



HAL
open science

Deciphering oxygen distribution and hypoxia profiles in the tumor microenvironment: a data-driven mechanistic modeling approach

Pawan Kumar, Matthieu Lacroix, Pierrick Dupré, Janan Arslan, Lise Fenou, Béatrice Orsetti, Laurent Le Cam, Daniel Racoceanu, Ovidiu Radulescu

► To cite this version:

Pawan Kumar, Matthieu Lacroix, Pierrick Dupré, Janan Arslan, Lise Fenou, et al.. Deciphering oxygen distribution and hypoxia profiles in the tumor microenvironment: a data-driven mechanistic modeling approach. *Physics in Medicine and Biology*, 2024, 69 (12), pp.125023. 10.1088/1361-6560/ad524a . hal-04612770

HAL Id: hal-04612770

<https://hal.sorbonne-universite.fr/hal-04612770v1>

Submitted on 14 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

PAPER • OPEN ACCESS

Deciphering oxygen distribution and hypoxia profiles in the tumor microenvironment: a data-driven mechanistic modeling approach

To cite this article: P Kumar *et al* 2024 *Phys. Med. Biol.* **69** 125023

View the [article online](#) for updates and enhancements.

You may also like

- [Machine Learning meets Kepler: Inverting Kepler's Equation for All vs All Conjunction Analysis](#)
Kevin Thomas Jordi Otto, Simon Burgis, Kristian Kersting et al.
- [Incorporation of zinc oxide and cobalt doped zinc oxide nano-rods in commercial sunscreen sample to optimize its UV-ray absorption](#)
Komal Ansari, Rehana Riaz, Fatima Gull et al.
- [Persistent spiking activity in neuromorphic circuits incorporating post-inhibitory rebound excitation](#)
Archita Hore, Sharba Bandyopadhyay and Saswat Chakrabarti



PAPER

OPEN ACCESS

RECEIVED
8 March 2024REVISED
10 May 2024ACCEPTED FOR PUBLICATION
30 May 2024PUBLISHED
14 June 2024

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Deciphering oxygen distribution and hypoxia profiles in the tumor microenvironment: a data-driven mechanistic modeling approach

P Kumar^{1,2,*} , M Lacroix^{3,4} , P Dupré^{3,4} , J Arslan^{2,5} , L Fenou³ , B Orsetti³ , L Le Cam^{3,4} ,
D Racoceanu² and O Radulescu^{1,*}

¹ Laboratory of Pathogens and Host Immunity, University of Montpellier, CNRS, INSERM, Montpellier, France

² Sorbonne Université, CNRS, INSERM, AP-HP, Inria, Paris Brain Institute (ICM), Paris, France

³ Institut de Recherche en Cancérologie de Montpellier (IRCM), INSERM U1194, University of Montpellier, Institut régional du Cancer de Montpellier (ICM), Montpellier, France

⁴ Equipe labélisée Ligue Contre le Cancer, Paris, France

⁵ Centre for Eye Research Australia, University of Melbourne, Royal Victorian Eye & Ear Hospital, East Melbourne, Australia

* Authors to whom any correspondence should be addressed.

E-mail: pawan.kumar1@umontpellier.fr and ovidiu.radulescu@umontpellier.fr

Keywords: oxygen gradient, hypoxia, mechanistic modeling, tumor heterogeneity, reaction-diffusion model

Supplementary material for this article is available [online](#)

Abstract

Objective. The distribution of hypoxia within tissues plays a critical role in tumor diagnosis and prognosis. Recognizing the significance of tumor oxygenation and hypoxia gradients, we introduce mathematical frameworks grounded in mechanistic modeling approaches for their quantitative assessment within a tumor microenvironment. By utilizing known blood vasculature, we aim to predict hypoxia levels across different tumor types. **Approach.** Our approach offers a computational method to measure and predict hypoxia using known blood vasculature. By formulating a reaction-diffusion model for oxygen distribution, we derive the corresponding hypoxia profile. **Main results.** The framework successfully replicates observed inter- and intra-tumor heterogeneity in experimentally obtained hypoxia profiles across various tumor types (breast, ovarian, pancreatic). Additionally, we propose a data-driven method to deduce partial differential equation models with spatially dependent parameters, which allows us to comprehend the variability of hypoxia profiles within tissues. The versatility of our framework lies in capturing diverse and dynamic behaviors of tumor oxygenation, as well as categorizing states of vascularization based on the dynamics of oxygen molecules, as identified by the model parameters. **Significance.** The proposed data-informed mechanistic method quantitatively assesses hypoxia in the tumor microenvironment by integrating diverse histopathological data and making predictions across different types of data. The framework provides valuable insights from both modeling and biological perspectives, advancing our comprehension of spatio-temporal dynamics of tumor oxygenation.

1. Introduction

Within several types of solid tumors, increased metabolic demands and abnormal vasculature lead to inadequate oxygen availability. This state is termed hypoxia and the corresponding regions of the tumor microenvironment are defined as hypoxic (McKeown 2014). Hypoxia, considered as one of the hallmarks of most solid tumors, compromises the functionality of cells and even their survival, resulting in the adaptation of (cancer) cells and their dynamics. These adaptations, which involve changes at both the cellular and the population levels, impinge on tumor progression by influencing drug resistance, invasiveness, angiogenesis, and vasculogenic mimicry (D'Aguanno *et al* 2021). Members of the hypoxia-inducible factor (HIF) family of transcription factors play a central role in the response to hypoxia by influencing multiple signaling pathways

and gene regulation that promote neovascularization, invasion, migration, adhesion, metastasis, metabolism, DNA replication, protein synthesis and phenotypic switching (from proliferative to invasive) (Harris 2002, Widmer *et al* 2013, Al Tameemi *et al* 2019). Hypoxia-driven changes contribute to the intra-tumor heterogeneity, thereby affecting accurate diagnosis. Another crucial effect of hypoxia is reduced treatment efficacy, as tumor cells in hypoxic regions do not respond well to many treatments including those based on chemotherapies or radiation therapy. Such effects contribute to the aggressiveness of several tumor types including melanoma, the most aggressive, complex, and heterogeneous skin cancer (Bedogni and Powell 2009, Grzywa *et al* 2017, D'Aguzzo *et al* 2021). The alleviation of hypoxia in a tissue region can be achieved through several strategies involving an increment in oxygen delivery or the decreased oxygen consumption by the cancer cells (Gallez *et al* 2017). The former is a widely accepted procedure (Suh *et al* 2006, Horsman and Van der Kogel 2009) and few studies, e.g. (Secomb *et al* 1995, Lin and Maity 2015, Gallez *et al* 2017), suggest that the latter can also be efficient in overcoming hypoxia and its consequences. Hence, a better assessment of oxygen distribution, hypoxia profiles, and O_2 consumption in the tumor microenvironment would improve prognosis and evaluate therapeutic options.

Mathematical modeling has proven to be beneficial in describing complex dynamics, offering powerful and cost-efficient tools for making predictions in various fields. In the context of tumor dynamics, mathematical models are valuable for reconciling spatiotemporal and phenotypic heterogeneity, as well as designing improved treatment strategies (Altrock *et al* 2015, Hodgkinson *et al* 2022). Numerous computational and mathematical models have been proposed to describe tissue oxygenation, especially focusing on the distribution of oxygen in both cancerous and regular tissues (Hsu and Secomb 1989, Secomb *et al* 1993, 2004, Goldman 2008, Toma-Dasu and Dasu 2013, Grimes *et al* 2016).

The cylinder model, initially proposed in Krogh (1919), laid the foundation for calculating oxygen flow based on an analytical expression dependent on the difference in oxygen tension in steady-state. Subsequently, several partial differential equation (PDE) models of the reaction-diffusion type were developed to represent the oxygen spatial distribution (Hsu and Secomb 1989, Secomb *et al* 1993, 2004, Dasu *et al* 2003, Powathil *et al* 2012, Skeldon *et al* 2012, Toma-Dasu and Dasu 2013, Welter *et al* 2016). These models involve different forms of source terms and boundary conditions. The resulting PDEs were numerically solved through discretization methods (finite-difference or finite element) or numerical integration, as seen in the alternative approach based on Green's function method (Hsu and Secomb 1989, Grimes *et al* 2016). This method was later modified in Secomb *et al* (1993, 2004) to account for more complex vascular architecture.

Despite these advancements, only a few works explicitly discuss hypoxia in connection with real-world tumor samples or experimental data. For instance, in Powathil *et al* (2012), the predicted hypoxia, observed using the needle electrode technique, is qualitatively compared with CD31⁶ stained images. However, most formulations are limited to synthetic data and qualitative results, lacking direct application and connection with actual tumor tissues and experimental outcomes. Hence, there is still a need for accurate predictive models of hypoxia that take into account tissue heterogeneity.

Here, we propose a physical model that predicts hypoxia in heterogeneous tumor microenvironments from stained tissue images. For this aim, we develop a data-driven mathematical framework to model the distribution of oxygen and hypoxia in the tumor microenvironment, incorporating mechanistic inferences. We produce experimental data in the form of high-pixel images illustrating staining of markers of blood vessels, hypoxia, and cell nuclei. The identification of these three elements is detailed as follows:

- **Blood Vessels:** The CD31 marker is a proxy for blood vessels. CD31, also known as PECAM-1 (platelet endothelial cell (EC) adhesion molecule 1), is a 130 kDa transmembrane glycoprotein belonging to the immunoglobulin superfamily. It is present on the surface of platelets, monocytes, macrophages, and neutrophils and is a constituent of the endothelial intercellular junction (Pusztaszeri *et al* 2006). ECs form a continuous lining (endothelium) along the inner surface of blood vessels, providing a barrier between circulating blood and surrounding tissues. By identifying these cells, the desired blood vessels are marked.
- **Hypoxia:** Intra-tumoral hypoxia is responsible for activating the key transcription factor HIF, which mediates the activation of several defense mechanisms for tumor cells. HIF-1 is a heterodimeric transcription factor comprising an HIF-1 α subunit and an HIF-1 β subunit (Mabjeesh and Amir 2014). HIF-1, promoted by intra-tumoral hypoxia, regulates Carbonic Anhydrase IX (CA9) expression. CA9 is considered one of

⁶ The abbreviation 'CD31' refers to 'Cluster of Differentiation 31', a biomarker commonly utilized for identifying blood vessels, particularly in tumor tissues, as detailed in the next paragraph.

the best-characterized HIF-1 targets (van den Beucken *et al* 2009), serving as an intrinsic marker for tumor hypoxia⁷.

- Cell Nuclei: The DAPI marker, which stands for 4',6-diamidino-2-phenylindole, is used to identify cell nuclei (Tarnowski *et al* 1991, kapuscinski 1995). DAPI is a standard marker for detecting nuclei.

The images acquired upon immuno-detection of CD31, CA9 and DAPI on tissue sections prepared from Patient-Derived Tumor Xenograft (PDX) mouse models illustrate intra- and inter-tumor heterogeneity in oxygen and hypoxia distribution. Leveraging this dataset, we develop a mechanistic framework to model oxygen and hypoxia distribution, predicting the hypoxia profile for a tumor tissue based on known vasculature.

Our study successfully recovers the biologically relevant heterogeneity of oxygenation and identifies crucial parameters involved in reducing hypoxia. As hypoxia promotes radio/chemo-resistance (Vaupel and Mayer 2007, Lin and Maity 2015), these results are also valuable for predicting the efficacy of therapies. We modeled oxygen distribution using a steady-state solution of a reaction-diffusion PDE. For a known distribution of O₂, we modeled hypoxia based on the dynamics of CA9 expression, used as a proxy for HIF-mediated transcription. While similar theoretical and computational models for oxygen distribution exist in the literature (Carlson *et al* 2011, Toma-Dasu and Dasu 2013, Kempf *et al* 2015), the novelty of our work lies in identifying and validating a predictive theoretical model using tumor samples and addressing the challenge of hypoxia heterogeneity. More precisely, we demonstrate the derivation of a space-dependent coefficient PDE model that predicts oxygen and hypoxia gradients. This is achieved by starting from a static model with space-invariant coefficients and identifying proxies for space-dependent coefficients directly from images of stained tissues.

The remainder of the article is organized as follows. The methods used in this work are detailed in section 2, covering information about the experimentally produced data, data processing, quantification of blood vessel diversity, mathematical modeling of oxygen and hypoxia, numerical simulation, parameter estimation, and validation. The obtained results are discussed in section 3. The study is concluded with a discussion in section 4 and a summary in section 5.

2. Methods

2.1. Experimental data

2.1.1. Primary derived xenograft (PDX)

Tissues from three cancer types are engrafted: breast (Triple Negative Breast Cancer #0056, #0520, #1090; Her2+ Breast Cancer #1092), ovarian (High-Grade Serous Ovarian Cancer #0721), and pancreatic (#1087, #1137, #1905) as listed in table 1.

Once the PDX model is established and frozen at passage 3, a single fragment of frozen tumor, approximately 8 mm³, is implanted into the inter-scapular fat pads of 3 to 4-week-old female Swiss-nude mice. Tumor growth is measured weekly with a caliper, up to the ethical limit of 1500 mm³ (du Manoir *et al* 2014, Colombo *et al* 2015).

2.1.2. Immunohistochemistry (IHC) and immunofluorescence (IF)

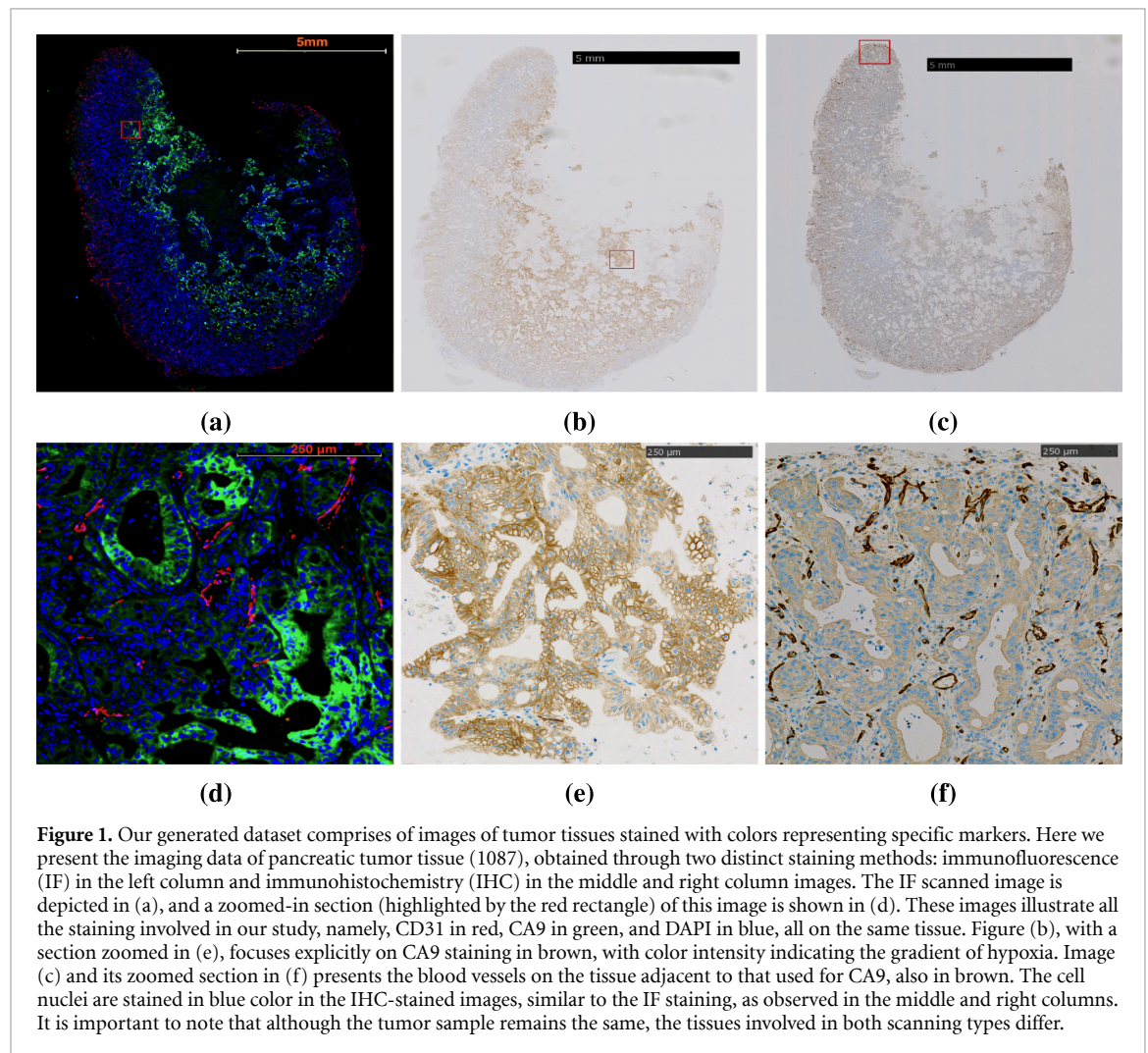
PDX samples are fixed in 4% neutral-buffered formalin for 24 h and then paraffin embedded. The FFPE tissues are sectioned into 4 μm sections and processed for IHC or IF. CD31 (Abcam, #ab28364) and CA9 (Roche, #760-6080) stainings are carried out on a VENTANA Discovery Ultra automated staining instrument (Ventana Medical Systems) following the manufacturer's instructions. For IF, after applying the primary antibodies, sections are incubated for 1 h at room temperature with secondary antibodies coupled to Alexa-488, Alexa-647 (Life Technologies) and nuclei are counterstained with 4',6-diamidino-2-phenylindole (DAPI, Sigma). Sections are then washed in PBST, and coverslipped using Mowiol mounting solution. For image analyses, sections are scanned on a 3DHISTECH scanner.

Analysis of CA9 and CD31 expression by IHC is conducted on adjacent tissue sections spaced four microns apart. The outcome reveals a brown color, indicating the presence of the targeted proteins, accompanied by cell nuclei depicted in blue, as exemplified in figure 1 (middle and right columns). Conversely, IF stainings are performed on the same tissue section with different target proteins represented by specific colors. Specifically, CD31, CA9, and DAPI are shown in red, green, and blue colors, respectively. The left column images in figure 1 visually demonstrate these stains in the mentioned colors.

⁷ In this study, we sometimes use the terms hypoxia and degree/measure of hypoxia interchangeably. Hypoxia is a state/condition of tissue environment represented by CA9 marker density in this work.

Table 1. Produced experimental data: In this work, we perform staining on three different PDX tumor types, namely breast, ovarian, and pancreatic. This table lists these tumor samples, with the first column containing the name of the sample used during the experiment. For each PDX, both staining techniques are applied to different tissue samples, generating one stained for IF and one for each CD31 and CA9 with IHC scans. Except for the tumor 0520, where IF is performed on multiple tissues of that tumor sample.

| Sample name | Tumor type | Tissue | # sections | Method |
|-------------|------------|--------|------------|---------|
| 1090 | Breast | PDX | 1 | IHC, IF |
| 1092 | Breast | PDX | 1 | IHC, IF |
| 0056 | Breast | PDX | 1 | IHC, IF |
| 0520 | Breast | PDX | 53 | IF |
| 0721 | Ovarian | PDX | 1 | IHC, IF |
| 1087 | Pancreatic | PDX | 1 | IHC, IF |
| 1137 | Pancreatic | PDX | 1 | IHC, IF |
| 1905 | Pancreatic | PDX | 1 | IHC, IF |



2.2. Heterogeneity in data: blood vessel diversity

Inter- and intra-tumoral heterogeneity is considered a challenging factor affecting therapeutic responses, and such heterogeneity is evident in our dataset, particularly concerning blood vessels, as depicted in figure 2. The images reveal diversity in blood vessel architecture and geometry, showing long and large capillaries in figure 2(a) and smaller ones in figure 2(g). Capillaries are densely packed in patches figures 2(g) and (i), while they are rare in the second patch of figures 2(d) and (h). This diversity may arise from the underlying heterogeneity within and between tumors and the angle at which the 3D tissue is sliced for 2D imaging. Since effective oxygen distribution relies on blood vessel distribution in the tissue, quantifying this heterogeneity is crucial, especially in our modeling process discussed in the subsequent sections.

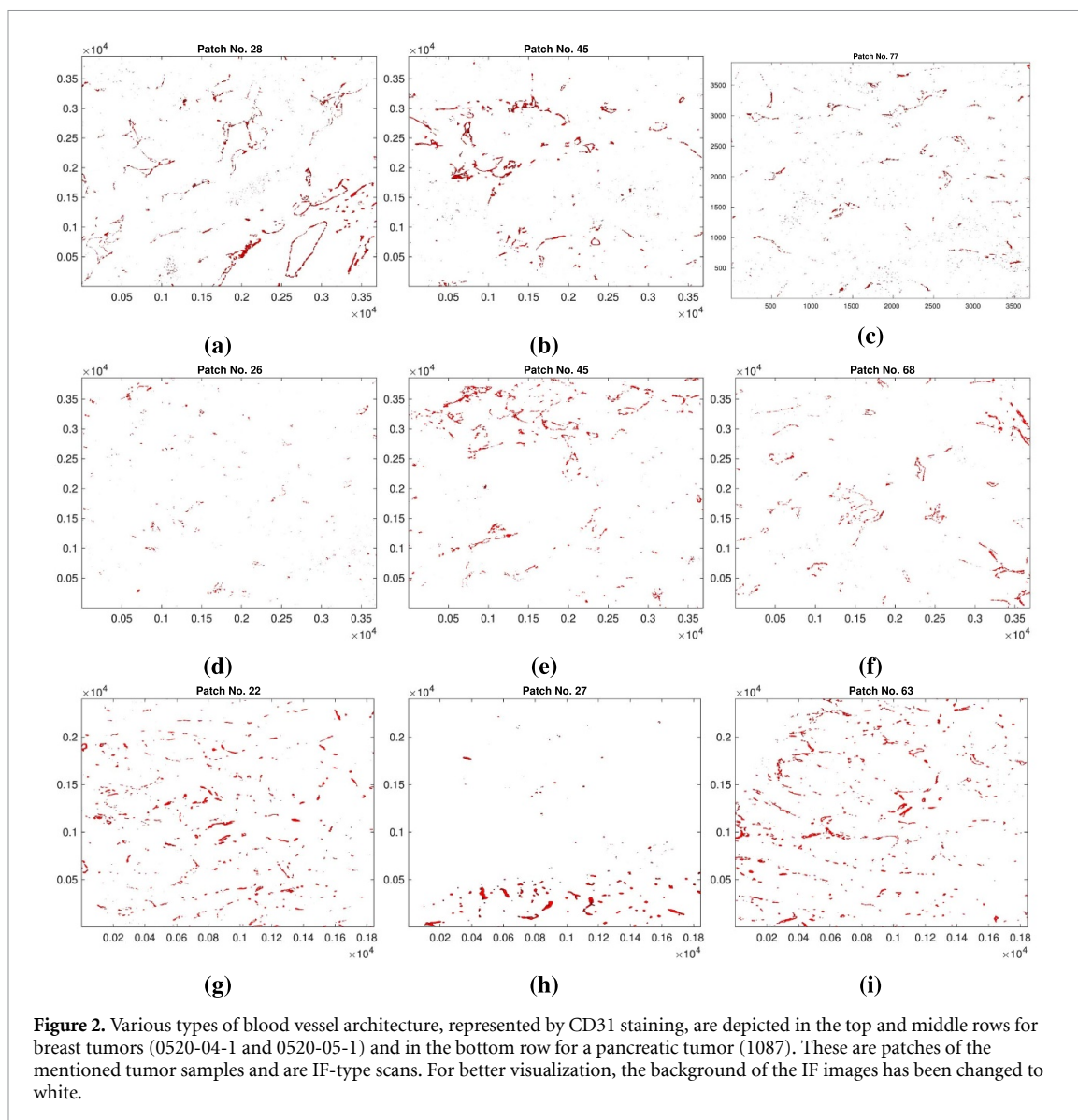


Figure 2. Various types of blood vessel architecture, represented by CD31 staining, are depicted in the top and middle rows for breast tumors (0520-04-1 and 0520-05-1) and in the bottom row for a pancreatic tumor (1087). These are patches of the mentioned tumor samples and are IF-type scans. For better visualization, the background of the IF images has been changed to white.

2.3. Data preprocessing

In order to utilize this data for spatial modeling, we apply some preprocessing techniques. Since the IHC stains of CA9 and CD31 are not conducted on the same samples, we perform image registration to obtain aligned images. Image registration is performed using two methods:

- Control point-based method: In this feature-based registration method, we manually identify a few matching features in both images (CD31 and CA9 stained IHC images) - specifically, matching tissue shapes. We choose 10 such points identical for each IHC scan pair. This selection of points is followed by estimating an affine transformation that differentiates one image (CD31 stained image, considered as the fixed one) from the other image (CA9 stained, considered as the moving one) geometrically. Applying the estimated transformation to the moving image aligns it with the fixed image.
- Optimization-based method: In this intensity-based registration method, we estimate the transformation using a regular step gradient descent optimization with mean square error metric configuration using the function `imregtform` in MATLAB⁸.

For the images utilized in this study, the control point-based method outperforms the intensity-based method. However, the latter proves useful when marking identical features in the images is challenging. Following the completion of registration, we employ color thresholding in MATLAB to isolate specific

⁸ MATLAB Version: 9.13.0.2126 072 (R2022b), The MathWorks Inc. Natick, Massachusetts, United States.

marker colors in the tissue images. This allows us to obtain distinct stains for each sample; for samples stained with the IF method, color selection, and corresponding thresholding are executed in Zen software⁹ to achieve the same outcome. Furthermore, given that these images are high-pixel images (i.e. a single image can be 20k × 20k in height and width), we extract small, representative patches from the whole image, particularly focusing on areas of interest for the subsequent steps outlined in the following sections.

2.4. Extraction of marker densities

After the necessary processing, we obtain images for the respective markers for each tissue sample. The upper row of figure 3 depicts the CD31, CA9, and DAPI marker staining images obtained from the preprocessing of the IF-based method for a patch of a breast tumor sample (1090). To facilitate modeling based on these images, spatial marker densities are required. This goal is achieved through three methods:

- Intensity-based approach: In the first method, a 2D Gaussian filter¹⁰ (G) is applied to the images, resulting in pixel intensity-based density.
- Binary-pixel approach: The second method extends the first by obtaining a binary image with non-zero pixels for staining. Then, a Gaussian filter is applied to get the density.
- Kernel density estimator approach: The third method is based on kernel density estimate¹¹ for the image parts covered by the non-zero pixels. In this case, a Gaussian kernel is defined for each point¹² on a considered grid in non-zero pixel regions, and the resulting density is the sum of all the kernels for the image.

The mentioned approaches are applied to the three markers based on their specific characteristics:

- For CA9: The intensity-based approach is applied to extract density for CA9, as the intensity and gradient are meaningful to mark the level of hypoxia. This method smoothens the images and ensures a continuous representation of hypoxia distribution.
- For CD31: The binary-pixel approach is used to obtain CD31 marker density. Initially, it is crucial to binarize the vascular and non-vascular regions for the blood vessels, followed by obtaining the density. Additionally, the resulting binarized image is further smoothed using a Gaussian filter.
- For DAPI: The kernel density estimator approach is applied to obtain cell density. This method captures the image portions covered by non-zero pixels and the proximity of the cell nuclei, forming a map of cells close to high density and low density for rarely packed cells.

The resultant densities from both these methods are demonstrated in the lower row of figure 3.

2.5. Heterogeneity quantification

To quantify the heterogeneity of blood vessels, we employ Claramunt entropy (Claramunt 2005) within patches. This entropy metric, a modified version of the Shannon entropy, takes into account the spatial distribution of blood vessels. Assuming that an increase in distance between entities of the same type and a decrease in distance between entities of different types will result in increased entropy, the spatial entropy H_s is defined as follows:

$$H_s = - \sum_{i=1}^n \frac{d_i^{\text{int}}}{d_i^{\text{ext}}} p_i \log_2(p_i) \quad (2.1)$$

d_i^{int} denotes the average distance between the entities (blood vessels) of the same type i (i.e. intra-distance), while d_i^{ext} represents the average distance between the entities of a given type i and the entities of the all other types (i.e. extra-distance), here p_i denotes the proportion of each entity i , with n being the number of such entities.

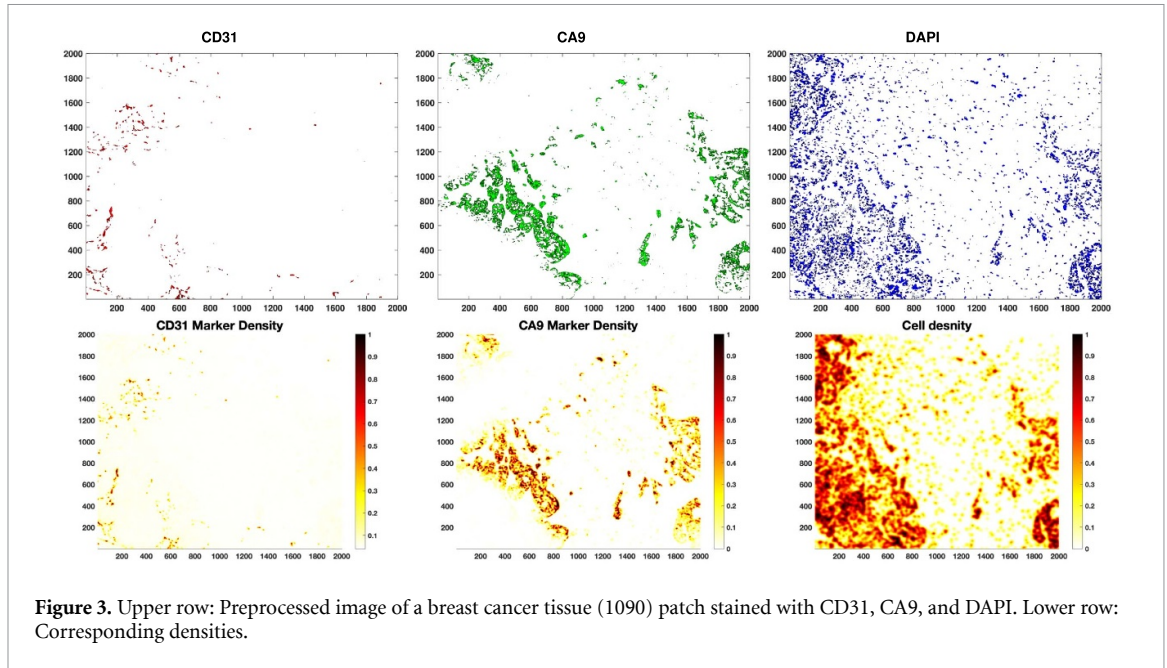
To apply the entropy method previously mentioned, we first categorize image patches based on the geometrical variations of blood vessels present within them. These variations are determined by the number of pixels in a region covered by each capillary in a given patch. To ensure consistency across all images, we

⁹ ZEISS ZEN (blue edition), Version 3.6.095.01 000, Carl Zeiss Microscopy GmbH, Germany.

¹⁰ The Gaussian filter G for image I in 2D is defined as $I * G$, where $G = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2}{2\sigma^2}}$, with σ being the standard deviation with zero mean applied to each point $\mathbf{x} = (x, y)$ on a 2D grid.

¹¹ The density $\hat{p}_n(\mathbf{x})$ is estimated by $\hat{p}_n(\mathbf{x}) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{\mathbf{x}_i - \mathbf{x}}{h}\right)$, with $K(\mathbf{x})$ being a Gaussian kernel function given by $K(\mathbf{x}) = \frac{1}{2\pi} e^{-\frac{(\mathbf{x}-\mathbf{x}_i)^2}{2}}$, h being the smoothing bandwidth, \mathbf{x}_i be the observed data point, \mathbf{x} is the value where kernel function is computed and n being the number of data points involved.

¹² In both methods, a point (in Euclidean space) is represented by the pixel coordinates of an image.



classify these variations into five categories (\mathcal{C}_i) based on pixel area¹³ (\mathcal{A}_i), $i \in [1, 5]$ of blood vessels, ranging from very small to massive vessels

$$\begin{aligned}
 \mathcal{C}_1 &:= 0 < \mathcal{A}_1 \leq 50 \\
 \mathcal{C}_2 &:= 50 < \mathcal{A}_2 \leq 100 \\
 \mathcal{C}_3 &:= 100 < \mathcal{A}_3 \leq 250 \\
 \mathcal{C}_4 &:= 250 < \mathcal{A}_4 \leq 500 \\
 \mathcal{C}_5 &:= \mathcal{A}_5 > 500.
 \end{aligned} \tag{2.2}$$

In the context of the formulation presented in equation (2.1), the term p_i denotes the percentage of all non-zero pixels associated with the category \mathcal{C}_i . For a particular patch, as depicted in figure 4(a), these categories are graphically represented using five distinct colors, as shown in figure 4(b). Each color signifies a category \mathcal{C}_i , with the boundaries of individual blood capillaries drawn by these colors. Consequently, intra-distances d_i^{int} refer to the distances between vessels of the same color, whereas inter-distances d_i^{ext} are the distances between vessels of different colors. The measurement of distances is achieved using the centroid of each vessel, identified as the connected pixels within the image. This approach comprehensively characterizes the spatial relationships between blood vessels, forming the basis for the subsequent entropy calculation outlined in equation (2.1).

Applying a moving window approach, we compute the entropy for patches within each sample using equation (2.1). As illustrated in figure 5(a), we initially divide an image sample into 100 patches. Subsequently, we calculate the diversity index (H_s) for each patch, as one of the patches exemplified in the zoomed-in form in figure 5(b). The resulting diversity map for this tissue sample is depicted in figure 5(c). The colormap in figure 5(c) visually represents the diversity of blood vessel architecture within a tumor sample. This quantification of entropy provides valuable insights into the spatial distribution and variation of blood vessels, serving as a basis for identifying similar patches during the validation process.

2.6. Modeling

Utilizing the data mentioned above, which includes the densities of blood vessels, hypoxia, and cell nuclei, we develop a mathematical framework and physical models to investigate the relationship between oxygen and hypoxia distribution within tumor tissues. We begin by forming a model of oxygen distribution.

¹³ Throughout this classification, the term ‘area’ is used to denote the number of pixels in a connected component, which remains independent of the actual pixel dimensions in each image. When applying equation (2.1), we require the proportion of such connected components, making it independent of the actual pixel length and breadth. Similarly, distances are calculated based on pixel coordinates, so the ratio of distances involved in equation (2.1) is also independent of actual pixel dimensions. This approach ensures that the comparison of H_s between different image patches is meaningful.

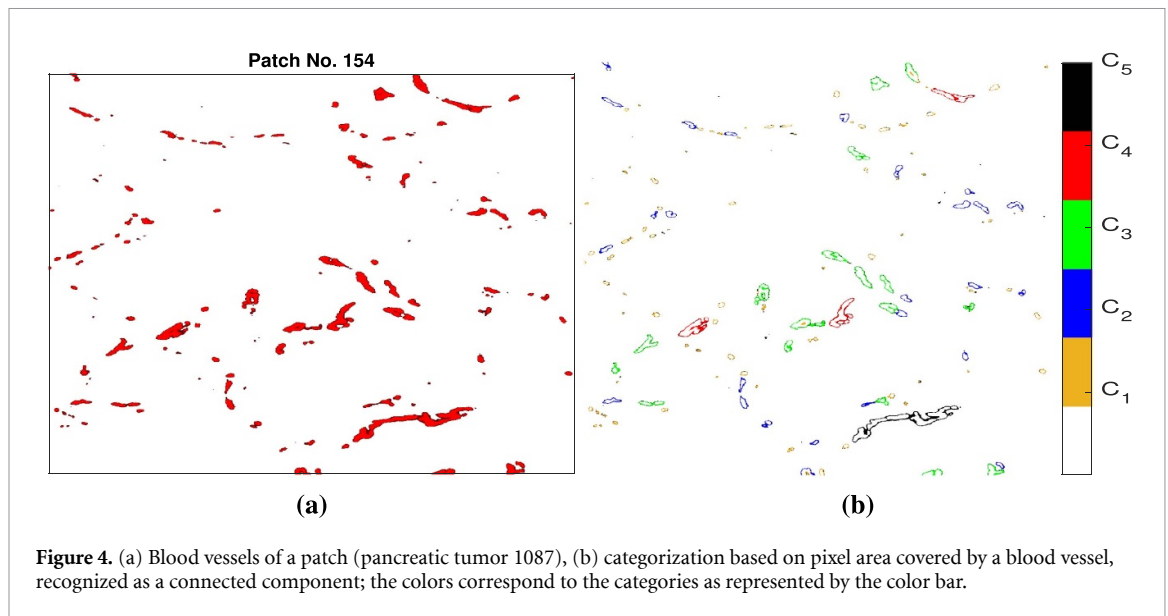


Figure 4. (a) Blood vessels of a patch (pancreatic tumor 1087), (b) categorization based on pixel area covered by a blood vessel, recognized as a connected component; the colors correspond to the categories as represented by the color bar.

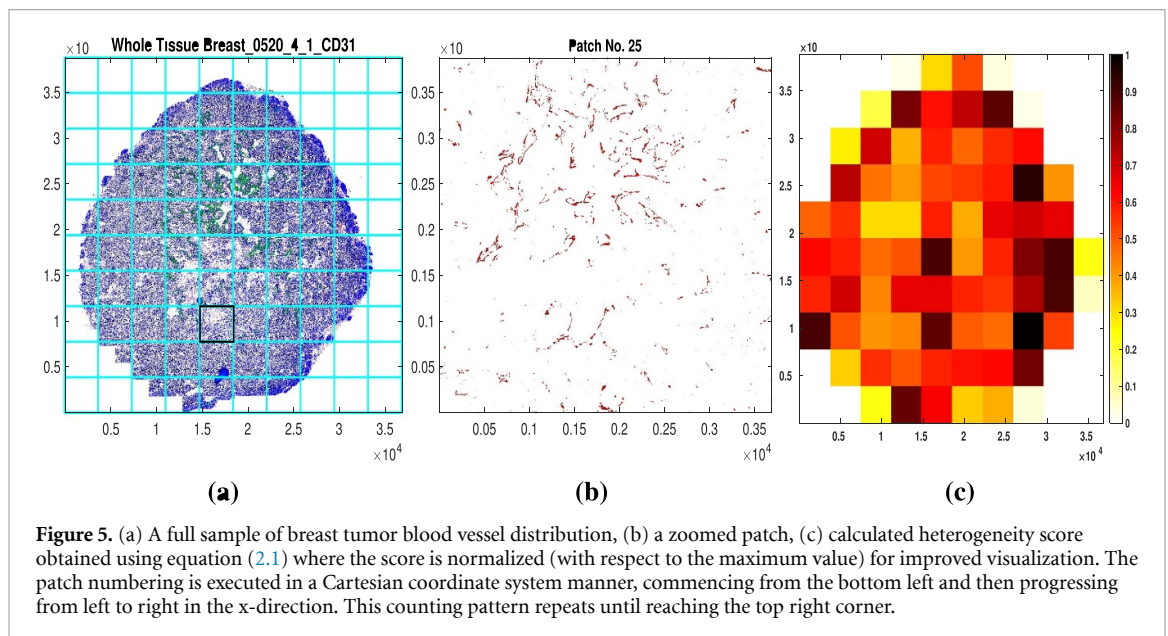


Figure 5. (a) A full sample of breast tumor blood vessel distribution, (b) a zoomed patch, (c) calculated heterogeneity score obtained using equation (2.1) where the score is normalized (with respect to the maximum value) for improved visualization. The patch numbering is executed in a Cartesian coordinate system manner, commencing from the bottom left and then progressing from left to right in the x-direction. This counting pattern repeats until reaching the top right corner.

2.6.1. Model of oxygen distribution

The dynamics of oxygen distribution within the microenvironment are influenced by various factors, including blood vessel architecture, geometry, functional state, permeability, the metabolic activity of cells, and oxygen consumption. In this study, we specifically focus on the following aspects:

- **Production of O_2 :** Our model assumes a direct correlation between oxygen production and blood vessel density in the tissue. Increased vessel density, particularly of capillaries, enhances oxygen exchange between blood and tissues. As detailed in section 2.4, estimating vessel density involves considering both overall density and spatial distribution. By incorporating vessel positions, we account for each vessel's neighborhood, crucial for determining local oxygen supply.
- **Neighborhood Determination:** The appropriate neighborhood for O_2 distribution is determined by the ability of O_2 molecules to migrate the longest distance from their source vessels. To model this, we consider the random motion of oxygen molecules in the tissue, hence employing a linear diffusion to quantify this process. This method is a PDE counterpart of the experimentally observed diffusion length, which is applied in some previous studies as in Turner *et al* (2002) and Vaupel *et al* (2021).
- **Consumption in the Tissue:** Oxygen undergoes consumption in the tissue, primarily by cells and other chemical reactions in the tumor microenvironment. The uptake of oxygen is a crucial factor affected by the cellular metabolic state and the state of the tumor microenvironment. From a treatment perspective, a

decrease in consumption can be utilized to reduce hypoxia (Gallez *et al* 2017). In our modeling approach, we consider uptake directly proportional to the available oxygen concentration, encapsulating all the influencing factors into a unified representation.

To facilitate our modeling process, we introduce some notations for the variables involved. We consider the following:

- $u(t, \mathbf{x})$: the oxygen concentration depending upon time t and space \mathbf{x} ,
- D : diffusion coefficient of O_2 (a constant),
- $V(\mathbf{x})$: blood vessel density,
- α : oxygen uptake coefficient,
- β : oxygen production coefficient.

Given the considerations outlined for modeling, along with the declared variables, we formulate a reaction-diffusion model for oxygen distribution in the following form:

$$\frac{\partial u}{\partial t} = D\Delta u - \alpha u + \beta V. \quad (2.3)$$

In this formulation, the modulation of oxygen concentration over time is governed by constant diffusion (represented by the first summand on the right-hand side of equation (2.3)). Additionally, there is a constant uptake proportional to the existing O_2 distribution (second summand) and oxygen production linked to the blood vessel density (third summand). This formulation aligns with previous models developed in Kelly and Brady (2006), Skeldon *et al* (2012) and Welter *et al* (2016). Notably, in our data-driven approach, we have established an explicit relationship between blood vessel density obtained from data and oxygen production.

Further, as oxygen dynamics is rapid enough and it reaches the corresponding equilibrium state very quickly, as in Kelly and Brady (2006) and Welter *et al* (2016), we consider the steady state from the reaction-diffusion model:

$$D\Delta u - \alpha u + \beta V = 0, \quad (2.4)$$

with boundary conditions: $\nabla u \cdot n = 0$, which ensures no oxygen flux exchange between the considered domain and outside, with n representing the outward normal to the surface boundary (Welter *et al* 2016).

The model is equation (2.3) ensures the supply of oxygen by the available blood vessels; however, the production should not be unbounded. To incorporate this phenomenon, we define a limited production in another form of the model:

$$D\Delta u - \alpha u + \beta \frac{V}{K_V + V} = 0,$$

where K_V is the carrying capacity of the blood vessel density. Rewriting $\tilde{V} = \frac{V}{K_V}$ and dropping the tilde for simplicity, we get the steady-state form as:

$$D\Delta u - \alpha u + \beta \frac{V}{1 + V} = 0, \quad (2.5)$$

with the same no-flux boundary condition.

2.6.2. Modeling of hypoxia-oxygen relationship

As mentioned in the introduction, hypoxic regions result from poor oxygenation. Therefore, the relationship between hypoxia and O_2 would be formulated in an inversely proportional manner.

One initial choice for modeling this relation is to consider a linear decay of hypoxia with an increase in O_2 . Let $h(u)$ represent hypoxia as a function of normalized oxygen distribution¹⁴, and the linear decay can be expressed as:

$$h(u) = \gamma(1 - u), \quad (2.6)$$

here, γ is a dimensionless constant conversion coefficient.

¹⁴ The oxygen distribution is normalized by its maximum value in the considered domain.

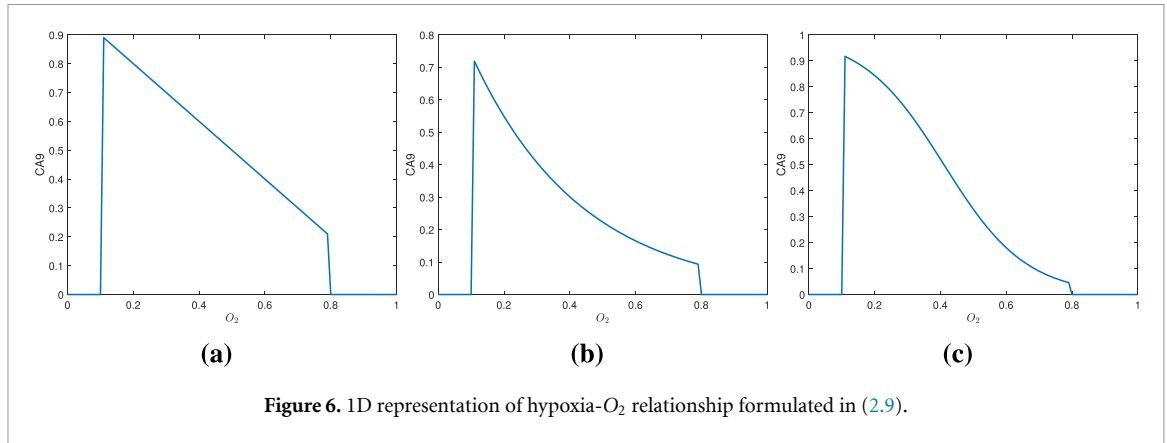


Figure 6. 1D representation of hypoxia- O_2 relationship formulated in (2.9).

Another formulation could involve an exponential decay:

$$h(u) = \gamma e^{-k_1 u}, \tag{2.7}$$

with k_1 being a dimensionless constant representing the rate of decay. Such decay has also been observed in Qutub and Popel (2006).

A more general approach is to consider a sigmoidal form:

$$h(u) = \frac{\gamma k_1}{k_1 + e^{(k_2(u-k_3))}}. \tag{2.8}$$

In this case, k_i , where $i \in [1, 3]$, are constants that govern a more controlled decrement of oxygen leading to hypoxia.

Utilizing the three models mentioned above, we obtain the corresponding hypoxia profiles, each representing the decay of hypoxia (linear, exponential, or sigmoidal). Furthermore, a switch-like¹⁵ relationship between O_2 and hypoxia is proposed in previous works, such as in Ginouves et al (2008), Cavadas et al (2013) and Nguyen et al (2015). We incorporate this response with the lower and upper thresholds of O_2 corresponding to non-zero hypoxia, denoted as O_l and O_h ; this relationship is modeled as follows:

$$h(u) = \begin{cases} 0, & \text{if } u \leq O_l, \\ \gamma(1-u) \text{ or } \gamma e^{-k_1 u} \text{ or } \frac{\gamma k_1}{k_1 + e^{(k_2(u-k_3))}}, & \text{if } O_l < u < O_h, \\ 0, & \text{if } u \geq O_h. \end{cases} \tag{2.9}$$

For the particular values of O_2 thresholds ($O_l = 0.2$ and $O_h = 0.8$, figure 6 illustrates the hypoxia- O_2 relationship represented by (2.9).

The upper bound (O_h) is evident as hypoxia decreases to zero in the presence of sufficient oxygen (normalized value = 1). This decrease can manifest as a linear decay (figure 6(a)), a sharp exponential decay (figure 6(b)), or a sigmoidal-type decay (figure 6(c)). The lower bound (O_l) models two processes. First, in cases of severe and prolonged hypoxia, HIF-1 levels decline, as suggested in Ginouves et al (2008), Saxena and Jolly (2019), Korbecki et al (2021). As HIF regulates CA9 (Hirota and Semenza 2006, Mabweesh and Amir 2014), which is a well-characterized HIF-1 target (van den Beucken et al 2009). Hence, below the lower threshold (O_l), the hypoxia $h(u)$ switches to zero. Secondly, this lower bound is also useful to incorporate prediction for non-cellular regions, where CA9 and CD31 markers are absent simultaneously. However, in our simulations, these parameters are obtained from the underlying data using an estimation method.

2.6.3. Final model system

Hence, our developed model system involves a PDE (equation (2.4)) for oxygen distribution and an algebraic relation between O_2 and hypoxia profile in (equation (2.9)) based on the mentioned dataset, which will be utilized to estimate the involved parameters. This estimation is performed with the exponential decay form of equation (2.9). We have also tested the model with bounded production described by (2.5) but did not observe an increase in training and validation performances.

¹⁵ A sharp transition between states, in this case, from normoxic to (severe) hypoxic state or vice-versa.

Table 2. Training dataset.

| Sample name | Tumor type | Tissue | # sections | Method | No. of patches |
|-------------|------------|--------|------------|--------|----------------|
| 1092 | Breast | PDX | 1 | IF | 81 |
| 0056 | Breast | PDX | 1 | IF | 76 |
| 0520 | Breast | PDX | 53 | IF | 4045 |
| 1137 | Pancreatic | PDX | 1 | IF | 77 |
| 1905 | Pancreatic | PDX | 1 | IF | 72 |

Table 3. Validation dataset.

| Sample name | Tumor type | Tissue | # sections | Method |
|-------------|------------|--------|------------|---------|
| 1090 | Breast | PDX | 1 | IF |
| 1092 | Breast | PDX | 1 | IHC |
| 1087 | Pancreatic | PDX | 1 | IF, IHC |
| 1905 | Pancreatic | PDX | 1 | IHC |
| 0721 | Ovarian | PDX | 1 | IHC |

2.7. Numerical simulation methods

To estimate the parameters in our formulated system, we numerically solve the PDE presented in equation (2.3). We use a finite-difference scheme where the spatial domain is discretized into 400×400 grid points. A five-point stencil (central-difference) discretization is applied to the diffusion term on these grid points. The resultant system of linear equations consists of 160 000 unknowns, representing the values of unknown ($u(\mathbf{x})$) at the grid points. The simulations are conducted within the MATLAB environment.

Upon solving the PDE in equation (2.3) for a given set of parameters, the O_2 distribution u is obtained at the discretized points. Subsequently, hypoxia $h(u)$ is calculated using equation (2.9) at the considered grid points within the domain. This forms the basis for further analysis and parameter estimation.

2.8. Estimation of parameters

Model parameters are determined using an optimization method. To do this, we first define a loss function as:

$$\text{loss function} = \|h(u) - H\|, \quad (2.10)$$

where $h(u)$ is the computed hypoxia from (2.9), H represents the hypoxia profile obtained from the CA9 scan and $\|\cdot\|$ denotes the square of L^2 -norm per grid point. The loss function is computed within the considered domain where either of the scans (CD31 or CA9) is non-zero. The minimization of this loss function provides estimates for the parameters.

The optimization is performed in MATLAB using the built-in `MultiSearch` function, where 100 to 500 random values are chosen for the initial set of parameters in the optimization process. This function uses multiple random starting points and generates optimal and also sub-optimal solutions. The search of optima is parallelized and calls upon the sequential quadratic programming (sqp) algorithm via `fmincon` in MATLAB. Additionally, we have also explored the non-linear least square fit method and `GlobalSearch` algorithm, which is based on the scatter-search method of generating trial points and automatically selecting 1000 starting points. Our findings indicate that the `MultiSearch` with `fmincon` equipped with the sqp algorithm outperforms other methods, at least as far as our problem is concerned.

2.9. Training and validation datasets

We divide the dataset into two sections: one for training the parameters and the other for validating the obtained model predictions. In this partition, we ensure diversity in both datasets by including all types of tumor samples and scan types. This selection ensures that the same tumor tissue is not used for both training and validation, and a variety of tumor types are included in each set to validate predictions effectively. Our framework is designed to accommodate different scan types for staining. Therefore, we chose IF for training, while validation is performed on both IF and IHC types, as illustrated in table 2. Training on IF samples is convenient due to the same tissue used for staining. Additionally, for validation (as shown in table 3), we include both scan types and tumor types to comprehensively test the predictive capabilities of our modeling framework. Next, we further divide each sample into small patches and train the model on each of them, as shown in table 2, indicating the number of patches.

2.10. Training

As depicted in figure 5(a), we partition each image into 100 patches and independently train the parameters for each patch. The final model system discussed in section 2.6.3 is solved using a no-flux boundary condition for each patch. This condition accounts for the inflow and outflow of oxygen between neighboring patches, considering the presence of blood vessels on either side. Therefore, no flux is exchanged at the boundaries. To assess the training performance against a worst-case scenario, we compute the minimized loss function and compare it with the mean of 50 random values sampled from the hypoxia profile at each grid point, which is generated during the discretization of the PDE. This evaluation involves generating 50 random values (ranging between 0 and 1) at each point and calculating their mean. The obtained mean value represents the potential hypoxia value at that point in a worst-case scenario. Further, in the training process, we record the parameter values corresponding to the 50% higher than the global optimum for each patch. Executing this process for the entire training set enables us to obtain optimal and sub-optimal parameters and an interval for each parameter for each patch, considering sub-optimal solutions 50% higher than the global optimum. This information provides insights into the uncertainty of the estimated parameters across different patches in the image.

2.11. Stratification of tissue regions

Upon completing the training phase, the trained framework is applied to classify a patch and identify similar patches from the trained dataset based on the blood vessel architecture. For a given patch from a tissue sample outside the training data, the classification process involves the following steps:

- Assigning Heterogeneity Score (H_s): Initially, calculate the heterogeneity score for the patch's blood vessel density using the method outlined in section 2.5 to quantify hypoxic regions.
- Proportion of Non-zero Pixel (N_p) Calculation: Next, we calculate the proportion of non-zero pixels representing blood vessels in the patch. As discussed in section 2.4, a binary pixel method is used to detect blood vessels in the images, providing a measure of the vessel-filled region in the patch.
- Initial Interval Determination: Determine two intervals with a variance of $\pm 10\%$ above and below both calculated values: $(\max(0, N_p - 0.1N_p), N_p + 0.1N_p)$ and $(\max(0, H_s - 0.1H_s), H_s + 0.1H_s)$, ensuring that the minimum values remain non-negative.
- Identification of similar patches: Once heterogeneity and non-zero pixels intervals are obtained for the target patch, identify patches with trained parameters, whose two scores lie within the intervals. These identified patches are considered 'similar' patches, and their parameters can be utilized to predict the hypoxia profile. Initiate an iterative search starting with the initial interval. The algorithm searches for trained patches within this interval. The search iteratively increases the interval until there is at least one set of values satisfying both conditions specified by these two intervals, providing a similar patch.

Hence, this classification ensures that similar patches contain equivalent vascular regions and that their spatial spread is comparable.

2.12. Validation of the physical model

This step involves predicting the hypoxia profile for a patch and subsequently for an entire image, then comparing the predictions with the actual data. The process is as follows:

2.12.1. Validation for a patch

To predict the hypoxia profile of a patch, the following steps are undertaken, starting from image preprocessing to solving a PDE, ultimately providing insights into the hypoxia profile of the tumor tissue sample.

- Input Image: Tumor tissue image stained with CD31 and DAPI, representing blood vessels and cell nuclei, respectively.
- Necessary Preprocessing: Implementation of necessary preprocessing steps as mentioned in section 2.1 to obtain the respective markers' densities.
- Patch Extraction: We divide the input sample into small patches. Depending upon the image pixel size, we divide them into 25 to 100 small patches.
- Classification of a Patch: Next, we apply the stratification method to identify similar patches from the training dataset, thereby obtaining an interval for each parameter.
- Parameter Assignment for Each Patch: We obtain a range of parameters for the patch, and each value can be identified as a different snapshot of dynamic oxygenation. For a general case, parameters are assigned for the patch by taking the mean of the interval of the similarly trained patches.

- PDE Solution: With the obtained parameters, the PDE in equation (2.4) is numerically simulated for the patch to get the O_2 distribution.
- Cell Region Identification: Identification of the cell regions based on DAPI marker density is performed to mark the regions where cells can be present and to make predictions for hypoxia.
- Hypoxia Calculation: Application of the hypoxia equation (2.9) and then setting the obtained to zero for non-cellular regions provide the desired profile for hypoxia for the input patch.
- Result Image: The final result is the hypoxia profile for the input patch.

2.12.2. Validation on whole sample

We assume constant parameters in the formulated model system described by equations (2.3) and (2.9), which may not hold for all tumor tissue samples. We adopt a data-driven approach to retrieve space-dependent model parameters from the data to predict the hypoxia profile for an entire sample. This is accomplished through the following steps using a moving window approach:

- Partition the input sample into several patches (e.g. 100 patches per sample), as illustrated in figure 5(a).
- Apply a classification strategy based on the heterogeneity score of each patch, as detailed in section 2.2.
- For each patch, acquire parameters from similar patches in trained samples, following the process detailed for the single patch case.
- The obtained parameters for each patch result in space-dependent parameters that remain invariant within a patch, expressed by the equation:

$$\nabla \cdot (D(\mathbf{x}) \nabla u) - \alpha(\mathbf{x}) u + \beta(\mathbf{x}) V = 0. \quad (2.11)$$

- Together with the no-flux boundary condition, equation (2.11) serves as a data-driven, space-dependent counterpart of equation (2.4).

This adaptation enables the consideration of spatial heterogeneity in parameters within the tissue sample, providing a more realistic representation of oxygen and hypoxia profiles. To address the spatial dependence of the diffusion coefficient D and prevent numerical oscillations, especially with discontinuous diffusion coefficients in the simulations of equation (2.11), we employ a conservative form of the five-stencil discretization method, as described in section 2.7.

3. Results

3.1. Training the mechanistic model for oxygen-hypoxia relationship

A simple yet effective mathematical framework is formulated to predict hypoxia from a known vessel architecture. The framework is mechanistic and data-oriented, utilizing underlying data for model development, training, and subsequent validation. The approach combines a theoretical foundation with empirical insights from the available dataset, emphasizing a balance between mechanistic understanding and practical applicability.

Following the training protocol detailed in section 2.10, it is apparent that the proposed framework aligns effectively with the data, as evidenced by the representations in figures 7 and S2. These scan images result from two distinct staining techniques, as elucidated in section 2.1, and the modeling framework exhibits a robust fit with both types of scans. Beyond the visual appeal of the fit comparison, a quantitative assessment reveals a strong agreement as we compare the minimized loss function values with the worst-case scenario, as explained in section 2.10. This is observable in the figures presented in figures 7 and S2. This consistent trend persists across the entire training set, where each patch's minimized loss function value is significantly smaller than the worst values, as shown in the figure S1.

3.2. Biological heterogeneity and epistemic uncertainty are responsible for distributed physical parameters

Upon estimating parameters for all patches within the training set, a distribution for each parameter is obtained, as depicted in figure 8(a). This visualization shows the parameter distribution, revealing notable inter-patch heterogeneity. The observed hypoxia heterogeneity in experiments (Iakovlev *et al* 2007) is consistent with the model predictions, particularly in terms of inter-patch variation of parameters.

The first source of parametric variability is biological and integrates several factors. Tissue oxygenation is a complex process influenced by various factors beyond our modeling constraints, with one such factor being the dynamic nature of oxygenation, particularly in vascularization induced by normal or tumor-related processes. This intricate process involves a series of coordinated steps, commencing with tumor cells

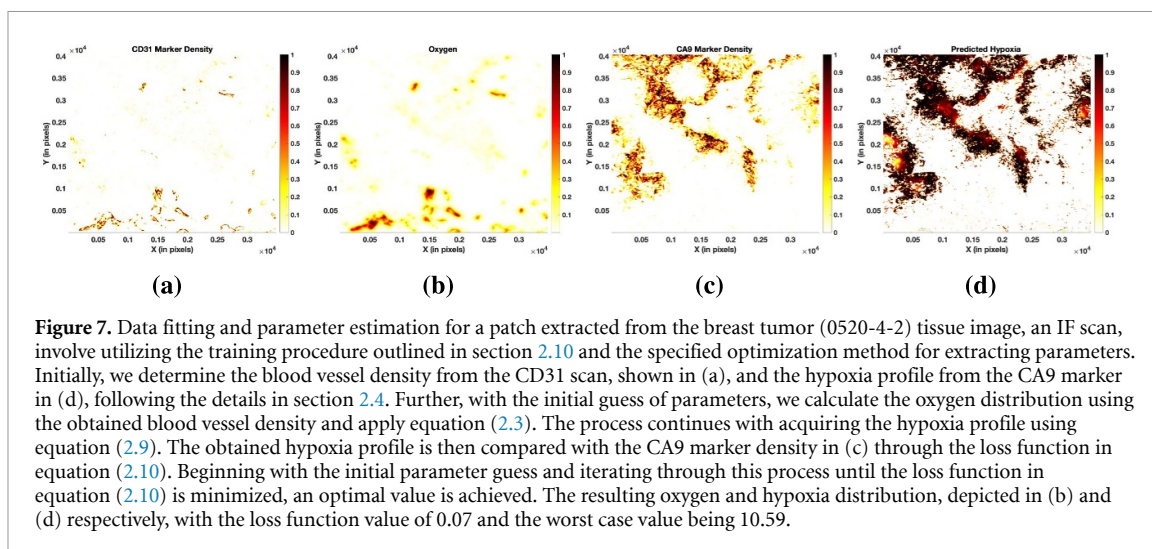


Figure 7. Data fitting and parameter estimation for a patch extracted from the breast tumor (0520-4-2) tissue image, an IF scan, involve utilizing the training procedure outlined in section 2.10 and the specified optimization method for extracting parameters. Initially, we determine the blood vessel density from the CD31 scan, shown in (a), and the hypoxia profile from the CA9 marker in (d), following the details in section 2.4. Further, with the initial guess of parameters, we calculate the oxygen distribution using the obtained blood vessel density and apply equation (2.3). The process continues with acquiring the hypoxia profile using equation (2.9). The obtained hypoxia profile is then compared with the CA9 marker density in (c) through the loss function in equation (2.10). Beginning with the initial parameter guess and iterating through this process until the loss function in equation (2.10) is minimized, an optimal value is achieved. The resulting oxygen and hypoxia distribution, depicted in (b) and (d) respectively, with the loss function value of 0.07 and the worst case value being 10.59.

overexpressing hypoxia-inducible regulators of angiogenesis, including vascular endothelial growth factor (VEGF). As a chemoattractant, VEGF promotes the proliferation of ECs lining the surrounding blood vessels. For a more detailed model and qualitative insights, particularly in the context of glioma, we refer to Kumar and Surulescu (2020).

The complexity of this process and the diversity in blood vessels and hypoxia profile are mirrored in the parameters. The observed inter-patch or inter-tumor variability is intricately linked to the underlying dynamics, facilitating an exploration of different possibilities and leveraging the mechanistic behavior of the model. These obtained distributions provide insights into the diverse, dynamic states of tissue oxygenation, as elaborated in subsequent sections.

Another source of parametric variability is epistemic, stemming from the fact that several parameter sets can result in similar fits. Therefore, trained parameters are inherently uncertain. In the optimization process, parameters corresponding to the minimum loss function, as defined in equation (2.10), are selected. However, introducing a bit more tolerance in the loss function results in an interval of parameters, reflecting epistemic parametric uncertainty. This type of variability results from insufficient training data, which can be mitigated by adding more data.

3.3. Prediction of hypoxia from stained tissue images using the physical model

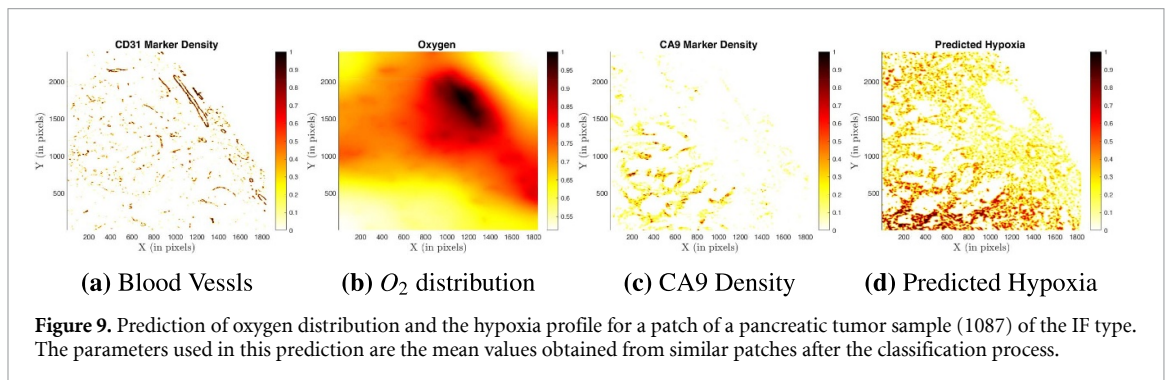
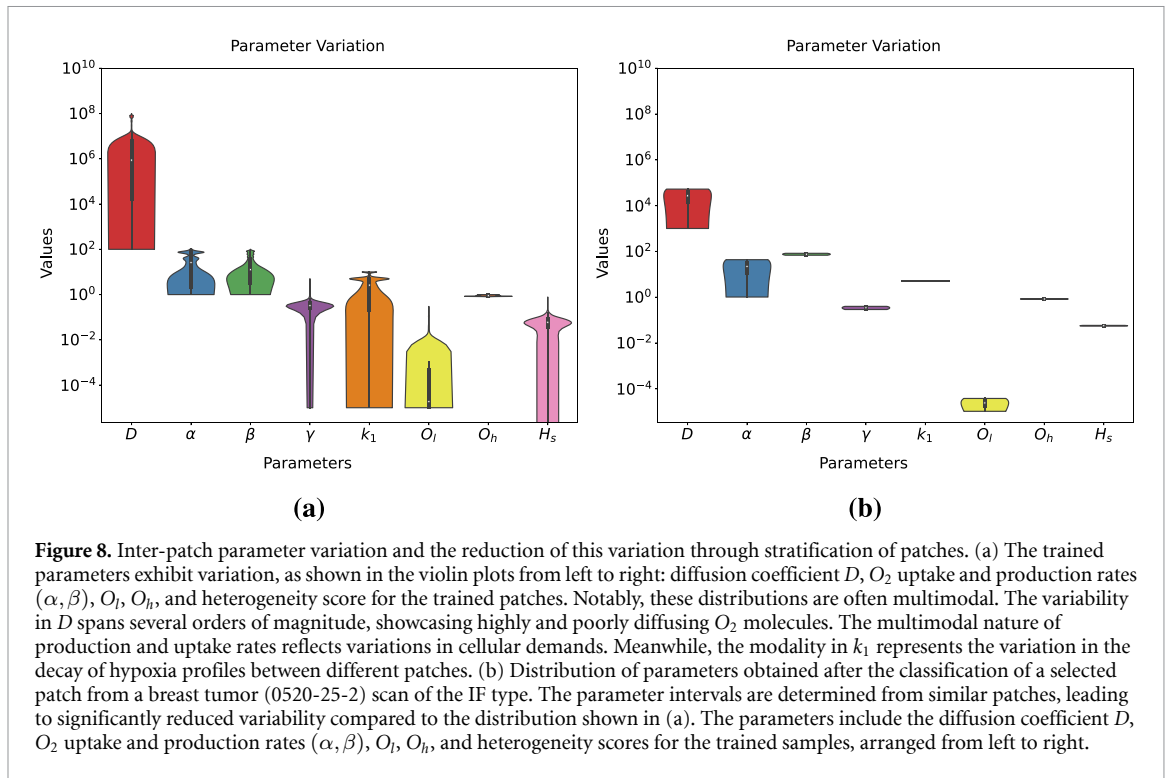
3.3.1. The stratification, based on blood vessel heterogeneity, reduces the variability of parameters

As we explore the variability of parameters across patches and consider the underlying dynamics, it becomes essential to narrow down options when making predictions using the framework. Indeed, model predictions depend on parameters that are *a priori* unknown. Our strategy is to use the heterogeneity score (spatial entropy H_s), which can be directly extracted from stained tissue images, to reduce the range of parameters required for prediction.

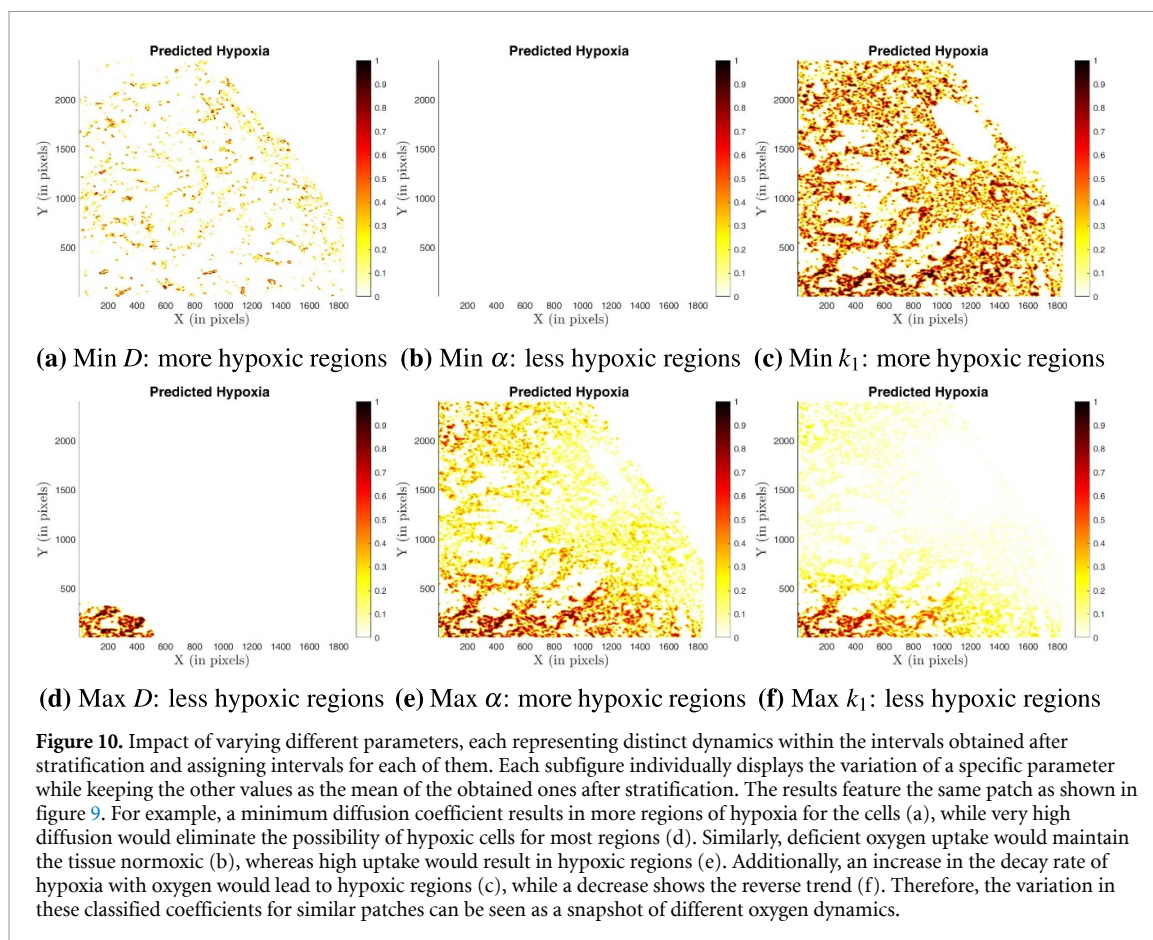
To assess the feasibility of this strategy, we employ the stratification method outlined in section 2.11 to associate a set of trained parameters with each new image patch for which we require prediction of hypoxia. This method effectively restricts the parameter range during the validation process. Applying the parameter assignment method, as detailed in section 2.12, reveals reduced variability in the parameters for individual patches—for example, one from breast tumor tissue (0520-25-2) and another from pancreatic tumor tissue (1087)—in the validation set. The former tissue image is an IF scan, while the latter is an IHC scan image. This reduced variability observation is evident from the distribution plots shown in figures 8(b) and S3.

3.3.2. Prediction based on a restricted set of parameters and assignment of a state of vascularization

Using the stratification method, we narrow down parameter intervals, resulting in a reduced range for each parameter. While not providing an exact value, this interval offers a range of possibilities reflecting the underlying dynamics in the tissue. The variations in individual parameters signify different underlying dynamics in the concerned tissue. To better understand these dynamics, we analyze the changes in specific parameters one at a time while keeping the other parameters at the mean of their obtained interval for a patch sample. For the given blood vessel in figure 9(a) of the patch, oxygen and hypoxia distribution is predicted, as shown in figures 9(b) and (d), whereas the CA9 marker density is depicted in figure 9(c). By exploring the range of parameters obtained for this patch, we observe the following:



- Diffusion coefficient (D): As mentioned, oxygen supply is assumed to be the diffusion of oxygen molecules produced by blood vessels. This diffusion is controlled by the diffusion coefficient (D). In a properly oxygenated tissue microenvironment, O_2 molecules would be acquiring the maximum limit of the diffusion coefficient, resulting in less hypoxia, as seen in figure 10(a). Conversely, less diffusion would lead to more hypoxia, as depicted in figure 10(d).
- Production and uptake coefficients: The production coefficient (β), related to blood vessels, characterizes the state of blood vessel functioning. Higher values of β reflect increased oxygen production, resulting in lower hypoxia. On the other hand, the Oxygen uptake (α) rate is modulated by the tumor microenvironment and cellular metabolism and plays a crucial role in the hypoxia profile. Increased uptake leads to more hypoxia, as demonstrated in figure 10(b) for a minimum value and for a maximum value of α in figure 10(e).
- Decay coefficient (k_1) and conversion coefficient (γ): The decay coefficient k_1 , which signifies the rate of decay of CA9 marker with increasing oxygen concentration, exhibits a high hypoxic response at minimum values (figure 10(f)), whereas maximum values result in low hypoxia (figure 10(c)). A high value of k_1 , such as the maximum selected parameter, suggests a microenvironment where hypoxia diminishes rapidly with the concentration of oxygen. This could indicate a tissue state with strong oxygen needs, that reacts abruptly to a lack of oxygen. Conversely, a low value represents a scenario where the tissue is less dependent on oxygen, as in the case of anaerobic metabolism. As γ is a conversion coefficient limiting the maximum value of hypoxia distribution, a high value of γ implies a state with high oxygen demands, while a lower value suggests lower oxygen dependence.



Combining these parameters can help profile the blood vessel's functioning, with several possibilities of parameter combinations reflecting different dynamic conditions of tumor oxygenation. Here, we focus on two scenarios: functional and non-functional blood vessels. Functional blood vessels may result from efficient diffusion, low oxygen uptake in tumor tissues, and a high decay coefficient of hypoxia with increasing O_2 . In contrast, the non-functional category includes less diffusion, less production, and high oxygen uptake—where even with high production, this could be the case for new vessel linings. The plots in figure 10 showcase plots corresponding to these scenarios, extracting the dynamic nature of the dynamics through this static model.

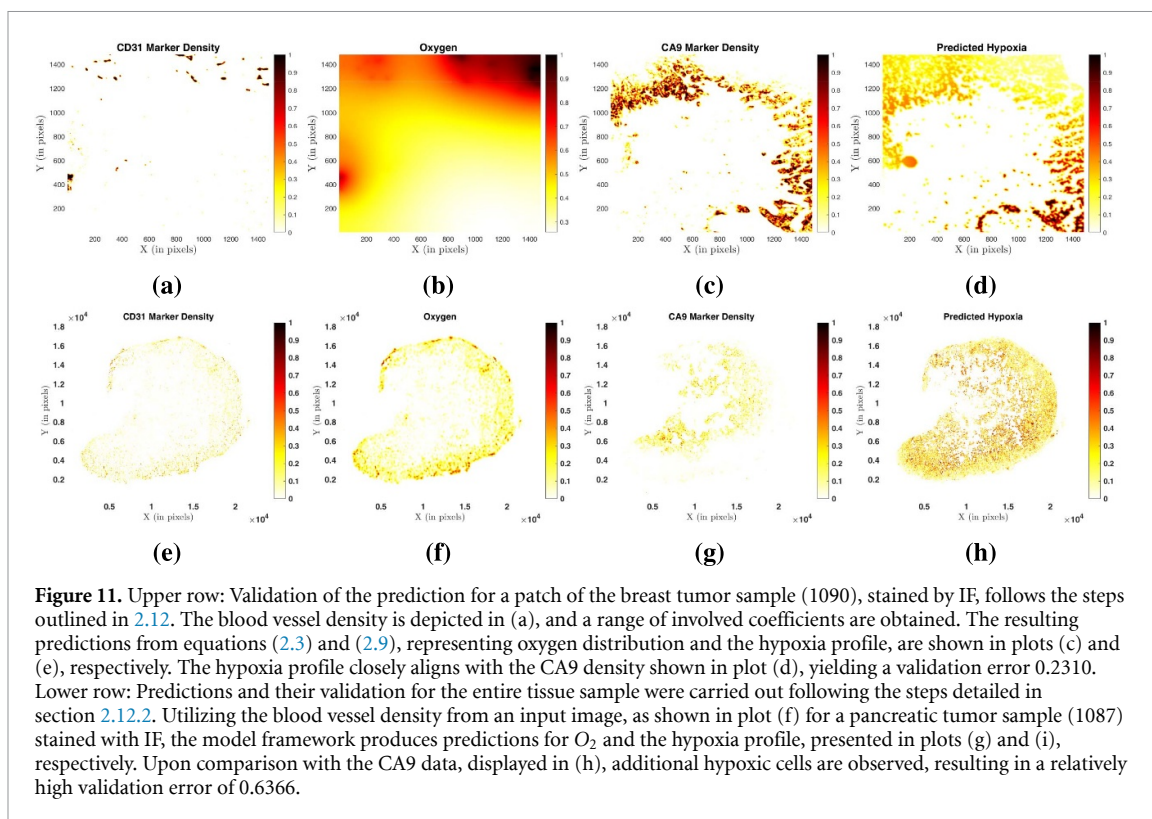
3.4. Validation of the physical model using diverse tumor samples

A desirable feature of a mathematical modeling framework is its predictive nature, coupled with a mechanistic explanation. Moreover, it is crucial to validate these predictions to ensure biological relevance. As detailed in section 2.12, we validate the model predictions using the outlined procedure, beginning with the patch validation. In this section, we present the results, demonstrating alignment with the mean values of the obtained parameters. Although we have intervals for each parameter, the validation results are consistent with the mean values for many of these samples. The validation was performed using a diverse set of tumor samples (see table 3).

3.4.1. Validation for patches

Our model predictions demonstrate strong alignment with the validation data, as evident in figures 11(a)–(e) and figure S4 for IF, and figure S6 for the IHC scan type. The predicted hypoxia profiles in figures 11(e) and S4(d) showcase the recovery of CA9 positive cells, utilizing the mean of the parameters after the classification process detailed in the previous section. While most CA9-positive cells are well-recovered, there are some additional tissue regions with cells marked as hypoxia-positive. However, these instances are relatively low compared to the well-aligned predictions. These additional positive markings may arise from various approximations and simplifications incorporated in our modeling process.

Despite the presence of some additional cells marked as CA9-positive with a slight gradient in the previously shown patch prediction, it is crucial to highlight that the modeling framework's predictions exhibit excellent alignment with patches devoid of CA9. This is evident in the results obtained for a patch from an ovarian tumor sample (0721) sample, illustrated in figure S5. The scatter distribution of blood



vessels in figure S5(a) is well-captured by the framework, leading to a uniformly distributed oxygen pattern and corresponding absence of hypoxic regions, as depicted in figures S5(b) and (d). Notably, this closely resembles the CA9 profile in figure S5(c) with a very small validation error.

3.4.2. Validation for whole sample

Detailed information at the patch level is valuable for examining microenvironments, but predicting hypoxia profiles for entire samples is equally crucial. Following the procedure outlined in section 2.12, we extend our predictions to whole samples, utilizing both IF and IHC scan types for testing. As illustrated in figures 11(e)–(h) for an IF scan from a breast tumor tissue and figure S7 for an IHC scan from a pancreas tumor, our predictions align well with the CA9 marker density. Similar to the observations at the patch level, our predictions cover CA9-positive cells and capture additional cells with a low hypoxia gradient.

It is essential to emphasize that identifying hypoxic cells is accurate; however, there are instances of false hypoxic regions being detected. This may be attributed to the coexistence of blood vessels and hypoxia or the non-availability of similar types of samples in the dataset.

4. Discussion

The well-established influence of hypoxia-driven alterations in the tumor microenvironment underscores the need for accurate assessments of oxygen and hypoxia distributions in effective cancer characterization and treatment design. We present a mechanistic framework applicable to various cancer types to predict hypoxic regions within the tumor microenvironment from tissue sections. This represents a crucial initial step towards a realistic model for hypoxia prediction and aiding in radiation dose estimation based on existing models (Dasu *et al* 2005, Titz and Jeraj 2008, Powathil *et al* 2012, Toma-Dasu and Dasu 2013). Understanding hypoxia is not only pivotal for radiation therapy efficacy but also essential for comprehending other hypoxia-induced phenomena, such as metabolic reprogramming, angiogenesis, tumor invasion and metastasis. In this study, we employ a simple mathematical approach capable of replicating biologically relevant hypotheses.

This data-driven yet mechanistic approach involves obtaining proxies for blood vessels, hypoxia, and cell nuclei (CD31, CA9, DAPI markers). We have formulated a mathematical model equation (2.3) to derive oxygen distribution from blood vessel density and determine hypoxia profiles equation (2.9). Optimization processes extract parameters by comparing model predictions with CA9 markers. The model is trained on tissue image patches, yielding parameter distributions. While similar results are obtained with models incorporating cell-dependent oxygen uptake (equation (S.1)) or introducing nonlinearity through limited

oxygen production and uptake equation (2.5) and (S.2), we choose to utilize the linear PDE (equation (2.3)). This decision is motivated by computational efficiency concerns associated with solving nonlinear PDEs multiple times for optimization. Furthermore, sticking with the linear model helps mitigate overfitting and reduce epistemic uncertainty stemming from additional parameters.

Assigning a class to each tissue patch proves beneficial to address the heterogeneity in blood vessels across different tissues and within the same tissue. We classify each patch based on its blood vessel architecture using the spatial entropy measure (refer to equation (2.1)). This classification is integral to our prediction framework, where the first step involves identifying the patch class and obtaining the corresponding parameters. Additionally, a range of parameters is obtained for each classified patch, enabling the observation of different dynamics within the tissue.

Recognizing the limitations posed by the assumption of constant diffusion, production, and uptake of O_2 molecules, we introduce a data-driven method to ascertain spatially varying coefficients. The method is based on the spatial entropy of patches from CD31 stained tissue images, summarizing the heterogeneity of blood vessels distributions, but the same approach can be adapted to other types of data such as those obtained from hematoxylin and eosin (H&E) stained tissue sections. Utilizing observed variations from trained and similar (classified) patches, we have employed a varying coefficient model (refer to equation (2.11)) to quantify this heterogeneity. Through this process, we have deduced a reaction-diffusion PDE wherein the coefficients exhibit variability between the considered patches. This approach facilitates the use of tailored rates for different patch types, enabling predictions for the entire tissue sample using a simplified model.

Our modeling predictions perform well for both types of scans, IHC and IF, commonly used in practice for staining different markers. Specifically, IF is preferred due to computational efficiency, eliminating the need for image registration. Additionally, it utilizes the same tissue in experiments, simplifying both experimental and computational aspects as the cell structures remain consistent. In addition, since H&E-stained tissue sections are produced as a routine procedure for cancer diagnosis, we expect our model to provide additional interesting information in the future that may help pathologists evaluate hypoxia on this type of standardized biological material.

We have identified additional hypoxic regions beyond CA9 marker densities, possibly due to the 2D approximation of the 3D blood vessel architecture. To address this, we plan to enhance our model by incorporating the 3D geometry and architecture of blood vessels, moving beyond the 2D slices used in the current estimation. While 3D reconstructions, exemplified in the study (Arslan *et al* 2023), provide a more precise representation, its practicality in clinical settings is constrained by tissue availability and the time consuming step of generating large series of adjacent tissue sections. To overcome this limitation, we are developing a framework to quantify the approximation error when using 2D datasets, crucial for evaluating hypoxia profiles in scenarios where 3D reconstruction is limited.

While CA9 is often considered an intrinsic marker for hypoxia, studies such as Hedley *et al* (2003), Mayer *et al* (2005) for uterine cervix cancer and Iakovlev *et al* (2007) for cervical carcinoma suggest a poor correlation between tumor oxygenation and CA9. This raises the need for exploring alternative markers or combinations of markers to improve the accuracy of hypoxia assessment. Additionally, the heterogeneity in oxygen consumption driven by cell metabolism is a crucial factor influencing hypoxia levels. Beyond the static assumptions and the 3D to 2D approximation, the dynamic state of cells may play a significant role in determining the hypoxia profile.

Predicting or estimating hypoxia without a time series dataset is valuable in various scenarios, especially given dynamic experiment challenges. Creating a profile based on the cell metabolic state and the functioning of blood vessels becomes essential, even if done less quantitatively, as attempted in our simulations for the diffusion coefficient, oxygen production, and uptake coefficients. Scores such as hypoxia scores (as seen in Van Laarhoven *et al* (2006)) may need more information, as they often overlook spatial heterogeneity in both cells and hypoxia when applied to an entire image. Experimental investigations involving cell metabolic states and other markers to assess the state of blood vessels would be essential for validating the proposed profiling approach.

5. Conclusion

This study introduces an innovative data-driven yet mechanistic tool for characterizing tumor oxygenation, effectively capturing the intricate dynamics of the tumor microenvironment. The computational framework successfully reproduces these complex dynamics, as demonstrated in the context of three distinct tumor types: breast, ovarian, and pancreatic. The results not only facilitate localized predictions for small tissue regions but also offer a data-driven approach to extract spatially varying rates related to oxygen (such as diffusion coefficient, production, and uptake rate). This spatially resolved information is crucial for capturing and predicting spatial heterogeneity across larger tissue samples. Additionally, based on the

derived oxygen-related rates, the study explores the potential to predict whether tumor vascularization is in a functioning or non-functioning state.

This approach has multiple applications. In fundamental research, it can serve as the basis for more realistic models of tumor microenvironment heterogeneity, applicable to tumors undergoing various types of treatment. Along this line of research, we intend to enhance current models by incorporating the dynamics of cellular adaptation to local hypoxia. As a further enhancement, we also plan to expand the model dimension from 2D to 3D and validate the 3D models using serial sectioning of tumors. In translational research, having effective tools for predicting hypoxia could prove valuable in guiding radiotherapy, especially when coupled with imaging techniques for tumor blood vessel reconstruction. These techniques include positron emission tomography, computed tomography, magnetic resonance angiography, ultrafast ultrasound imaging, or their combinations (Provost *et al* 2018, Facchin *et al* 2020, Moog *et al* 2022, Denis *et al* 2023).

Data availability statement

The data that support the findings of this study will be openly available following an embargo at the following URL/DOI: [10.5281/zenodo.10796880](https://doi.org/10.5281/zenodo.10796880). Data will be available from 1 November 2024.

Acknowledgments

We acknowledge financial support from ITMO Cancer on funds administered by INSERM (project MALMO) and a support by Mr Jean-Paul Baudecroux and The Big Brain Theory Program—Paris Brain Institute (ICM). We are grateful for the technical support of the Montpellier Rio Imaging (MRI) and animal (RAM) facilities. Histological analyses were performed thank to the ‘Réseau d’Histologie Expérimentale de Montpellier’ (RHEM), a facility supported by the SIRIC Montpellier Cancer (Grant INCa-DGOS-INSERM-ITMO Cancer 18004), the european regional development foundation and the Occitanie region (FEDER-FSE 2014–2020 Languedoc Roussillon), REACT-EU (Recovery Assistance for Cohesion and the Territories of Europe), IBiSA. We express our gratitude to Sophie Baize for her handling of the histological blocks. This work was also publicly funded through ANR (the French National Research Agency) under the ‘Investissements d’avenir’ programme with the reference ANR-16-IDEX-0006 (Projet I-Site Muse ‘MEL-ECO’).

Ethical statement

The research was conducted in accordance with the principles embodied in the Declaration of Helsinki and PDTX models were established from fresh tumor fragments obtained from the Pathology Department at the Institut du Cancer de Montpellier (ICM) after informed consent of the patients. The study is approved by the ethic committee for animal welfare of the region Languedoc Roussillon (Comité d’Ethique en Expérimentation Animal Languedoc Roussillon), which is an accredited institution of the french ‘MINISTERE DE L’ENSEIGNEMENT SUPERIEUR, DE LA RECHERCHE ET DE L’INNOVATION’ (Agreement Number CEEA-LR-14 140).

ORCID iDs

P Kumar  <https://orcid.org/0000-0003-4600-1028>
M Lacroix  <https://orcid.org/0000-0002-1268-9058>
J Arslan  <https://orcid.org/0000-0003-2683-3775>
L Fenou  <https://orcid.org/0009-0005-2888-3145>
L Le Cam  <https://orcid.org/0000-0003-0325-878X>
D Racocanu  <https://orcid.org/0000-0002-9416-1803>
O Radulescu  <https://orcid.org/0000-0001-6453-5707>

References

- Al Tameemi W, Dale T P, Al-Jumaily R M K and Forsyth N R 2019 Hypoxia-modified cancer cell metabolism *Front. Cell Dev. Biol.* **7** 00004
- Altrock P M, Liu L L and Michor F 2015 The mathematics of cancer: integrating quantitative models *Nat. Rev. Cancer* **15** 730–45
- Arslan J *et al* 2023 Efficient 3D reconstruction of whole slide images in melanoma *Medical Imaging 2023: Digital and Computational Pathology* vol 12471 (SPIE) p 124711S
- Bedogni B and Powell M B 2009 Hypoxia, melanocytes and melanoma—survival and tumor development in the permissive microenvironment of the skin *Pigment Cell Melanoma Res.* **22** 166–74

- Carlson D J, Keall P J, Loo B W, Chen Z J and Brown J M 2011 Hypofractionation results in reduced tumor cell kill compared to conventional fractionation for tumors with regions of hypoxia *Int. J. Radiat. Oncol. Biol. Phys.* **79** 1188–95
- Cavadas M A, Nguyen L K and Cheong A 2013 Hypoxia-inducible factor (HIF) network: insights from mathematical models *Cell Commun. Signal.* **11** 42
- Claramunt C 2005 A spatial form of diversity *Spatial Information Theory: Int. Conf., COSIT 2005 (Ellicottville, NY, USA, 14 September–18 September 2005) Proc. 7* (Springer) pp 218–31
- Colombo P E et al 2015 Ovarian carcinoma patient derived xenografts reproduce their tumor of origin and preserve an oligoclonal structure *Oncotarget* **6** 28 327–28 340
- D’Aguanno S, Mallone F, Marengo M, Del Bufalo D and Moramarco A 2021 Hypoxia-dependent drivers of melanoma progression *J. Exp. Clin. Cancer Res.* **40** 1–32
- Daşu A, Toma-Daşu I and Karlsson M 2003 Theoretical simulation of tumour oxygenation and results from acute and chronic hypoxia *Phys. Med. Biol.* **48** 2829–42
- Daşu A, Toma-Daşu I and Karlsson M 2005 The effects of hypoxia on the theoretical modelling of tumour control probability *Acta Oncol.* **44** 563–71
- Denis L et al 2023 Sensing ultrasound localization microscopy for the visualization of glomeruli in living rats and humans *eBioMedicine* **91** 104578
- du Manoir S et al 2014 Breast tumor PDXs are genetically plastic and correspond to a subset of aggressive cancers prone to relapse *Mol. Oncol.* **8** 431–43
- Facchin C et al 2020 Concurrent imaging of vascularization and metabolism in a mouse model of paraganglioma under anti-angiogenic treatment *Theranostics* **10** 3518–32
- Gallez B, Neveu M-A, Danhier P and Jordan B F 2017 Manipulation of tumor oxygenation and radiosensitivity through modification of cell respiration. A critical review of approaches and imaging biomarkers for therapeutic guidance *Biochim. Biophys. Acta—Bioenerg.* **1858** 700–11
- Ginouès A, Ilc K, Macías N, Pouysségur J and Berra E 2008 PHDs overactivation during chronic hypoxia “desensitizes” HIF α and protects cells from necrosis *Proc. Natl Acad. Sci. USA* **105** 4745–50
- Goldman D 2008 Theoretical models of microvascular oxygen transport to tissue *Microcirculation* **15** 795–811
- Grimes D R, Kannan P, Warren D R, Markelc B, Bates R, Muschel R and Partridge M 2016 Estimating oxygen distribution from vasculature in three-dimensional tumour tissue *J. R. Soc. Interface* **13** 20160362
- Grzywa T M, Paskal W and Włodarski P K 2017 Intratumor and intertumor heterogeneity in melanoma *Trans. Oncol.* **10** 956–75
- Harris A L 2002 Hypoxia—a key regulatory factor in tumour growth *Nat. Rev. Cancer* **2** 38–47
- Hedley D, Pintilie M, Woo J, Morrison A, Birle D, Fyles A, Milosevic M and Hill R 2003 Carbonic anhydrase IX expression, hypoxia and prognosis in patients with uterine cervical carcinomas *Clin. Cancer Res.* **9** 5666–74 (available at: <https://aacrjournals.org/clincancerres/article/9/15/5666/202195/Carbonic-Anhydrase-IX-Expression-Hypoxia-and?searchresult=1>)
- Hirota K and Semenza G L 2006 Regulation of angiogenesis by hypoxia-inducible factor 1 *Crit. Rev. Oncol. Hematol.* **59** 15–26
- Hodgkinson A, Trucu D, Lacroix M, Le Cam L and Radulescu O 2022 Computational model of heterogeneity in melanoma: designing therapies and predicting outcomes *Front. Oncol.* **12** 1245
- Horsman M R and van der Kogel A J 2009 Therapeutic approaches to tumour hypoxia *Basic Clin. Radiobiol.* **20** 233
- Hsu R and Secomb T W 1989 A Green’s function method for analysis of oxygen delivery to tissue by microvascular networks *Math. Biosci.* **96** 61–78
- Iakovlev V V, Pintilie M, Morrison A, Fyles A W, Hill R P and Hedley D W 2007 Effect of distributional heterogeneity on the analysis of tumor hypoxia based on carbonic anhydrase IX *Lab. Invest.* **87** 1206–17
- Kapuscinski J 1995 Dapi: a DNA-specific fluorescent probe *Biotech. Histochem.* **70** 220–33
- Kelly C J and Brady M 2006 A model to simulate tumour oxygenation and dynamic [18F]-Fmiso PET data *Phys. Med. Biol.* **51** 5859–73
- Kempf H, Bleicher M and Meyer-Hermann M 2015 Spatio-temporal dynamics of hypoxia during radiotherapy *PLoS One* **10** e0133357
- Korbecki J, Simińska D, Gaśowska-Dobrowolska M, Listos J, Gutowska I, Chlubek D and Baranowska-Bosiacka I 2021 Chronic and cycling hypoxia: drivers of cancer chronic inflammation through HIF-1 and NF- κ b activation: a review of the molecular mechanisms *Int. J. Mol. Sci.* **22** 10701
- Krogh A 1919 The number and distribution of capillaries in muscles with calculations of the oxygen pressure head necessary for supplying the tissue *J. Physiol.* **52** 409–15
- Kumar P and Surulescu C 2020 A flux-limited model for glioma patterning with hypoxia-induced angiogenesis *Symmetry* **12** 1870
- Lin A and Maity A 2015 Molecular pathways: a novel approach to targeting hypoxia and improving radiotherapy efficacy via reduction in oxygen demand *Clin. Cancer Res.* **21** 1995–2000
- Mabjeesh N J and Amir S 2007 Hypoxia-inducible factor (HIF) in human tumorigenesis *Histol. Histopathol.* **22** 559–72
- Mayer A, Höckel M and Vaupel P 2005 Carbonic anhydrase IX expression and tumor oxygenation status do not correlate at the microregional level in locally advanced cancers of the uterine cervix *Clin. Cancer Res.* **11** 7220–5
- McKeown S R 2014 Defining normoxia, physoxia and hypoxia in tumours—implications for treatment response *Br. J. Radiol.* **87** 1–12
- Moog S et al 2022 Preclinical evaluation of targeted therapies in *Sdhb*-mutated tumors *Endocrine-Related Cancer* **29** 375–88
- Nguyen L K et al 2015 A dynamic model of the hypoxia-inducible factor 1 α (HIF-1 α) network *J. Cell Sci.* **128** 422
- Powathil G, Kohandel M, Milosevic M and Sivaloganathan S 2012 Modeling the spatial distribution of chronic tumor hypoxia: implications for experimental and clinical studies *Comput. Math. Methods Med.* **2012** 1–11
- Provost J et al 2018 Simultaneous positron emission tomography and ultrafast ultrasound for hybrid molecular, anatomical and functional imaging *Nat. Biomed. Eng.* **2** 85–94
- Pusztaszeri M P, Seelentag W and Bosman F T 2006 Immunohistochemical expression of endothelial markers CD31, CD34, von willebrand factor and FLI-1 in normal human tissues *J. Histochem. Cytochem.* **54** 385–95
- Qutub A A and Popel A S 2006 A computational model of intracellular oxygen sensing by hypoxia-inducible factor HIF1 α *J. Cell Sci.* **119** 3467–80
- Saxena K and Jolly M K 2019 Acute vs. chronic vs. cyclic hypoxia: their differential dynamics, molecular mechanisms and effects on tumor progression *Biomolecules* **9** 1–27
- Secomb T W, Hsu R, Dewhirst M W, Klitzman B and Gross J F 1993 Analysis of oxygen transport to tumor tissue by microvascular networks *Int. J. Radiat. Oncol. Biol. Phys.* **25** 481–9
- Secomb T W, Hsu R, Ong E T, Gross J F and Dewhirst M W 1995 Analysis of the effects of oxygen supply and demand on hypoxic fraction in tumors *Acta Oncol.* **34** 313–6

- Secomb T W, Hsu R, Park E Y and Dewhirst M W 2004 Green's function methods for analysis of oxygen delivery to tissue by microvascular networks *Ann. Biomed. Eng.* **32** 1519–29
- Skeldon A C, Chaffey G, Lloyd D J, Mohan V, Bradley D A and Nisbet A 2012 Modelling and detecting tumour oxygenation levels *PLoS One* **7** 1–17
- Suh J H et al 2006 Phase III study of efaproxiral as an adjunct to whole-brain radiation therapy for brain metastases *J. Clin. Oncol.* **24** 106–14
- Tarnowski B I, Spinale F G and Nicholson J H 1991 Dapi as a useful stain for nuclear quantitation *Biotech. Histochem.* **66** 296–302
- Titz B and Jeraj R 2008 An imaging-based tumour growth and treatment response model: investigating the effect of tumour oxygenation on radiation therapy response *Phys. Med. Biol.* **53** 4471–88
- Toma-Dasu I and Dasu A 2013 Modelling tumour oxygenation, reoxygenation and implications on treatment outcome *Comput. Math. Methods Med.* **2013** 1–9
- Turner K J, Crew J P, Wykoff C C, Watson P H, Poulson R, Pastorek J, Ratcliffe P J, Cranston D and Harris A L 2002 The hypoxia-inducible genes VEGF and CA9 are differentially regulated in superficial vs invasive bladder cancer *Br. J. Cancer* **86** 1276–82
- van den Beucken T et al 2009 Hypoxia-induced expression of carbonic anhydrase 9 is dependent on the unfolded protein response *J. Biol. Chem.* **284** 24204–12
- Van Laarhoven H W, Kaanders J H, Lok J, Peeters W J, Rijken P F, Wiering B, Ruers T J, Punt C J, Heerschap A and Van Der Kogel A J 2006 Hypoxia in relation to vasculature and proliferation in liver metastases in patients with colorectal cancer *Int. J. Radiat. Oncol. Biol. Phys.* **64** 473–82
- Vaupel P, Flood A B and Swartz H M 2021 Oxygenation status of malignant tumors vs. normal tissues: critical evaluation and updated data source based on direct measurements with pO₂ microsensors *Appl. Magn. Reson.* **52** 1451–79
- Vaupel P and Mayer A 2007 Hypoxia in cancer: significance and impact on clinical outcome *Cancer Metastasis Rev.* **26** 225–39
- Welter M, Fredrich T, Rinneberg H and Rieger H 2016 Computational model for tumor oxygenation applied to clinical data on breast tumor hemoglobin concentrations suggests vascular dilatation and compression *PLoS One* **11** 1–42
- Widmer D S, Hoek K S, Cheng P F, Eichhoff O M, Biedermann T, Raaijmakers M I, Hemmi S, Dummer R and Levesque M P 2013 Hypoxia contributes to melanoma heterogeneity by triggering HIF1 α -dependent phenotype switching *J. Investigative Dermatol.* **133** 2436–43