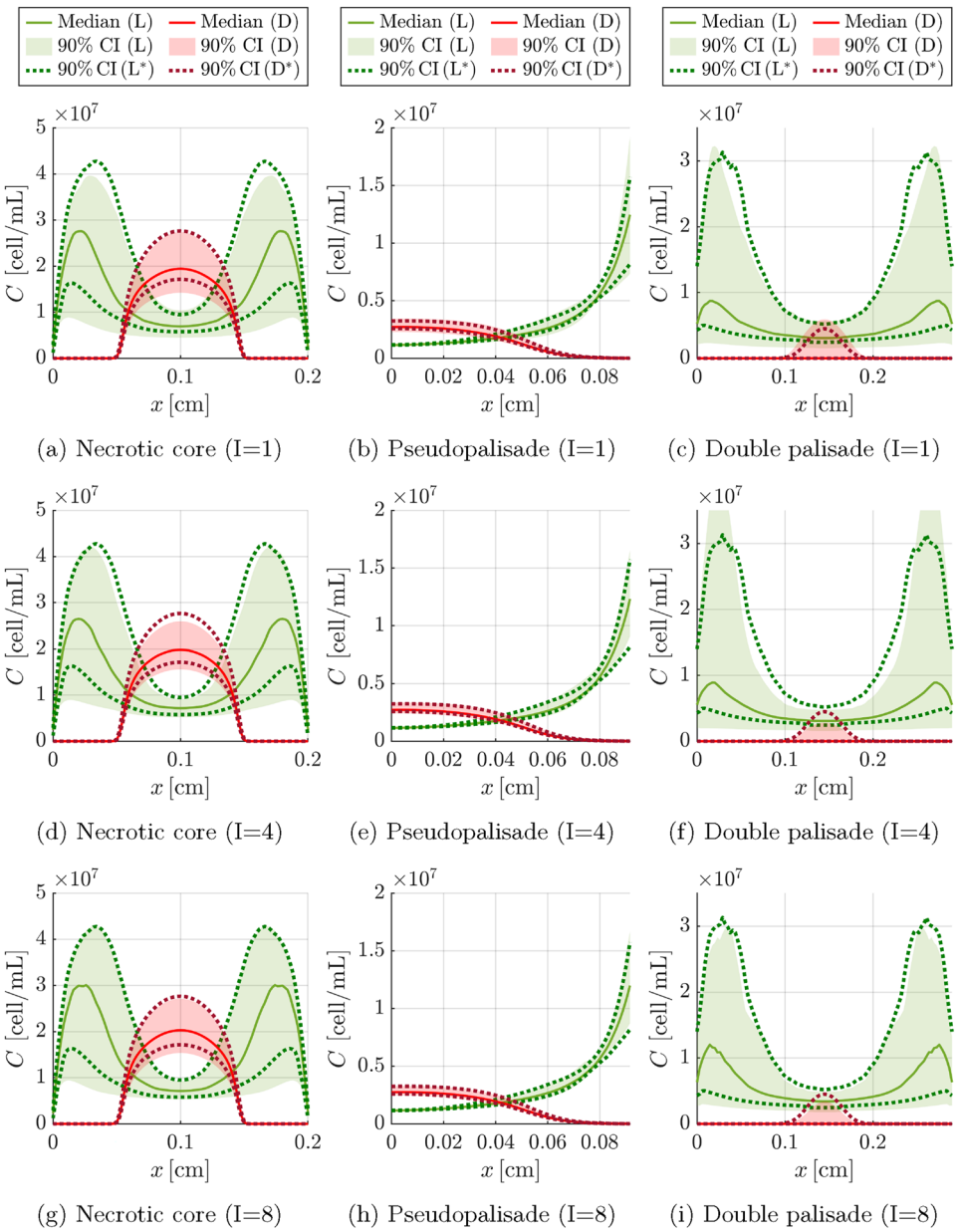


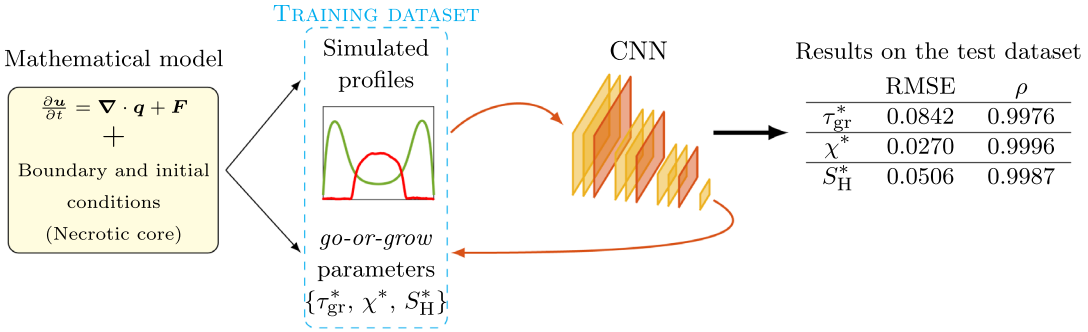
Fig. 14 Evolution of the prediction of cell concentration. The 90% confidence intervals (CI) of the model prediction for live (L) and dead (D) cell concentrations are compared to the ground truth model (indicated in the legend with L* and D* for live and dead cells respectively) used for data generation for iterations $I = 1$, $I = 4$ and $I = 8$. It can be seen that the predicted band becomes closer to the ground truth along the iterations for the three experimental configurations analyzed

RMSE ≤ 0.1 and $\rho \geq 0.99$ and there were no signs of overfitting during the network training.

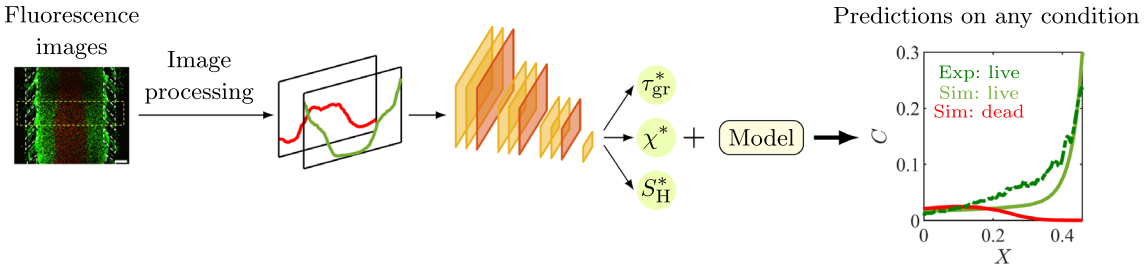
Figure 15b shows a scheme of the CNN evaluation. Once the training is completed, the resulting CNN can be used to predict the parameters governing the evolution of a GBM sample cultured in a microfluidic device. These parameters can in turn be fed to the mathematical model and used to simulate and predict the evolution in different conditions. In particular, in Fig. 15b it can be seen how the obtained parameters accurately predict the formation of the pseudopalisade, even if the network was trained only with results from the necrotic core experiment.

The parameters derived from the network closely resemble those determined through the mathematical model using conventional fitting methods, except for the proliferation rate. Nonetheless, the experimental outcomes are faithfully replicated in the current study. This indicates the limited relevance of proliferation in these experiments, which have a short duration with the cells subjected to low oxygen levels that they mainly show migratory activity.

In summary, the presented approach falls within the context of intelligent microfluidics, combining the increasing power of microfluidics and deep learning tools to characterize the behavior of GBM cultures. The work goes in the direction of patient specific modeling since it allows real time extraction of the parameters that determine how a specific tumor behaves, eventually allowing prospective in silico predictions of prognosis and the in silico testing of



(a) Network training.



(b) Network evaluation.

Fig. 15 Scheme of the methodology. Subfigure **a** depicts the training process, using in silico concentration profiles and the corresponding parameters that generated them obtained from the mathematical model, together with the metrics obtained after training on the test dataset. Subfigure **b** shows how the network can be evaluated to determine patient-specific parameters governing the evolution of a specific GBM cell culture, which can be used to predict the evolution of the tumor under different conditions

potential treatments at a personalized level to support decision making.

To conclude, it must be noted that the methodology herein presented can only learn the parameters of a previously defined mathematical model. Thus, it must be assumed that cell behavior is adequately described with the mathematical model. However, there could be functions that better describe some phenomena than the ones currently used. For example, the *go-or-grow* behavior could have been described with other activation functions such as heaviside functions or hyperbolic tangents. To overcome these limitations, the next section explores the problem from a wider perspective, allowing the unraveling of cell behavior in a non-parametric way.

5 Hybrid Approaches

5.1 Scientific Machine Learning Approaches to GBM Progression

Parametric models are corseted by the mathematical relations that describe the *a priori* assumed biological hypotheses, so they present an obvious modeling bias. Besides, in many experimental investigations, the main goal is to understand the intrinsic mechanisms that control these biological processes, rather than to be able to predict new outcomes. This knowledge goes further than the numerical value of a specific parameters and is generally related to concepts with biological meaning, such as whether the metabolic change is sharp or smooth, reversible or irreversible, localized or distributed in different oxygen regimes, or presents one or many different levels of transition.

There have been several attempts to unravel the mechanisms governing cellular metabolism, combining data and knowledge at different levels, particularly in mathematical oncology [123]. Indeed, the combination of ML learning capacity power and mechanistic modeling knowledge is assessed with hope [124]. Some works have proposed the use of Physics-Informed Neural Networks in biology in combination with cell culture data as a way to discover hidden mechanistic relationships, using the Fisher-Kolmogorov-Petrovsky-Piskunov equation as a benchmark problem, a concept that has been coined as Biologically-Informed Neural Networks (BINN) [125]. When dealing with cell cultures, some of the variables are accessible via measurements, whereas other variables are internal and hidden, accessible only via the use of mathematical models, so unraveling

complex metabolic mechanisms may be challenging [126, 127]. In addition, microfluidic systems are capable of reproducing physiological or pathological microenvironments in a way that is not easy to achieve using traditional cell culture systems, thus bridging the gap between traditional in vitro research and in vivo models, shedding light on the tumor process as well as increasing the effectiveness of drug development [128].To achieve these physiological and pathological conditions, cell cultures are subjected to variable external stimuli, that translate into different source terms and boundary conditions for mathematical models.

Physically Guided Neural Networks with Internal Variables (PGNNIV) have appeared recently as a method for the combination of conventional physical modeling and Machine Learning methods [44, 129]. In contrast to other Scientific Machine Learning methods, such as Physics-Informed Neural Networks (PINNS) [130, 131], the physics does not *constrain*, but only *guides* the network learning capacity, as the measured data may be supplied to endow the network with explanatory capacity [68]. In addition to having demonstrated its predictive and explanatory capabilities in Computational Solid Mechanics [43], PGNNIV have also been applied to the GBM progression problem both for understanding the *go-or-grow* mechanism and for predicting the cell culture evolution under different external stimuli (boundary conditions) [132].

Suppose we have the conditions of Sect. 3 where we can assume a 1D problem. If we discretize Eqs. (11) we obtain

$$\begin{aligned}\dot{\boldsymbol{u}} &= \boldsymbol{D}_x\left(\pi_1 \boldsymbol{D}_x\boldsymbol{u} - \pi_2\boldsymbol{\pi}_{\text{go}} \odot \boldsymbol{u} \odot \boldsymbol{D}_x\boldsymbol{v}\right)\\ &\quad + \boldsymbol{\pi}_{\text{gr}} \odot \boldsymbol{u} \odot (1-\boldsymbol{u}) - \pi_3\boldsymbol{\pi}_{\text{d}} \odot \boldsymbol{u},\end{aligned} \tag{31a}$$

$$\dot{\boldsymbol{v}} = \boldsymbol{D}_x\left(\pi_4 \boldsymbol{D}_x\boldsymbol{v}\right) - \pi_5\boldsymbol{\pi}_{\text{c}} \odot \boldsymbol{u}, \tag{31b}$$

where $\boldsymbol{D}_x$ is the (discretized) differential operator, we have used the symbol $\odot$ for indicating pointwise multiplication, and $\boldsymbol{\pi}_{\text{go}} = \boldsymbol{\pi}_{\text{go}}(\boldsymbol{v})$, $\boldsymbol{\pi}_{\text{gr}} = \boldsymbol{\pi}_{\text{gr}}(\boldsymbol{v})$, $\boldsymbol{\pi}_{\text{d}} = \boldsymbol{\pi}_{\text{d}}(\boldsymbol{v})$ and $\boldsymbol{\pi}_{\text{c}} = \boldsymbol{\pi}_{\text{c}}(\boldsymbol{v})$ are vector functions that act at each nodal value. The framework presented permits the consideration of functional relationships, that is, the value of $\boldsymbol{\pi}$ at a point x could depend on the value of the field v on the whole computational domain. However, the underlying nature of the *go-or-grow* framework allows us to consider $\pi : x \mapsto \pi(v(x)), x \in [0;1]$, or equivalently, for the vector $\boldsymbol{\pi}$, $\pi_j = F(v_j)$, generating a restriction to sparse graphs for the network architecture. Consequently, we obtain sparse tensors and operators [129]. With respect to the time integration, for each ODE $\dot{\boldsymbol{y}} = \boldsymbol{f}(\boldsymbol{y})$, the system is discretized using forward Euler method, that is, $\boldsymbol{y}(t+\Delta t) - \boldsymbol{y}(t) \simeq \Delta t \boldsymbol{f}(\boldsymbol{y}(t))$. This approximation leads to the discretization:

$$\boldsymbol{y}^{n+1} = \boldsymbol{y}^n + (\Delta t)\boldsymbol{f}(\boldsymbol{y}^n).$$

where $\boldsymbol{y} = (\boldsymbol{u}, \boldsymbol{v})$ and $\boldsymbol{f} = (\boldsymbol{f}_1, \boldsymbol{f}_2)$ with

$$\begin{aligned}\boldsymbol{f}_1(\boldsymbol{u},\boldsymbol{v}) &= \boldsymbol{D}_x\left(\pi_1\boldsymbol{D}_x\boldsymbol{u} - \pi_2\boldsymbol{\pi}_{\text{go}} \odot \boldsymbol{u} \odot \boldsymbol{D}_x\boldsymbol{v}\right) + \boldsymbol{\pi}_{\text{gr}} \odot \boldsymbol{u} \odot (1-\boldsymbol{u}) - \pi_3\boldsymbol{\pi}_{\text{d}} \odot \boldsymbol{u},\\ \boldsymbol{f}_2(\boldsymbol{u},\boldsymbol{v}) &= \boldsymbol{D}_x\left(\pi_4\boldsymbol{D}_x\boldsymbol{v}\right) - \pi_5\boldsymbol{\pi}_{\text{c}} \odot \boldsymbol{u}.\end{aligned}$$

Considering $\boldsymbol{Y} = \boldsymbol{y}^{n+1}$, $\boldsymbol{X} = \boldsymbol{y}^n$ and $\boldsymbol{\Pi} = (\boldsymbol{\pi}_{\text{go}}, \boldsymbol{\pi}_{\text{gr}}, \boldsymbol{\pi}_{\text{d}}, \boldsymbol{\pi}_{\text{c}})$, the time-discretized version of Eqs. (31) may be formulated as a PGNNIV problem

$$\boldsymbol{Y} = \mathsf{Y}(\boldsymbol{X}, \boldsymbol{\Pi}), \tag{32a}$$

$$\boldsymbol{\Pi} = \mathsf{H}(\boldsymbol{X}). \tag{32b}$$

where Y is the predictive network and H the explanatory network. For a general PGNNIV problem, Y is an ANN that may need to be fitted [43, 44], but for this particular problem it is possible to define

$$\mathsf{Y}(\boldsymbol{X}, \boldsymbol{\Pi}) = \boldsymbol{Y}(\boldsymbol{X}, \boldsymbol{\Pi}) = \boldsymbol{X} + (\Delta t)\boldsymbol{f}(\boldsymbol{X}), \tag{33}$$

where the function $\boldsymbol{f}$ is known and can be encoded in a ANN structure using standard ANN operators, such as tensor pointwise multiplication and division, scalar multiplication and convolution with known filters [132].

5.2 Learning Cell Response to Hypoxia

The PGNNIV approach has demonstrated utility in learning the metabolic switch between proliferative and

Table 8 Different functional relationships for the *go-or-grow* behavior

Migration activation	Model parameters	Model name
$\pi_{\text{go}}(v) = \mathbb{I}_{[0;\theta]}(v)$	θ	Heaviside, Binary step
$\pi_{\text{go}}(v) = \left(1 - \frac{v}{\theta}\right)\mathbb{I}_{[0;\theta]}(x)$	θ	Piecewise linear
$\pi_{\text{go}}(v) = \frac{v}{v+\theta}$	θ	Michaelis-Menten
$\pi_{\text{go}}(v) = \frac{1}{2}\left(1 - \tanh\left(\frac{v-\theta}{\Delta\theta}\right)\right)$	$\theta, \Delta\theta$	Sigmoid

The different functions include features such as different smoothness and nonlinearities

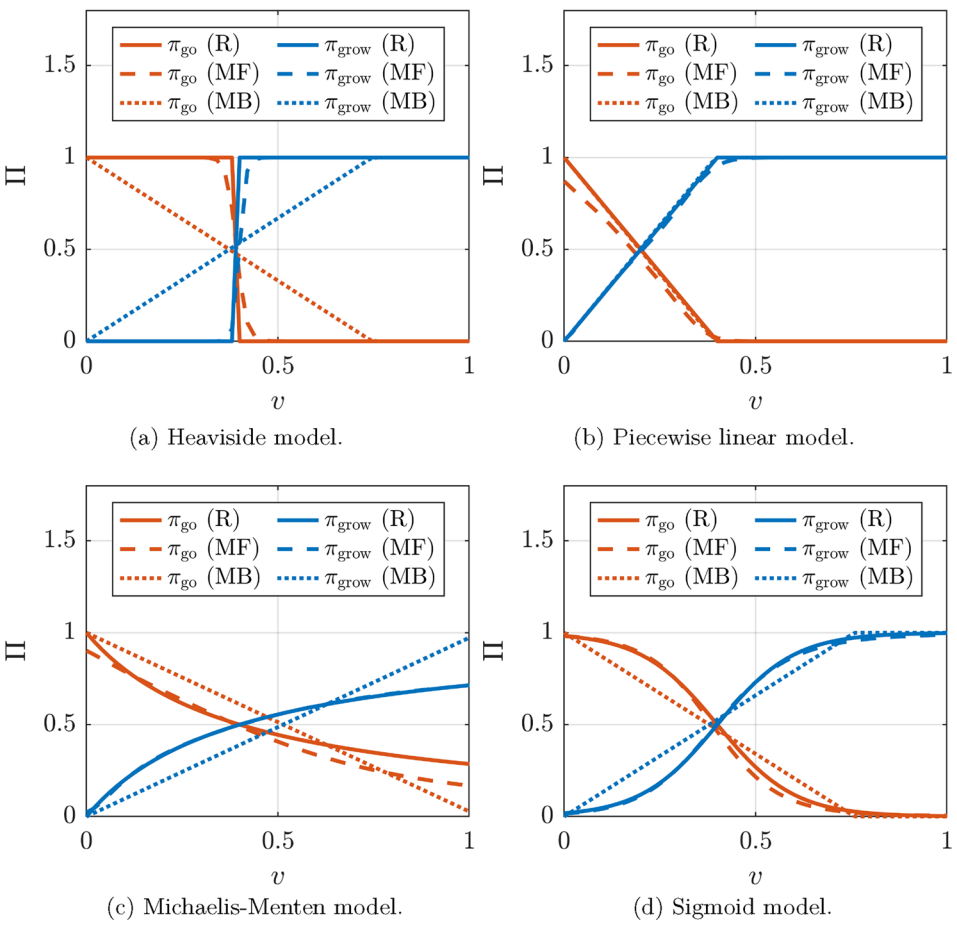
migratory activity under hypoxic conditions [68, 132]. Indeed, assuming as in Sect. 2.2 that $\pi_3 \ll 1$, and $\pi_c(v) = \frac{v}{v+\kappa}$, the explanatory capacity of the PGNNIV has been able to unravel the structure of the function π_{go} and π_{gr} for the different metabolic switches presented in Table 8, where $\pi_{gr}(v) = 1 - \pi_{go}(v)$.

This has been demonstrated using in silico data generation, following these different benchmark models [132], with $\pi_1 = \pi_2 = \pi_3 = 1$, $\pi_4 = 0.05$ and $\kappa = 2$. The dataset is publicly available in Ref. [133]. Regarding the different benchmark models, the values have been set to $\theta = \Delta\theta = 2$ as defined in Table 8. Note that in Ref. [132], the model presented has more parameters, but these extra parameters may be bypassed by an appropriate selection of the dimensional variables (characteristic length, characteristic time, characteristic cell density and characteristic oxygen level). Also, many different experimental configurations are recreated in order to ensure data variability and data representation of all the relevant phenomena. The details of the different experimental configurations and the training procedure are explained in Ref. [132].

In Fig. 16 we can see the capacity of the explanatory network for unveiling the metabolic switch for the four different models presented in Table 8. The results using PGNNIV (Model-Free approach) are also compared to the ones that would have been obtained by prescribing a piecewise linear model (Model-Based approach) for the *go-or-grow* switch, as was presented in Sect. 3.

5.3 Predicting GBM Evolution

The correct characterization of the *go-or-grow* switch has important consequences from the prediction point of view. Once the model $\mathbf{\Pi}$ has been learned, it is represented by the network topology together with all the network parameters (weights and biases). Therefore it may be encapsulated as a one input - two output *black-box* and inserted in any numerical integration scheme. It has been shown that the use of PGNNIV for encapsulating the model $v \mapsto (\Pi_{go}, \Pi_{gr})$ improves the accuracy of GBM progression prediction for certain configurations where hypoxia plays a major role in the dynamics [132]. Even if a prescribed and then calibrated parametric model is able to reproduce accurately the cell culture progression under a fully hypoxic and a fully normoxic



(a) Heaviside model. (b) Piecewise linear model.

(c) Michaelis-Menten model. (d) Sigmoid model.

Fig. 16 Explanatory capacity of the PGNNIV approach. Learned Π_{go} and Π_{gr} functions for four different in silico datasets generated with the four different benchmark models tested. The adjusted piecewise linear model is also shown in each case. R: real model; MF: Model-free approach; MB: Model-based approach

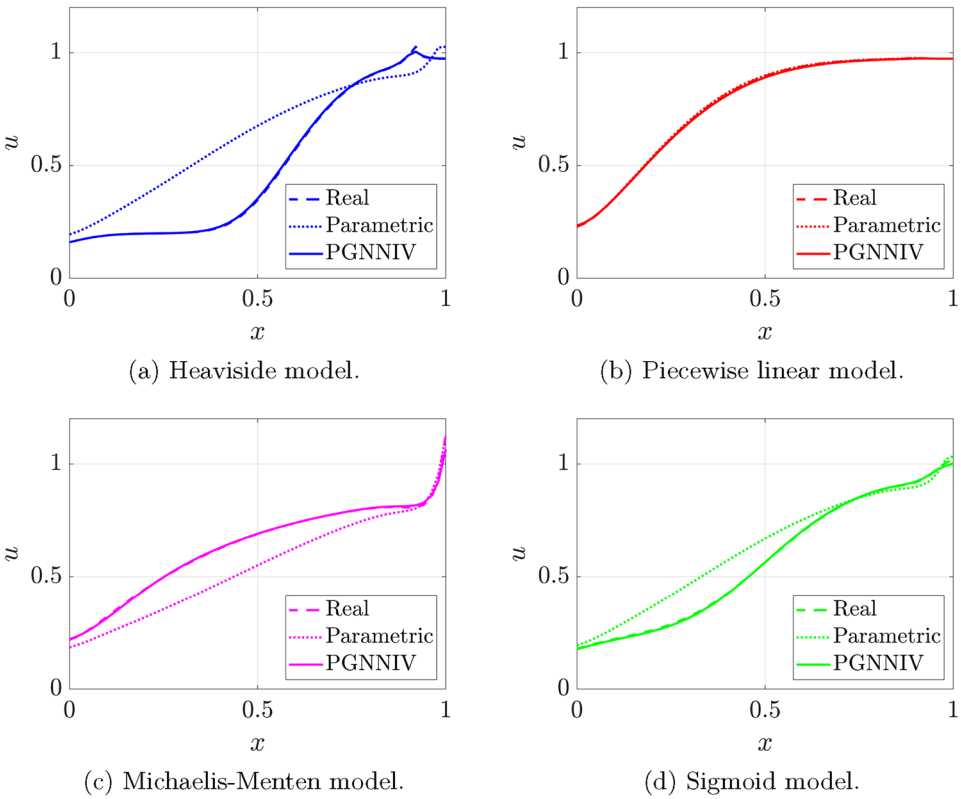


Fig. 17 Predictive capacity of the PGNNIV approach. Final dimensionless cell profile estimated using the PGNNIV approach for four different in silico data-sets generated with the four different benchmark models tested. The profile obtained using a parametric fit on the data, using the piecewise linear model is also shown in each case

environment, PGNNIV improves the prediction error for configurations with oxygen gradients, as it can be seen in Fig. 17. Indeed, the specific features of the model have a greater impact for oxygen levels in the transition between normoxic and hypoxic behavior since it is in this case when the differences between the different models has a greater impact on the cell evolution.

6 Discussion

6.1 Overview of the Different Methods

We have seen that many different approaches have been used to analyze cell cultures in microfluidic devices, ranging from conventional parametric models, to modern approaches under the scope of PIML methods. The degree of knowledge on the cell culture evolution process and the data availability has been exploited in different levels, as illustrated in Fig. 18:

1. In Sect. 2, we presented how GBM progression in microfluidic devices may be described using closed mathematical models. That is, the functions Π_{go}, Π_{gr}, Π_d and Π_c are fixed, which may be represented via $\boldsymbol{\Pi}(S) = \boldsymbol{F}(S)$ where $\boldsymbol{F}$ is a known vector function and $\boldsymbol{\Pi} = (\Pi_{go}, \Pi_{gr}, \Pi_d, \Pi_c)$.

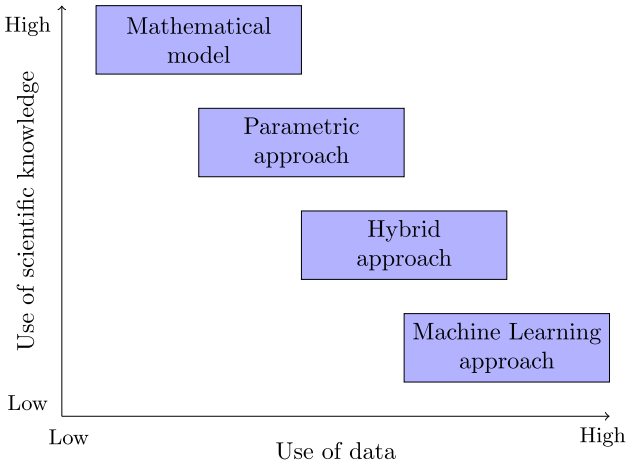


Fig. 18 Classification of the different approaches to GBM in terms of physics and data usage. The approaches proposed in each section use in a different manner the data availability and the existing knowledge

2. In Sect. 3, the ignorance about the model is reduced to the value of selected model parameters that are fitted using experimental data. The functions Π_{go}, Π_{gr}, Π_{d} and Π_{c} are replaced by parametric dependencies to be calibrated, that is $\boldsymbol{\Pi}(S) = \boldsymbol{F}(S;\boldsymbol{\theta})$. This fitting may be deterministic, obtaining an estimation of the parameters, $\hat{\boldsymbol{\theta}}$, or probabilistic, having $\boldsymbol{\Theta}$ as a random vector representing the model parameters. For probabilistic approaches more data would be generally needed.
3. In Sect. 5, the knowledge of the parametric structure is relaxed by allowing a non-parametric[2] dependence for the *go-or-grow* switch, at the cost of needing more data for accurately characterizing this transition. In that case the functions Π_{go}, Π_{gr}, Π_{d} and Π_{c} are replaced by ML models, e.g. ANN models, represented via $\boldsymbol{\Pi}(S) = \mathbf{F}(S)$, where $\mathbf{F}$ is an ANN.
4. In Sect. 4, the model is assumed as perfectly known but the values of the parameters are inferred directly from the data (fluorescence microscopy images) using Deep Learning techniques. Using the previous notations, what is obtained is a function that relates a new piece of data with its associated parameters $\boldsymbol{\theta}$, which is analogous to the previously presented concepts but using Computer Vision and Deep Learning techniques instead of classical fitting.

It has been demonstrated that, even if the parametric approaches are able to reproduce GBM progression accurately, non-parametric approaches are more flexible when the metabolic switch is unknown, inducing a better prognosis of cell culture evolution [68]. In that sense, Table 9 summarizes the strengths and weaknesses of each approach and particularizes a specific suitable context for each of them.

6.2 Open Possibilities

Discovery of hidden cell metabolisms is of major interest in systems biology. Indeed, unraveling the changes of the cell behavior when exposed to different stimuli can focus attention on the mechanisms driving different cell signaling pathways [134, 135]. As a result, parameters such as the hypoxic threshold S_H are replaced by richer non-parametric behaviors.

Moreover, from a mechanistic perspective, different metabolic pathways may be tested in silico using computational approaches [136], to assess whether a candidate pathway is compatible with the metabolic changes discovered by means of PGNNIV. Therefore, this knowledge on the macroscopic cell metabolic behavior at the population level is important not only from an epistemic point of view, for modeling purposes, but also as a promising tool for molecular biologists, in their attempt to isolate and define the different signaling pathways, thus providing a deeper understanding of the underlying mechanisms [137]. Indeed, one of the main concerns about GBM is tumor heterogeneity. The frameworks presented here are able to address this problem at its core. Regarding inter-patient heterogeneity, both the data-driven and the hybrid approaches are particularly suitable for that purpose. The data-driven approach allows the characterization of each cell culture (from a particular patient) with its own model parameters. This heterogeneity is even better captured using non-parametric models for unveiling different metabolic pathways. With respect to spatial heterogeneity, the methods presented incorporate heterogeneous oxygen and cell distributions in space which in turn affects cell behavior. In addition, if the aim is to incorporate spatial parameter dependence, the presented models are limited to homogeneous behavior (model parameters or model ANN components are assumed to be space-independent), but

Table 9 Summary of the strengths and weaknesses of each of the presented approaches

	Strengths	Limitations	Example of application
Known metabolism (mathematical model)	No data requirements, Fast	Needs for validation	Prognosis when no uncertainty on the model parameters
Adjusted metabolism (parametric approach)	Robust, Interpretable	Computationally demanding	Design of experiments
Unveiled metabolism (hybrid approach)	Flexible, Patient specific	Data demanding, Lack of interpretability	Basic research and drug discovery
Assigned metabolism (ML approach)	Non-intrusive, Real time	Lower prediction capacity	Therapy targeting for personalized medicine

We present a specific situation in which each one of the presented approaches is particularly useful, but there may be many others

[2] Actually, the parameters are the ones related with the ANN, but with the term "non-parametric" we refer to models that do not have prescribed structural parameters.

their adaptation to more complex scenarios incorporating spatial heterogeneity is possible. In particular, hybrid approaches using PGNNIV have demonstrated suitability for incorporating spatial heterogeneity [129].

Sometimes, there are some energetic constraints, directly derived from the first and second laws of thermodynamics, that restrict the accessible states in a biological system [138–140]. These constraints are translated into macroscopic ones in a continuum population model. We have seen that a common constraint is to assume that $\Pi_{gr} + \Pi_{go} = 1$. However, this is a special case of a more general constraint $G(\Pi_{gr}, \Pi_{go}) = 0$ for a given function G, which could be founded on an energetic argument about the resources available for the cell to grow or proliferate. All these extra constraints may be incorporated in the PGNNIV framework by expressing some relational equations between variables explicitly, or by adding appropriate penalty terms [44]. As in any ML approach, care must be taken when interpreting the results and deriving conclusions, as the learning methods can suffer from overfitting. A suitable strategy (train and test approaches, cross-validation, validation trials and so forth) is therefore crucial for drawing generalizable conclusions.

6.3 Model Limitations

It has been shown that a number of different models, while following different approaches, are able to reproduce GBM cell culture experiments in microfluidic devices with reasonable accuracy. However, there is still room for improvement. We acknowledge that microfluidic experiments do not fully replicate the in vivo tumor microenvironment and have inherent limitations such as the limited duration of experiments, shorter than tumor development in patients. Nevertheless, these devices offer a reasonably biomimetic platforms to study tumor behavior under controlled conditions. The models presented in this work have been designed to capture the mechanisms driving this behavior rather than to replicate the complexity of in vivo progression. Future work could focus on adapting and fitting these models to clinical datasets [141] that capture tumor evolution over extended periods. From the modeling point of view, there is a mismatch between predicted and experimental results close to the lateral channels and an overestimation of the live cell population at the center of the chamber experiments with high cell concentration (see Fig. 8). These errors could be explored by including the oxygen flow obstruction effect that may be the result of high density cultures, that is, by specifying $D = D(C_1, C_2)$. However, for such modeling investigation, experimental measurements of the oxygen level are most likely required but are not straightforward.

The inclusion of the effect of the mechanical cues related with the ECM, as well as the mechanical interactions between cells, has been formulated in the general framework presented in Sect. 2, but the particular relations between the mechanical fields and the cell populations, as well as the value of the parameters has not been explored yet due to experimental difficulties when working with hydrogels in microfluidic devices. Indeed, to carry out long-time experiments to ensure that ECM remodeling becomes relevant, or to set up spatially controlled ECM mechanical properties is difficult from a technical point of view. Hence, these aspects are generally under-explored although some attempts have been made [142, 143]. A similar remark applies to heterogeneous thermal or electrical fields in order to characterize the impact of electrotaxis or thermotaxis. In addition, if the measurement space is not large enough and/or the variables are not highly uncorrelated, there may be a serious problem of parameter structural or practical identifiability [144], although this could be detected in the stochastic parameter fitting.

Another model improvement would be to include the effect of history on the cell metabolic behavior. The study of the (reversible or not) changes suffered by a cell due to the action of external stimuli, changing the cell gene expression (phenotype) but not affecting the genetic content (genotype) is known as epigenetics [145] and is a hot topic of current research. The inclusion of epigenetic features in GBM cell behavior has been explored recently using multiphenotypic models [17], damage theory using internal variables [146], or non-local PDEs in the phenotypic variable [147, 148].

6.4 Future Perspectives and Outlook

There is a growing consensus that ML and multiscale mechanistic modeling can naturally complement each other to create robust predictive models integrating knowledge about cell metabolism, to manage ill-posed problems and explore massive design spaces [149]. From the conventional distinction between the bottom-up (phenomenological) and top-down (data-driven) [150] strategies, new hybrid strategies have arisen, taking advantage of the new technological facilities in the cell culture (e. g. [151]) or medical image (e. g. [152]) fields, many of them incorporating elements of PDEs, ML and Bayesian estimation [153].

The combination of microfluidic technologies, mechanistic mathematical modeling and machine learning tools let us glimpse some steps in the direction of personalized medicine [154], enabling the utilization of tissues extracted from different patients to build specific digital twins [155]. Once characterized, the tumoral population may be in silico subjected to different stimuli or therapeutic strategies [156, 157]. The digital twin can even be interrogated with new questions to fine-tune the best treatment for the specific

patient, in a quick and cheap loop, always supervised by the clinician.

7 Conclusion

Microfluidic technologies have opened up enormous possibilities in the field of cell culture monitoring. Thanks to their high throughput character and their ability to recreate biomimetic microenvironments, they can be combined with biomathematical frameworks for the study of tumor progression, towards the development of better preclinical models. The ability to recreate three-dimensional and time-varying environments opens the door to the use of sophisticated mathematical models for characterizing tumor cell culture evolution, ranging from the use of partial differential equations to machine learning tools, for analyzing tumor dynamics. We have demonstrated this claim by applying this principle to the analysis of glioblastoma progression in microfluidic devices.

Depending on the data availability, on the experimental facilities and on the particular scientific goal, the knowledge inclusion and the data assimilation may be implemented in different ways. We have illustrated here that parametric models are the most explanatory and easily interpretable, and the abundance of data can be exploited for obtaining parameter and prediction uncertainties as well as in the design of proper experiments. At the other end of the spectrum, pure Machine Learning approaches are practical, easy to use and have enormous predictive power, although sometimes they cannot give true insights into the underlying biological phenomena that lead to a given outcome. In between these two solutions, there is a whole family of alternatives that attempt to unravel in a more or less specific way the mechanistic nature of the cellular metabolic processes linked to tumor emergence and progression.

In this review paper, we have presented an overview of the different approaches for a very particular problem, but it is clear that most of the tools and ideas can be extrapolated to other types of tumors and experimental setups with a similar structure of mathematical models, whose hypotheses must be adapted according to biological knowledge and available evidence. The combination of the presented technological, mathematical and computational techniques not only facilitates research in areas such as biotechnology and biomedical engineering, but brings the results closer to clinical practice.

Appendix A: Global Sensitivity Analysis of the Parametric Models

Using the Morris Method [69], we have performed a global sensitivity analysis of the parametric models describing glioblastoma (GBM) evolution in microfluidic devices that were presented in Sect. 2.1.2 of the main text..

Formulation of the Parametric Models

Following the global formulation presented in Sect. 2.1.2, we particularize it here for the three models presented (although the 2-phenotype model has already been discussed in the main text, we include it here for completion). In all the cases, the governing equations can be expressed as:

$$\frac{\partial C_i}{\partial t} = \nabla \cdot \left(K_{D,i} \nabla C_i - \sum_{j=1}^{m} K_{C,i,j} \nabla C'_j \right) + C_i F_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} \left(C_i F_{ij} + C_j F_{ji} \right), \tag{A1a}$$

$$\frac{\partial C'_i}{\partial t} = \nabla \cdot \left(K'_{D,i} \nabla C'_i \right) + \sum_{j=1}^{n} C_j F'_{ij}. \tag{A1b}$$

Table 10 Parameters of the 3-phenotype model

Parameter	Symbol	Dimensions
Cell pedesis (normoxic)	P_1	$L^2 \cdot T^{-1}$
Cell pedesis (hypoxic)	P_2	$L^2 \cdot T^{-1}$
Growth characteristic time (normoxic)	τ_1	T
Growth characteristic time (hypoxic)	τ_2	T
Saturation concentration	C_{sat}	L^{-3}
Switch time normoxic - hypoxic	τ_{12}	T
Switch time hypoxic - normoxic	τ_{21}	T
Switch time hypoxic - necrotic	τ_{23}	T
Uptake rate (normoxic)	α_1	$Y \cdot L^3 \cdot T^{-1}$
Uptake rate (hypoxic)	α_2	$Y \cdot L^3 \cdot T^{-1}$
Oxygen diffusion rate	D	$L^2 \cdot T^{-1}$
Hypoxic threshold	S_{H}	Y
Anoxic threshold	S_{A}	Y
Michaelis-Menten constant	k_{m}	Y

The parameters of the model are summarized together with their dimensions. L represents length, T time and Y the physical magnitude representing oxygen level

3-Phenotype Model [65]

This model describes the evolution of normoxic C_1, hypoxic C_2 and necrotic C_3 cell concentrations as well as oxygen concentration $S = C_1'$. The particular terms in governing equations are expressed as:

$$
\begin{aligned}
&K_{D,1} = P_1, && K_{C,1,1} = 0, && F_1 = \frac{1}{\tau_1}\left(1 - \frac{C_1 + C_2 + C_3}{C_{\text{sat}}}\right),\\
&K_{D,2} = P_2, && K_{C,2,1} = 0, && F_2 = \frac{1}{\tau_2}\left(1 - \frac{C_1 + C_2 + C_3}{C_{\text{sat}}}\right),\\
&K_{D,3} = 0, && K_{C,3,1} = 0, && F_3 = 0,\\
&F_{12} = \frac{1}{\tau_{12}} s_{12}, && s_{21} = \frac{1}{\tau_{21}} s_{21}, && F_{13} = 0,\\
&F_{31} = 0, && F_{23} = \frac{1}{\tau_{23}} s_{23}, && K'_{D,1} = D,\\
&F'_{11} = \alpha_1 \frac{S}{k_{\text{m}} + S}, && F'_{21} = \alpha_2 \frac{S}{k_{\text{m}} + S}, && F'_{31} = 0,
\end{aligned}
$$

where s_{ij} are the switch functions from phenotype i to j. In particular:

$$
s_{12} = \begin{cases} 1 \text{ if } S \le S_{\text{H}} \\ 0 \text{ if } S > S_{\text{H}} \end{cases}, \quad s_{21} = \begin{cases} 0 \text{ if } S \le S_{\text{H}} \\ 1 \text{ if } S > S_{\text{H}} \end{cases}, \quad s_{23} = \begin{cases} 1 \text{ if } S \le S_{\text{A}} \\ 0 \text{ if } S > S_{\text{A}} \end{cases}.
$$

Table 11 Parameters of the 4-phenotype model

Parameter	Symbol	Dimensions
Cell pedesis (normoxic)	P_1	$L^2 \cdot T^{-1}$
Cell pedesis (hypoxic)	P_2	$L^2 \cdot T^{-1}$
Cell pedesis (malignant)	P_3	$L^2 \cdot T^{-1}$
Chemotaxis coefficient	χ	$L^2 \cdot T^{-1} \cdot Y^{-1}$
Growth characteristic time (normoxic)	τ_1	T
Growth characteristic time (hypoxic)	τ_2	T
Growth characteristic time (malignant)	τ_3	T
Saturation concentration	C_{sat}	L^{-3}
Switch time normoxic - hypoxic	τ_{12}	T
Switch time hypoxic - malignant	τ_{23}	T
Switch time malignant - hypoxic	τ_{32}	T
Switch time hypoxic - necrotic	τ_{24}	T
Uptake rate (normoxic)	α_1	$Y \cdot L^3 \cdot T^{-1}$
Uptake rate (hypoxic)	α_2	$Y \cdot L^3 \cdot T^{-1}$
Uptake rate (malignant)	α_3	$Y \cdot L^3 \cdot T^{-1}$
Oxygen diffusion rate	D	$L^2 \cdot T^{-1}$
Hypoxic threshold	S_{H}	Y
Anoxic threshold	S_{A}	Y
Michaelis-Menten constant	k_{m}	Y

The parameters of the model are summarized together with their dimensions. L represents length, T time and Y the physical magnitude representing oxygen level

The model parameters are summarized in Table 10.

4-Phenotype Model [17]

This model adds a malignant phenotype and, hence, considers the evolution of normoxic C_1, hypoxic C_2, malignant C_3 and

Table 12 Parameters of the 2-phenotype model

Parameter	Symbol	Dimensions
Cell pedesis	P	$L^2 \cdot T^{-1}$
Chemotaxis coefficient	χ	$L^2 \cdot T^{-1} \cdot Y^{-1}$
Growth characteristic time	τ_{gr}	T
Saturation concentration	C_{sat}	L^{-3}
Death characteristic time	τ_{d}	T
Uptake rate	α	$Y \cdot L^3 \cdot T^{-1}$
Oxygen diffusion rate	D	$L^2 \cdot T^{-1}$
Hypoxic threshold	S_{H}	Y
Anoxic threshold	S_{A}	Y
Anoxia sensitivity	ΔS_{A}	Y
Michaelis-Menten constant	k_{m}	Y

The parameters of the model are summarized together with their dimensions. L represents length, T time and Y the physical magnitude representing oxygen level

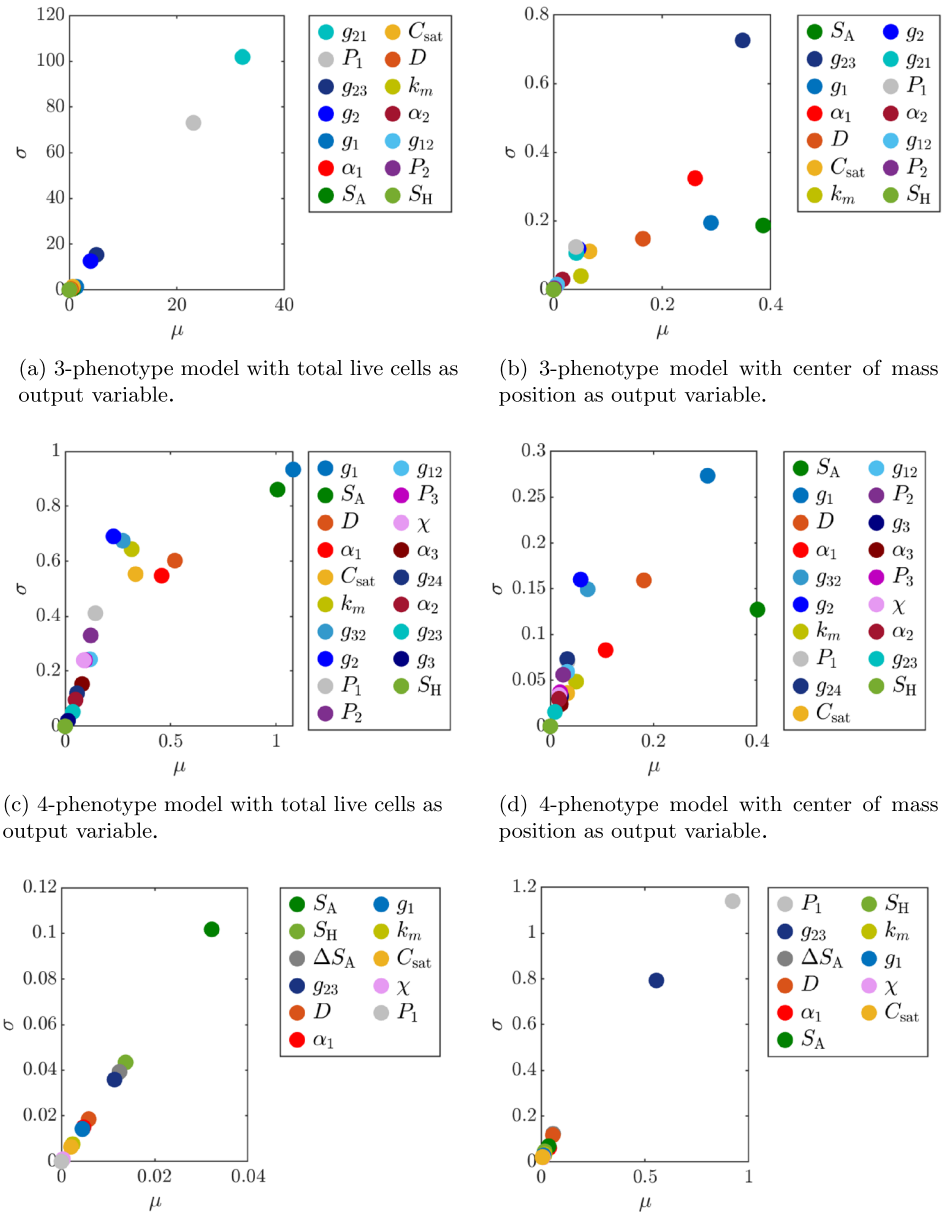
necrotic C_4 cell concentrations together with oxygen concentration $S = C_1'$. In this case, the particular terms of governing equations are expressed as:

$$
\begin{aligned}
&K_{D,1} = P_1, \quad K_{C,1,1} = 0, \quad F_1 = \frac{1}{\tau_1}\left(1 - \frac{C_1 + C_2 + C_3 + C_4}{C_{\text{sat}}}\right),\\
&K_{D,2} = P_2, \quad K_{C,2,1} = \chi, \quad F_2 = \frac{1}{\tau_2}\left(1 - \frac{C_1 + C_2 + C_3 + C_4}{C_{\text{sat}}}\right),\\
&K_{D,3} = P_3, \quad K_{C,3,1} = 0, \quad F_3 = \frac{1}{\tau_3}\left(1 - \frac{C_1 + C_2 + C_3 + C_4}{C_{\text{sat}}}\right),\\
&K_{D,4} = 0, \quad K_{C,4,1} = 0, \quad F_4 = 0,\\
&F_{12} = \frac{1}{\tau_{12}} S_{12}, \quad F_{21} = 0, \quad F_{13} = 0,\\
&F_{31} = 0, \quad F_{14} = 0, \quad F_{41} = 0,\\
&F_{23} = \frac{1}{\tau_{23}} S_{23}, \quad F_{32} = \frac{1}{\tau_{32}} S_{32}, \quad F_{24} = \frac{1}{\tau_{24}} S_{24},\\
&F_{34} = 0, \quad F_{43} = 0, \quad K'_{D,1} = D,\\
&F'_{11} = \alpha_1 \frac{S}{k_{\text{m}} + S}, \quad F'_{21} = \alpha_2 \frac{S}{k_{\text{m}} + S}, \quad F'_{31} = \alpha_3 \frac{S}{k_{\text{m}} + S}\\
&F'_{41} = 0,
\end{aligned}
$$

Table 13 Ranges of the parameters in the sensitivity analysis

Parameter	Symbol	Min. value	Max. value	Units	Models
Cell pedesis (normoxic)	P_1	$1 \cdot 10^{-13}$	$1 \cdot 10^{-10}$	$\text{cm}^2 \cdot \text{s}^{-1}$	2p, 3p, 4p
Cell pedesis (hypoxic)	P_2	$1 \cdot 10^{-12}$	$1 \cdot 10^{-9}$	$\text{cm}^2 \cdot \text{s}^{-1}$	3p, 4p
Cell pedesis (malignant)	P_3	$1 \cdot 10^{-12}$	$1 \cdot 10^{-9}$	$\text{cm}^2 \cdot \text{s}^{-1}$	4p
Chemotaxis coefficient	χ	$1 \cdot 10^{-9}$	$1 \cdot 10^{-8}$	$\text{cm}^2 \cdot \text{s}^{-1} \cdot \text{mmHg}^{-1}$	2p, 4p
Growth rate (normoxic)	g_1	$1 \cdot 10^{-7}$	$1 \cdot 10^{-5}$	s	2p, 3p, 4p
Growth rate (hypoxic)	g_2	$1 \cdot 10^{-7}$	$1 \cdot 10^{-5}$	s	3p, 4p
Growth rate (malignant)	g_3	$1 \cdot 10^{-7}$	$1 \cdot 10^{-5}$	s	4p
Saturation concentration	C_{sat}	$5 \cdot 10^{6}$	$5 \cdot 10^{8}$	$\text{cell} \cdot \text{cm}^{-3}$	2p, 3p, 4p
Switch rate normoxic - hypoxic	$1/\tau_{12}$	$3 \cdot 10^{-5}$	$3 \cdot 10^{-4}$	s	3p, 4p
Switch rate hypoxic - normoxic	$1/\tau_{21}$	$1 \cdot 10^{-6}$	$1 \cdot 10^{-5}$	s	3p
Switch rate hypoxic - malignant	$1/\tau_{23}$	$1 \cdot 10^{-6}$	$1 \cdot 10^{-5}$	s	4p
Switch rate malignant - hypoxic	$1/\tau_{32}$	$1 \cdot 10^{-6}$	$1 \cdot 10^{-5}$	s	4p
Switch time hypoxic - necrotic	τ_{24}	$3 \cdot 10^{-5}$	$3 \cdot 10^{-4}$	s	2p, 3p, 4p
Uptake rate (normoxic)	α_1	$5 \cdot 10^{-10}$	$1 \cdot 10^{-8}$	$\text{cell} \cdot \text{mmHg} \cdot \text{cm}^3 \cdot \text{s}^{-1}$	2p, 3p, 4p
Uptake rate (hypoxic)	α_2	$5 \cdot 10^{-10}$	$1 \cdot 10^{-8}$	$\text{cell} \cdot \text{mmHg} \cdot \text{cm}^3 \cdot \text{s}^{-1}$	3p, 4p
Uptake rate (malignant)	α_3	$1 \cdot 10^{-10}$	$5 \cdot 10^{-9}$	$\text{cell} \cdot \text{mmHg} \cdot \text{cm}^3 \cdot \text{s}^{-1}$	4p
Oxygen diffusion rate	D	$1 \cdot 10^{-9}$	$5 \cdot 10^{-8}$	$\text{cm}^2 \cdot \text{s}^{-1}$	2p, 3p, 4p
Hypoxic threshold	S_{H}	2	10	mmHg	2p, 3p, 4p
Anoxic threshold	S_{A}	0.1	2	mmHg	2p, 3p, 4p
Anoxia sensitivity	ΔS_{A}	0.01	10	mmHg	2p
Michaelis-Menten constant	k_{m}	1	5	mmHg	2p, 3p, 4p

The minimum and maximum values of the parameters of all the models are displayed, together with the models in which each parameter appears (2p is the 2-phenotype model, 3p the 3-phenotype model and 4p the 4-parameter model). Note that the switch time from hypoxic to normoxic phenotypes is equivalent to the death characteristic time in the 2p model



(a) 3-phenotype model with total live cells as output variable.

(b) 3-phenotype model with center of mass position as output variable.

(c) 4-phenotype model with total live cells as output variable.

(d) 4-phenotype model with center of mass position as output variable.

(e) 2-phenotype model with total live cells as output variable.

(f) 2-phenotype model with center of mass position as output variable.

Fig. 19 Results of the Morris sensitivity analysis of the models. The figures depict the importance of the model parameters in terms of the morris sensitivity indices (μ and σ)

where, again, s_{ij} are phenotype switch functions:

2-Phenotype Model [67]

$$s_{12} = \begin{cases} 1 \text{ if } S \leq S_H \\ 0 \text{ if } S > S_H \end{cases}, \quad s_{23} = \begin{cases} 0 \text{ if } S \leq S_H \\ 1 \text{ if } S > S_H \end{cases}, \quad s_{23} = \begin{cases} 1 \text{ if } S \leq S_H \\ 0 \text{ if } S > S_H \end{cases}, \quad s_{24} = \begin{cases} 1 \text{ if } S \leq S_A \\ 0 \text{ if } S > S_A \end{cases}.$$

Table 11 provides a summary of the model parameters.

This model only considers two phenotypes: normoxic C_1 and necrotic C_2 as well as oxygen concentration $S = C_1'$. The terms of the governing equations are given by:

$$K_{D,1} = P, \qquad K_{C,1,1} = \chi \Pi_{go} F_{go}, \qquad F_1 = \frac{1}{\tau_{gr}} \Pi_{gr} \left(1 - \frac{C_1 + C_2}{C_{sat}}\right),$$

$$F_{12} = \frac{1}{\tau_d} \frac{1}{2} \left(1 - \tanh\left(\frac{S - S_A}{\Delta S_A}\right)\right), \qquad F_{21} = 0, \qquad K_{D,2} = 0,$$

$$K_{C,2,1} = 0, \qquad F_2 = 0, \qquad K'_{D,1} = D,$$

$$F'_{11} = \alpha \frac{S}{k_m + S}, \qquad F'_{21} = 0.$$

In this case, the activation functions are:

$$\pi_{go} = \begin{cases} 1 - S/S_H & \text{if } S \le S_H \\ 0 & \text{if } S > S_H \end{cases}, \qquad \pi_{go} = \begin{cases} S/S_H & \text{if } S \le S_H \\ 1 & \text{if } S > S_H \end{cases},$$

$$F_{gr} = \begin{cases} 1 - C_1/C_{sat} & \text{if } C_1 \le C_{sat} \\ 0 & \text{if } C_1 > C_{sat} \end{cases},$$

and the parameters are summarized in Table 12.

Overview of the Morris Method

The Morris method, also known as the Elementary Effect (EE) method, utilizes a discrete approximation of the average value of the partial derivatives over the input space. The method relies on a one-factor-at-a-time (OAT) experimental design where the effects of a single factor on the output are assessed sequentially. The individual randomized experimental designs are known as trajectories. Each trajectory consists of $\alpha + 1$ points, where α is the number of input variables analyzed and only one input variable changes at each step by a fixed step size, Δ. The elementary effect (EE_i) for each input variable is then computed by comparing the model output at two consecutive points:

$$EE_i(\boldsymbol{\theta}) = \frac{\hat{Q}(\boldsymbol{\theta} + \Delta \boldsymbol{e}_i) - \hat{Q}(\boldsymbol{\theta})}{\Delta}, \tag{A2}$$

with $\boldsymbol{\theta}$ the vector of input variables and $\boldsymbol{e}_i$ is the unit vector in the direction of the i-th axis for $i = 1 \ldots, \alpha$. Finally, the distribution of elementary effects is summarized using the mean effect μ_i , indicating overall influence, and σ_i measuring non-linearity and interactions. The method's main advantage is the lack of assumptions on the inputs and functional dependency of the output to inputs such as monotonicity or linearity.

The analysis was carried out for the three models presented above, considering all their parameters and taking 10 trajectories. It is worth noting for the characteristic times, the analysis has been carried out taking the inverse parameter, that is, $g_i = 1/\tau_i$ The ranges of the parameters are displayed in Table 13.

We analyzed the sensitivity of the parameters with respect to two output variables: the total number of live cells N and the position of center of mass of the cell distribution x_{CM}, both calculated at the end of a pseudopalisade simulation. The concentration of live cells can be obtained as:

$$C_{live} = \sum_{i=1}^{m-1} C_i,$$

and then we can integrate this variable to obtain the total number of cells (per unit area)

$$N = \int_0^L C_{live}\, dx.$$

The position of the center of mass is obtained as:

$$x_{CM} = \frac{\int_0^L x \cdot C_{live}\, dx}{\int_0^L C_{live}\, dx}.$$

To enable a more straightforward comparison, both magnitudes are normalized with respect to the saturation concentration C_{sat} and the domain length L respectively.

Results

Figure 19 displays the results of the sensitivity analysis for the three models and the two computed metrics. Parameters with high influence on the output variables are located in the upper right corner of the plots, having both high mean elementary effect μ (measuring the general importance of the parameter in the model) and high standard deviation of the elementary effect σ (measuring nonlinear effects of the parameter).

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. The authors would like to acknowledge the financial support from the CDTI project, Chineka (EXP 00149754 / IDI-20220819) as well as Financial support from the GBM_IMMUNE project (PID2021-126051OB-C41:/AEI/10.13039/501100011033/

FEDER, UE) funded by the Ministry of Science and Innovation, the state administration agency (AEI), and the European Regional Development Fund and project PID2022-138572OB-C44, funded by MCIN/AEI/10.13039/501100011033/FEDER, UE. We also acknowledge the support from the Government of Aragon, Spain (group DGA-T62_23R).

Declarations

Conflict of interests The authors declare that they have no Conflict of interest.

Ethical Approval Not applicable.

References

1. Tan AC, Ashley DM, López GY, Malinzak M, Friedman HS, Khasraw M (2020) Management of glioblastoma: state of the art and future directions. CA Cancer J Clin 70(4):299–312
2. Delgado-López P, Corrales-García E (2016) Survival in glioblastoma: a review on the impact of treatment modalities. Clin Transl Oncol 18(11):1062–1071
3. Ostrom QT, Gittleman H, Farah P, Ondracek A, Chen Y, Wolinsky Y, Stroup NE, Kruchko C, Barnholtz-Sloan JS (2013) Cbtrus statistical report: primary brain and central nervous system tumors diagnosed in the united states in 2006–2010. Neuro Oncol 15(suppl–2):1–56
4. Brat DJ, Castellano-Sanchez AA, Hunter SB, Pecot M, Cohen C, Hammond EH, Devi SN, Kaur B, Van Meir EG (2004) Pseudopalisades in glioblastoma are hypoxic, express extracellular matrix proteases, and are formed by an actively migrating cell population. Can Res 64(3):920–927
5. Rong Y, Durden DL, Van Meir EG, Brat DJ (2006) 'pseudopalisading'necrosis in glioblastoma: a familiar morphologic feature that links vascular pathology, hypoxia, and angiogenesis. J Neuropathol Exp Neurol 65(6):529–539
6. Brat DJ, Van Meir EG (2004) Vaso-occlusive and prothrombotic mechanisms associated with tumor hypoxia, necrosis, and accelerated growth in glioblastoma. Lab Invest 84(4):397
7. Dapash M, Hou D, Castro B, Lee-Chang C, Lesniak MS (2021) The interplay between glioblastoma and its microenvironment. Cells 10(9):2257
8. Jalili-Firoozinezhad S, Miranda CC, Cabral JM (2021) Modeling the human body on microfluidic chips. Trends Biotechnol 39(8):838–852
9. Sackmann EK, Fulton AL, Beebe DJ (2014) The present and future role of microfluidics in biomedical research. Nature 507(7491):181–189
10. Shin Y, Han S, Jeon JS, Yamamoto K, Zervantonakis IK, Sudo R, Kamm RD, Chung S (2012) Microfluidic assay for simultaneous culture of multiple cell types on surfaces or within hydrogels. Nat Protoc 7(7):1247–1259
11. Guckenberger DJ, Groot TE, Wan AM, Beebe DJ, Young EW (2015) Micromilling: a method for ultra-rapid prototyping of plastic microfluidic devices. Lab Chip 15(11):2364–2378
12. Jiménez-Torres JA, Peery SL, Sung KE, Beebe DJ (2016) Lumenext: a practical method to pattern luminal structures in ecm gels. Adv Healthcare Mater 5(2):198–204
13. Boussommier-Calleja A, Li R, Chen MB, Wong SC, Kamm RD (2016) Microfluidics: a new tool for modeling cancer-immune interactions. Trends Cancer 2(1):6–19
14. Zervantonakis IK, Hughes-Alford SK, Charest JL, Condeelis JS, Gertler FB, Kamm RD (2012) Three-dimensional microfluidic model for tumor cell intravasation and endothelial barrier function. Proc Natl Acad Sci 109(34):13515–13520
15. Jeon JS, Bersini S, Gilardi M, Dubini G, Charest JL, Moretti M, Kamm RD (2015) Human 3d vascularized organotypic microfluidic assays to study breast cancer cell extravasation. Proc Natl Acad Sci 112(1):214–219
16. Bersini S, Jeon JS, Dubini G, Arrigoni C, Chung S, Charest JL, Moretti M, Kamm RD (2014) A microfluidic 3d in vitro model for specificity of breast cancer metastasis to bone. Biomaterials 35(8):2454–2461
17. Ayuso JM, Monge R, Martínez-González A, Virumbrales-Muñoz M, Llamazares GA, Berganzo J, Hernández-Laín A, Santolaria J, Doblaré M, Hubert C et al (2017) Glioblastoma on a microfluidic chip: generating pseudopalisades and enhancing aggressiveness through blood vessel obstruction events. Neuro Oncol 19(4):503–513
18. Ayuso JM, Virumbrales-Muñoz M, Lacueva A, Lanuza PM, Checa-Chavarria E, Botella P, Fernández E, Doblare M, Allison SJ, Phillips RM et al (2016) Development and characterization of a microfluidic model of the tumour microenvironment. Sci Rep 6(1):1–16
19. Hatzikirou H, Deutsch A, Schaller C, Simon M, Swanson K (2005) Mathematical modelling of glioblastoma tumour development: a review. Math Models Methods Appl Sci 15(11):1779–1794
20. Suarez C, Maglietti F, Colonna M, Breitburd K, Marshall G (2012) Mathematical modeling of human glioma growth based on brain topological structures: study of two clinical cases. PLoS ONE 7(6):39616
21. Swanson KR, Rockne RC, Claridge J, Chaplain MA, Alvord EC, Anderson ARA (2011) Quantifying the role of angiogenesis in malignant progression of gliomas: in silico modeling integrates imaging and histology. Cancer Res 71(24):7366–7375. https://doi.org/10.1158/0008-5472.CAN-11-1399
22. Scribner E, Saut O, Province P, Bag A, Colin T, Fathallah-Shaykh HM (2014) Effects of anti-angiogenesis on glioblastoma growth and migration: model to clinical predictions. PLoS ONE 9(12):115018
23. Kim Y, Jeon H, Othmer H (2016) The role of the tumor microenvironment in glioblastoma: A mathematical model. IEEE Trans Biomed Eng 64(3):519–527
24. Rejniak KA (2016) Systems biology of tumor microenvironment: quantitative modeling and simulations, vol 936. Springer, Cham
25. Schuetz TA, Becker S, Mang A, Toma A, Buzug TM (2013) Modelling of glioblastoma growth by linking a molecular interaction network with an agent-based model. Math Comput Model Dyn Syst 19(5):417–433
26. Lima E, Oden J, Hormuth D, Yankeelov T, Almeida R (2016) Selection, calibration, and validation of models of tumor growth. Math Models Methods Appl Sci 26(12):2341–2368

27. Buchlak QD, Esmaili N, Leveque J-C, Bennett C, Farrokhi F, Piccardi M (2021) Machine learning applications to neuroimaging for glioma detection and classification: an artificial intelligence augmented systematic review. J Clin Neurosci 89:177–198
28. Chen Z, Li N, Liu C, Yan S, et al (2022) Deep convolutional neural network-based brain magnetic resonance imaging applied in glioma diagnosis and tumor region identification. Contrast Media Mol Imaging 2022
29. Wang Z, Wang Y, Yang T, Xing H, Wang Y, Gao L, Guo X, Xing B, Wang Y, Ma W (2021) Machine learning revealed stemness features and a novel stemness-based classification with appealing implications in discriminating the prognosis, immunotherapy and temozolomide responses of 906 glioblastoma patients. Brief Bioinform 22(5):032
30. Qi D, Li J, Quarles CC, Fonkem E, Wu E (2023) Assessment and prediction of glioblastoma therapy response: challenges and opportunities. Brain 146(4):1281–1298
31. Bathaee Y (2017) The artificial intelligence black box and the failure of intent and causation. Harv JL Tech 31:889
32. Castelvecchi D (2016) Can we open the black box of AI? Nat News 538(7623):20
33. Lorenzo G, Ahmed SR, Hormuth II DA, Vaughn B, Kalpathy-Cramer J, Solorio L, Yankeelov TE, Gomez H (2024) Patient-specific, mechanistic models of tumor growth incorporating artificial intelligence and big data. Ann Rev Biomed Eng 26
34. Jarrett D, Stride E, Vallis K, Gooding MJ (2019) Applications and limitations of machine learning in radiation oncology. Br J Radiol 92(1100):20190001
35. Cuomo S, Di Cola VS, Giampaolo F, Rozza G, Raissi M, Piccialli F (2022) Scientific machine learning through physics-informed neural networks: where we are and what's next. J Sci Comput 92(3):88
36. Cueto E, Chinesta F (2023) Thermodynamics of learning physical phenomena. Arch Comput Methods Eng 30(8):4653–4666
37. Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L (2021) Physics-informed machine learning. Nat Rev Phys 3(6):422–440
38. Misyris GS, Venzke A, Chatzivasileiadis S (2020) Physics-informed neural networks for power systems. 2020 IEEE Power & Energy Society General Meeting (PESGM). IEEE, pp 1–5
39. Cai S, Mao Z, Wang Z, Yin M, Karniadakis GE (2021) Physics-informed neural networks (PINNS) for fluid mechanics: a review. Acta Mech Sin 37(12):1727–1738
40. Cai S, Wang Z, Wang S, Perdikaris P, Karniadakis GE (2021) Physics-informed neural networks for heat transfer problems. J Heat Transfer 143(6):060801
41. Flaschel M, Kumar S, De Lorenzis L (2021) Unsupervised discovery of interpretable hyperelastic constitutive laws. Comput Methods Appl Mech Eng 381:113852
42. Thakolkaran P, Joshi A, Zheng Y, Flaschel M, De Lorenzis L, Kumar S (2022) Nn-euclid: Deep-learning hyperelasticity without stress data. J Mech Phys Solids 169:105076
43. Ayensa-Jiménez J, Orera-Echeverría J, Doblare M (2025) Predicting and explaining nonlinear material response using deep physically guided neural networks with internal variables. Math Mech Solids 30(2):573–598
44. Ayensa-Jiménez J, Doweidar MH, Sanz-Herrera JA, Doblaré M (2021) Prediction and identification of physical systems by means of physically-guided neural networks with meaningful internal layers. Comput Methods Appl Mech Eng 381:113816
45. Moya B, Badías A, González D, Chinesta F, Cueto E (2023) Computational sensing, understanding, and reasoning: an artificial intelligence approach to physics-informed world modeling. Arch Comput Methods Eng: 1–18
46. Du G, Fang Q, Toonder JM (2016) Microfluidics for cell-based high throughput screening platforms-a review. Anal Chim Acta 903:36–50
47. Perthame B (2006) Transp Equations Biol. Springer, Basel
48. Hadeler KP (2000) Reaction transport equations in biological modeling. Math Comput Model 31(4–5):75–81
49. Murray JD (2002) Mathematical biology: I: an introduction, vol 17. Springer, New York
50. Murray JD (2003) Mathematical biology: II: spatial models and biomedical applications, vol 18. Springer, New York
51. Keller EF, Segel LA (1971) Traveling bands of chemotactic bacteria: a theoretical analysis. J Theor Biol 30(2):235–248
52. Arumugam G, Tyagi J (2021) Keller–Segel chemotaxis models: a review. Acta Appl Math 171(1):1–82
53. Xue C, Hwang HJ, Painter KJ, Erban R (2011) Travelling waves in hyperbolic chemotaxis equations. Bull Math Biol 73(8):1695–1733
54. Ayensa-Jiménez J, Doweidar MH, Doblaré M, Gaffney EA (2023) A mathematical modelling study of chemotactic dynamics in cell cultures: the impact of spatio-temporal heterogeneity. Bull Math Biol 85(10):98
55. Benzekry S, Lamont C, Beheshti A, Tracz A, Ebos JM, Hlatky L, Hahnfeldt P (2014) Classical mathematical models for description and prediction of experimental tumor growth. PLoS Comput Biol 10(8):1003800
56. Foryś U, Marciniak-Czochra A (2003) Logistic equations in tumour growth modelling. Int J Appl Math Comput Sci 13(3):317–325
57. Gompertz B (1825) On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies: in a letter to Francis Baily, ESQ FRS & C. Phil Trans R Soc Lond 24(115):513–583
58. Rocha JL, Aleixo SM (2013) An extension of Gompertzian growth dynamics: Weibull and fréchet models. Math Biosci Eng 10(2):379
59. Richards F (1959) A flexible growth function for empirical use. J Exp Bot 10(2):290–301
60. Tabatabai M, Williams DK, Bursac Z (2005) Hyperbolastic growth models: theory and application. Theor Biol Med Model 2(1):1–13
61. Jarrett AM, Lima EA, Hormuth DA, McKenna MT, Feng X, Ekrut DA, Resende ACM, Brock A, Yankeelov TE (2018) Mathematical models of tumor cell proliferation: a review of the literature. Expert Rev Anticancer Ther 18(12):1271–1286
62. Alfonso J, Talkenberger K, Seifert M, Klink B, Hawkins-Daarud A, Swanson K, Hatzikirou H, Deutsch A (2017) The biology and mathematical modelling of glioma invasion: a review. J R Soc Interface 14(136):20170490
63. Paguirigan AL, Beebe DJ (2008) Microfluidics meet cell biology: bridging the gap by validation and application of microscale techniques for cell biological assays. BioEssays 30(9):811–821
64. Logun M, Zhao W, Mao L, Karumbaiah L (2018) Microfluidics in malignant glioma research and precision medicine. Adv Biosyst 2(5):1700221
65. Martínez-González A, Calvo GF, Romasanta LAP, Pérez-García VM (2012) Hypoxic cell waves around necrotic cores in glioblastoma: a biomathematical model and its therapeutic implications. Bull Math Biol 74(12):2875–2896
66. Hatzikirou H, Basanta D, Simon M, Schaller K, Deutsch A (2012) 'Go or grow': the key to the emergence of invasion in tumour progression? Math Med Biol: J IMA 29(1):49–65
67. Ayensa-Jiménez J, Pérez-Aliacar M, Randelovic T, Oliván S, Fernández L, Sanz-Herrera JA, Ochoa I, Doweidar MH, Doblaré M (2020) Mathematical formulation and parametric analysis of in vitro cell models in microfluidic devices:

application to different stages of glioblastoma evolution. Sci Rep 10(1):1–21
68. Ayensa-Jiménez J (2022) Study of the effect of the tumour microenvironment on cell response using a combined simulation and machine learning approach. application to the evolution of glioblastoma. PhD thesis, Universidad de Zaragoza
69. Morris MD (1991) Factorial sampling plans for preliminary computational experiments. Technometrics 33(2):161–174
70. Pérez-García VM, Calvo GF, Belmonte-Beitia J, Diego D, Pérez-Romasanta L (2011) Bright solitary waves in malignant gliomas. Phys Rev E 84(2):021921
71. Gerlee P, Nelander S (2016) Travelling wave analysis of a mathematical model of glioblastoma growth. Math Biosci 276:75–81
72. Stepien TL, Rutter EM, Kuang Y (2018) Traveling waves of a go-or-grow model of glioma growth. SIAM J Appl Math 78(3):1778–1801
73. Odde DJ (2023) Glioblastoma cell invasion: Go? Grow? Yes. Oxford University Press
74. Skeel RD, Berzins M (1990) A method for the spatial discretization of parabolic equations in one space variable. SIAM J Sci Stat Comput 11(1):1–32
75. Morton KW (2019) Revival: numerical solution of convection-diffusion problems (1996). CRC Press, Boca Raton
76. Stramer B, Mayor R (2017) Mechanisms and in vivo functions of contact inhibition of locomotion. Nat Rev Mol Cell Biol 18(1):43–55
77. Lu X, Kang Y (2010) Hypoxia and hypoxia-inducible factors: master regulators of metastasis. Clin Cancer Res 16(24):5928–5935
78. Carreau A, Hafny-Rahbi BE, Matejuk A, Grillon C, Kieda C (2011) Why is the partial oxygen pressure of human tissues a crucial parameter? small molecules and hypoxia. J Cell Mol Med 15(6):1239–1253
79. Galluzzi L, Vitale I, Aaronson SA, Abrams JM, Adam D, Agostinis P, Alnemri ES, Altucci L, Amelio I, Andrews DW, et al. (2018) Molecular mechanisms of cell death: recommendations of the nomenclature committee on cell death 2018. Cell Death Differentiation: 1
80. Sendoel A, Hengartner MO (2014) Apoptotic cell death under hypoxia. Physiology 29(3):168–176
81. Meister A (1956) Advances in enzymology and related areas of molecular biology, Vol 17. Advances in enzymology—and related areas of molecular biology. Interscience Publishers Inc., New York. http://gen.lib.rus.ec/book/index.php?md5=8c34c4973d8bb6feed125599b81d6bb7
82. Tannock IF (1972) Oxygen diffusion and the distribution of cellular radiosensitivity in tumours. Br J Radiol 45(535):515–524
83. Pogue BW, O'Hara JA, Wilmot CM, Paulsen KD, Swartz HM (2001) Estimation of oxygen distribution in RIF-1 tumors by diffusion model-based interpretation of pimonidazole hypoxia and eppendorf measurements. Radiat Res 155(1):15–25
84. Secomb T, Hsu R, Dewhirst M, Klitzman B, Gross J (1993) Analysis of oxygen transport to tumor tissue by microvascular networks. Int J Radiat Oncol Biol Phys 25(3):481–489
85. Patel AA, Gawlinski ET, Lemieux SK, Gatenby RA (2001) A cellular automaton model of early tumor growth and invasion: the effects of native tissue vascularity and increased anaerobic tumor metabolism. J Theor Biol 213(3):315–331
86. Martínez-González A, Calvo GF, Pérez Romasanta LA, Pérez-García VM (2012) Hypoxic cell waves around necrotic cores in glioblastoma: a biomathematical model and its therapeutic implications. Bull Math Biol 74(12):2875–2896. https://doi.org/10.1007/s11538-012-9786-1. arXiv:1204.3809
87. Tang P-S (1933) On the rate of oxygen consumption by tissues and lower organisms as a function of oxygen tension. Q Rev Biol 8(3):260–274
88. Liepe J, Kirk P, Filippi S, Toni T, Barnes CP, Stumpf MPH (2014) A framework for parameter estimation and model selection from experimental data in systems biology using approximate bayesian computation. Nat Protoc 9(2):439–456
89. Ayensa-Jiménez J, Pérez-Aliacar M, Randelovic T, Sanz-Herrera JA, Doweidar MH, Doblaré M (2020) Analysis of the parametric correlation in mathematical modeling of in vitro glioblastoma evolution using copulas. Mathematics 9(1):27
90. Agosti A, Giverso C, Faggiano E, Stamm A, Ciarletta P (2018) A personalized mathematical tool for neuro-oncology: a clinical case study. Int J Non-Linear Mech
91. Daşu A, Toma-Daşu I, Karlsson M (2003) Theoretical simulation of tumour oxygenation and results from acute and chronic hypoxia. Phys Med Biol 48(17):2829
92. Sklar M (1959) Fonctions de repartition an dimensions et leurs marges. Publ Inst Statist Univ Paris 8:229–231
93. Jaworski P, Durante F, Härdle WK, Rychlik T (2010) Copula theory and its applications: proceedings of the workshop held in Warsaw, Vol. 198, 25–26 Sept 2009. Springer, Berlin, Heidelberg
94. McNeil A, Frey R, Embrechts P (2005) Quantitative risk management: concepts, techniques, and tools. Princeton University Press
95. Meucci A (2011) A new breed of copulas for risk and portfolio management. Risk 24(9):122–126
96. Solari S, Losada M (2011) Non-stationary wave height climate modeling and simulation. J Geophys Res: Oceans 116(C9)
97. Hyndman RJ (1996) Computing and graphing highest density regions. Am Stat 50(2):120–126
98. Fisher RA (1966) The design of experiments, vol 21. Oliver and Boyd, Edinburgh
99. Kreutz C, Timmer J (2009) Systems biology: experimental design. FEBS J 276(4):923–942
100. Faller D, Klingmüller U, Timmer J (2003) Simulation methods for optimal experimental design in systems biology. SIMULATION 79(12):717–725
101. Garre A, Peñalver-Soto JL, Esnoz A, Iguaz A, Fernandez PS, Egea JA (2019) On the use of in-silico simulations to support experimental design: a case study in microbial inactivation of foods. PLoS ONE 14(8):0220683
102. Wang Z, Sheikh H, Lee K, Georgakis C (2018) Sequential parameter estimation for mammalian cell model based on in silico design of experiments. Processes 6(8):100
103. Kontoravdi C, Pistikopoulos EN, Mantalaris A (2010) Systematic development of predictive mathematical models for animal cell cultures. Comput Chem Eng 34(8):1192–1198
104. Chaloner K, Verdinelli I (1995) Bayesian experimental design: a review. Statist. Sci. 10(3):273–304. https://doi.org/10.1214/ss/1177009939
105. Ryan EG, Drovandi CC, McGree JM, Pettitt AN (2016) A review of modern computational algorithms for Bayesian optimal design. Int Stat Rev 84(1):128–154
106. Rainforth T, Foster A, Ivanova DR, Bickford Smith F (2024) Modern Bayesian experimental design. Stat Sci 39(1):100–114
107. Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27(3):379–423
108. Becker AP, Sells BE, Haque SJ, Chakravarti A (2021) Tumor heterogeneity in glioblastomas: from light microscopy to molecular pathology. Cancers 13(4):761
109. Farina E, Nabhen JJ, Dacoregio MI, Batalini F, Moraes FY (2022) An overview of artificial intelligence in oncology. Fut Sci OA 8(4):787
110. Alekya B, Rao S, Pandya HJ (2019) Engineering approaches for characterizing soft tissue mechanical properties: a review. Clin Biomech 69:127–140
111. Cilla M, Pérez-Rey I, Martínez MA, Peña E, Martínez J (2018) On the use of machine learning techniques for the mechanical

characterization of soft biological tissues. Int J Numer Methods Biomed Eng 34(10):3121
112. Liu M, Liang L, Sun W (2019) Estimation of in vivo constitutive parameters of the aortic wall using a machine learning approach. Comput Methods Appl Mech Eng 347:201–217
113. Cabeza-Gil I, Ríos-Ruiz I, Calvo B (2020) Customised selection of the haptic design in c-loop intraocular lenses based on deep learning. Ann Biomed Eng 48(12):2988–3002
114. Eren F, Aslan M, Kanarya D, Uysalli Y, Aydin M, Kiraz B, Aydin O, Kiraz A (2022) Deepcan: a modular deep learning system for automated cell counting and viability analysis. IEEE J Biomed Health Inform 26(11):5575–5583
115. Van Valen DA, Kudo T, Lane KM, Macklin DN, Quach NT, DeFelice MM, Maayan I, Tanouchi Y, Ashley EA, Covert MW (2016) Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. PLoS Comput Biol 12(11):1005177
116. Piotrowski T, Rippel O, Elanzew A, Nießing B, Stucken S, Jung S, König N, Haupt S, Stappert L, Brüstle O et al (2021) Deep-learning-based multi-class segmentation for automated, non-invasive routine assessment of human pluripotent stem cell culture status. Comput Biol Med 129:104172
117. Ao Z, Cai H, Wu Z, Hu L, Nunez A, Zhou Z, Liu H, Bondesson M, Lu X, Lu X et al (2022) Microfluidics guided by deep learning for cancer immunotherapy screening. Proc Natl Acad Sci 119(46):2214569119
118. McIntyre D, Lashkaripour A, Fordyce P, Densmore D (2022) Machine learning for microfluidic design and control. Lab Chip 22(16):2925–2937
119. Galan EA, Zhao H, Wang X, Dai Q, Huck WT, Ma S (2020) Intelligent microfluidics: the convergence of machine learning and microfluidics in materials science and biomedicine. Matter 3(6):1893–1922
120. Riordon J, Sovilj D, Sanner S, Sinton D, Young EW (2019) Deep learning with microfluidics for biotechnology. Trends Biotechnol 37(3):310–324
121. Pérez-Aliacar M, Doweidar MH, Doblaré M, Ayensa-Jiménez J (2021) Predicting cell behaviour parameters from glioblastoma on a chip images: A deep learning approach. Comput Biol Med: 104547
122. MATLAB deep learning toolbox. The MathWorks, Natick (version R2020b)
123. Metzcar J, Jutzeler CR, Macklin P, Köhn-Luque A, Brüningk SC (2024) A review of mechanistic learning in mathematical oncology. Front Immunol 15:1363144
124. Hatzikirou H (2022) Combining dynamic modeling with machine learning can be the key for the integration of mathematical and clinical oncology: comment on "improving cancer treatments via dynamical biophysical models" M. Kuznetsov, J. Clairambault, V. Volpert. Phys Life Rev 40: 1–2
125. Lagergren JH, Nardini JT, Baker RE, Simpson MJ, Flores KB (2020) Biologically-informed neural networks guide mechanistic modeling from sparse experimental data. PLoS Comput Biol 16(12):1008462
126. Cazzaniga P, Damiani C, Besozzi D, Colombo R, Nobile MS, Gaglio D, Pescini D, Molinari S, Mauri G, Alberghina L et al (2014) Computational strategies for a system-level understanding of metabolism. Metabolites 4(4):1034–1087
127. Strutz J, Martin J, Greene J, Broadbelt L, Tyo K (2019) Metabolic kinetic modeling provides insight into complex biological questions, but hurdles remain. Curr Opin Biotechnol 59:24–30
128. Katt ME, Placone AL, Wong AD, Xu ZS, Searson PC (2016) In vitro tumor models: advantages, disadvantages, variables, and selecting the right platform. Front Bioeng Biotechnol 4:12
129. Muñoz-Sierra R, Ayensa-Jiménez J, Doblaré M (2025) On the application of physically-guided neural networks with internal variables to continuum problems. Mech Mater: 105317
130. Raissi M, Perdikaris P, Karniadakis GE (2017) Physics informed deep learning (part ii): Data-driven discovery of nonlinear partial differential equations. arXiv preprint arXiv:1711.10566
131. Raissi M (2018) Deep hidden physics models: deep learning of nonlinear partial differential equations. J Mach Learn Res 19(1):932–955
132. Ayensa-Jiménez J, Doweidar MH, Sanz-Herrera JA, Doblare M (2022) Understanding glioblastoma invasion using physically-guided neural networks with internal variables. PLoS Comput Biol 18(4):1010019
133. Ayensa-Jiménez J (2022) In silico data-set used to train the PGNNIV (1.0). Zenodo. [Data set]. https://doi.org/10.5281/zenodo.6349224
134. Lavrik I, Golks A, Krammer PH (2005) Death receptor signaling. J Cell Sci 118(2):265–267
135. Laplante M, Sabatini DM (2009) mtor signaling at a glance. J Cell Sci 122(20):3589–3594
136. Kholodenko BN (2006) Cell-signalling dynamics in time and space. Nat Rev Mol Cell Biol 7(3):165–176
137. Baker RE, Pena J-M, Jayamohan J, Jérusalem A (2018) Mechanistic models versus machine learning, a fight worth fighting for the biological community? Biol Let 14(5):20170660
138. Benzinger T (1971) Thermodynamics, chemical reactions and molecular biology. Nature 229(5280):100–102
139. Haynie DT (2001) Biological thermodynamics. Cambridge University Press
140. Dill K, Bromberg S (2010) Molecular driving forces: statistical thermodynamics in biology, chemistry, physics, and nanoscience. Garland Science
141. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M et al (2013) The cancer imaging archive (TCIA): maintaining and operating a public information repository. J Digit Imaging 26:1045–1057
142. Dou J, Mao S, Li H, Lin J-M (2019) Combination stiffness gradient with chemical stimulation directs glioma cell migration on a microfluidic chip. Anal Chem 92(1):892–898
143. Solbu AA, Caballero D, Damigos S, Kundu SC, Reis RL, Halaas Ø, Chahal AS, Strand BL (2023) Assessing cell migration in hydrogels: an overview of relevant materials and methods. Mater Today Bio 18:100537
144. Wieland F-G, Hauber AL, Rosenblatt M, Tönsing C, Timmer J (2021) On structural and practical identifiability. Curr Opin Syst Biol 25:60–69
145. Jaenisch R, Bird A (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 33(3):245–254
146. Perez-Aliacar M, Ayensa-Jimenez J, Doblare M (2023) Modelling cell adaptation using internal variables accounting for cell plasticity in continuum mathematical biology. bioRxiv, 2023–05
147. Celora GL, Byrne HM, Zois CE, Kevrekidis PG (2021) Phenotypic variation modulates the growth dynamics and response to radiotherapy of solid tumours under normoxia and hypoxia. J Theor Biol 527:110792
148. Celora GL, Byrne HM, Kevrekidis P (2023) Spatio-temporal modelling of phenotypic heterogeneity in tumour tissues and its impact on radiotherapy treatment. J Theor Biol 556:111248
149. Alber M, Tepole AB, Cannon WR, De S, Dura-Bernal S, Garikipati K, Karniadakis G, Lytton WW, Perdikaris P, Petzold L et al

(2019) Integrating machine learning and multiscale modeling-perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. NPJ Digit Med 2(1):1–11
150. Brady R, Enderling H (2019) Mathematical models of cancer: when to predict novel therapies, and when not to. Bull Math Biol 81:3722–3731
151. Narayanan H, Luna M, Sokolov M, Butté A, Morbidelli M (2022) Hybrid models based on machine learning and an increasing degree of process knowledge: application to cell culture processes. Ind Eng Chem Res 61(25):8658–8672
152. Chen Q, Ye Q, Zhang W, Li H, Zheng X (2023) Tgm-nets: a deep learning framework for enhanced forecasting of tumor growth by integrating imaging and modeling. Eng Appl Artif Intell 126:106867
153. Mascheroni P, Savvopoulos S, Alfonso JCL, Meyer-Hermann M, Hatzikirou H (2021) Improving personalized tumor growth predictions using a Bayesian combination of mechanistic modeling and machine learning. Commun Med 1(1):19
154. Liang B, Tan J, Lozenski L, Hormuth II DA, Yankeelov TE, Villa U, Faghihi D (2023) Bayesian inference of tissue heterogeneity for individualized prediction of glioma growth. IEEE Trans Med Imaging
155. Wu C, Lorenzo G, Hormuth DA, Lima EA, Slavkova KP, DiCarlo JC, Virostko J, Phillips CM, Patt D, Chung C, et al. (2022) Integrating mechanism-based modeling with biomedical imaging to build practical digital twins for clinical oncology. Biophys Rev 3(2)
156. Hernandez-Boussard T, Macklin P, Greenspan EJ, Gryshuk AL, Stahlberg E, Syeda-Mahmood T, Shmulevich I (2021) Digital twins for predictive oncology will be a paradigm shift for precision cancer care. Nat Med 27(12):2065–2066
157. Sager S (2023) Digital twins in oncology. J Cancer Res Clin Oncol 149(9):5475–5477