Macrophage sensitivity to microenvironmental cues influences spatial heterogeneity of tumours

Joshua A. Bull[1,*] and Helen M. Byrne[1]

1 Wolfson Centre for Mathematical Biology, Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK

* joshua.bull@maths.ox.ac.uk

Abstract

Cross-talk between tumour and immune cells, including macrophages, is complex, dynamic and contributes to tumour heterogeneity. In this paper, we introduce a hybrid agent-based model (ABM) to investigate how tumour-macrophage dynamics evolve over time and how they influence spatial patterns of tumour growth. Macrophage phenotype is determined by microenvironmental cues and governs the extent to which macrophages are pro- or anti- tumour, i.e., whether they infiltrate and attack tumour cells, or promote metastasis by guiding tumour cell migration towards blood vessels.

We perform extensive ABM simulations to investigate how changes in macrophage sensitivity to microenvironmental cues – specifically, cytokines produced by tumour cells and perivascular fibroblasts – alters their spatial and phenotypic distributions and the tumour's growth dynamics. We identify outcomes that include: compact tumour growth, tumour elimination, and diffuse patterns of invasion, characterised by clustering of tumour cells and macrophages around blood vessels.

We compare the ability of different statistics to characterise the diverse spatial patterns that the ABM generates. These include the weighted pair correlation function (wPCF), a new statistic that quantifies the spatial distributions of cells labelled with a continuous value (e.g., macrophage phenotype). We assess the ability of each statistic to discriminate between the various spatial patterns that the ABM exhibits.

By combining statistical analysis with ABM simulations, we show how mechanistic models can be used to generate synthetic data for validation of novel statistics (here, the wPCF) and to assess the extent to which specific statistical descriptions can distinguish different spatial patterns and model behaviours. Such statistics can then be applied to biological imaging data, such as multiplexed medical images, with increased confidence in the interpretation of the results. We show that the wPCF accurately describes differences in macrophage localisation between images, and posit that it would be a valuable tool for analysing multiplexed imaging data.

Author summary

Macrophage phenotype is regulated by complex microenvironmental cues. It affects their spatial position and behaviour. In solid tumours, the spatial distribution of macrophages can vary significantly, and correlate with patient prognosis.

In this paper, we use an agent-based model (ABM) to investigate how changes in sensitivity to tumour-induced microenvironmental cues affect macrophage phenotype and, in turn, how such phenotypic heterogeneity affects tumour composition and

morphology. We illustrate the wide range of tumour outcomes that can arise from changing macrophage sensitivity to microenvironmental cues.

We apply a variety of statistics to simulation outputs to characterise the range of observed spatial patterns. Different statistics identify different aspects of these patterns, with the most descriptive characterisations obtained from spatial statistics, such as the weighted pair correlation function (wPCF), that account for both phenotype and cell localisation.

More generally, this paper illustrates how ABMs can be used to generate synthetic data which mimics data that can be extracted from different imaging modalities. This synthetic data can be used to test new statistics, like the wPCF, and validate their use for future application to multiplex histological imaging.

Introduction

The immune landscape within solid tumours is complex and varied [1, 2], with innate and adaptive immune cells implicated in pro- and anti-tumour responses [3]. For example, high densities of tumour associated macrophages have been associated with poor prognosis in breast, prostate and head and neck cancer and with good prognosis in colorectal and gastric cancer [4, 5]. These differences may be due to differences in the numbers of pro- and anti-tumour macrophages in these cancers. Indeed, macrophages are often described as having an anti-tumour, "M_1" phenotype or a pro-tumour, "M_2" phenotype. Further the number of macrophages of each phenotype, their morphology and spatial distribution are associated with patient survival [6–9]. For example, in non-small cell lung cancer high infiltration of M_1 macrophages into tumour islets, but not tumour stroma, has been associated with increased patient survival [10].

Macrophage phenotype may be viewed as a continuous variable, with the balance of pro- and anti-tumour behaviours determined by integrating multiple, microenvironmental cues [11]. Until recently, it has been difficult to determine simultaneously macrophage phenotype and spatial position in clinical or experimental settings. Classical immunohistochemistry (IHC) resolves spatial positions but only distinguishes broad macrophage subtypes. Single-cell sequencing more accurately describes cell phenotypes but lacks information about spatial position. Multiplex imaging modalities, such as multiplexed IHC and imaging mass cytometry (IMC), provide information about the spatial position of different macrophage phenotypes, mapping expression levels of up to 40 different cellular markers [7, 12–14]. Post-processing these data can be challenging. Metrics such as cell counts ignore spatial information. More complex statistics, that account for cell position, include measurements of distances between nearby cells [15], metrics based on immune cell infiltration [16–19], and cell morphology and topology [20, 21]. Spatial statistics, such as the pair correlation function (PCF), have also been used to describe immune cell localisation within tumours [22].

In this paper, we investigate how macrophage-tumour interactions and, in particular, macrophage responses to environmental cues, impact tumour growth and composition. We develop a new off-lattice, force-based, agent-based model (ABM) to study the effect that macrophage recruitment and phenotype plasticity have on tumour dynamics, and use this model as a testbed for comparing statistics that can be applied to multidimensional spatially resolved imaging data. The ABM introduces a continuous and dynamic variable that describes macrophage phenotype, and accounts for paracrine signalling between the tumour cells and macrophages. Our model is motivated by observations of macrophage infiltration into tumour regions [23]. M_1-like macrophages extravasate and migrate towards the tumour in response to tumour-derived signals such as colony stimulating factor-1 (CSF-1) and chemokine C-C motif ligand 2 (CCL2).

Exposure to TGF-β in the tumour increases macrophage sensitivity to the chemokine C-X-C chemokine ligand type 12 (CXCL12) and drives them to an M_2-like phenotype. CXCL12 is produced by perivascular fibroblasts and biases the movement of M_2-like macrophages towards neighbouring blood vessels. As they migrate out of the tumour, M_2-like macrophages express epidermal growth factor (EGF), a tumour cell chemoattractant. In this way, the M_2-like macrophages facilitate metastasis by guiding the tumour cells towards the vasculature [23, 24].

We use our ABM to simulate the above behaviour and also to show how altering macrophage sensitivity to microenvironmental cues alters their spatial and phenotypic distributions and, in so doing, can generate a diverse range of tumour growth patterns. By varying two parameters associated with macrophage response to CSF-1 (describing the extravasation rate and chemotactic sensitivity to CSF-1 gradients) we identify outcomes including tumour elimination, compact tumour growth, and the formation of perivascular niches containing M_2-like macrophages and tumour cells. We also show that varying a larger number of the parameters that regulate macrophage responses to environmental cues generates a wider range of behaviours. We apply a range of statistics (e.g., cell counts and pair correlation functions) to synthetic data extracted from ABM simulations and investigate the extent to which they can describe the spatial and phenotypic patterns that the ABM generates. We also introduce the weighted pair correlation function in order to describe spatial variations in macrophage phenotype, and use our model to demonstrate its ability to distinguish different simulation outcomes.

Modelling approaches

Agent-based models (ABMs) are ideally suited for studying complex biological tissues. Each cell is represented by an agent whose behaviour is determined by subcellular variables (describing, for example, cell cycle state [25] or phenotype), and its interaction with the environment (e.g., through force laws describing mechanical interactions between cells). Hybrid ABMs combine this approach with other modelling techniques, for instance by using PDEs to describe the concentration of diffusible species throughout the domain or by using ODEs to describe the dynamics of subcellular variables.

ABMs describing tumour growth have been developed within a range of software frameworks, including PhysiCell [26], HAL (Hybrid Automata Library) [27], CompuCell3D [28, 29] and Chaste (Cancer, Heart and Soft Tissue Environment) [30–32]. These models have been used to study avascular and vascular tumour growth, tumour angiogenesis, tumour-immune interactions and also to simulate tumour responses to treatments that include chemotherapy, radiotherapy, immunotherapy and their combinations [36, 37]. In this paper, we use the Chaste framework to investigate the impact of microenvironmental cues on the spatial and phenotypic distribution of macrophages and their ability to control tumour growth.

Many models of macrophage-tumour interactions view macrophages as a homogeneous population, and do not distinguish phenotype. These include continuum models, formulated using partial differential equations (PDEs), which describe macrophage infiltration into avascular tumours [38, 39]. In [40] a hybrid model (combining both PDEs and ABM) was used to investigate the feasibility of engineering macrophages to target the delivery of cytotoxic drugs to hypoxic tumour sites. Norton *et al.* [41] developed a 3D lattice-based model of triple-negative breast cancer in which macrophages and fibroblasts interact with vasculature and tumour-derived cytokines such as CSF-1. Alternative models account for multiple macrophage subtypes (typically M_1 and M_2) [42–45], and their interactions with T-cells [46–48]. Eftimie [49] and El-Kenawi *et al.* [50] have developed models that view macrophage phenotype as a continuous variable whose dynamics are governed by environmental cues, such as pH.

The hybrid ABM developed in this paper builds on existing differential equation models [51, 52] and ABMs [52, 53] that focus on specific tumour-macrophage interactions, such as the CSF-1/EGF paracrine loop that mediates cross-talk between tumour cells and macrophages. Our model accounts for macrophage extravasation in response to tumour-derived CSF-1, their subsequent tumour infiltration, and the CSF-1/EGF paracrine loop that mediates cross-talk between tumour cells and macrophages. We represent macrophage phenotype as a continuous variable, p. Its dynamics depend on local levels of TGF-β and determine whether a macrophage is M_1-like (i.e., likely to kill tumour cells), M_2-like (i.e., benign towards tumour cells and expressing EGF), or of an intermediate phenotype, with intermediate behaviour.

Spatial statistics

A range of spatial statistics can be used to describe the spatial patterns of tumour cells, blood vessels, and macrophages that our ABM generates. The pair correlation function (PCF) quantifies spatial clustering and dispersion of a single cell type (e.g., macrophages or tumour cells) [34, 54–58]. The cross-PCF extends the PCF to account for multiple cell types, and has been used to identify correlations in the spatial patterns generated by two interacting cell populations [59, 60].

The PCF and cross-PCF quantify spatial correlation at multiple length scales for cell populations with discrete, categorical labels. In practice, many point clouds are labelled with continuous variables, or 'marks'. These marks may represent protein intensity in IMC, macrophage phenotype in our ABM, or the concentration of a diffusible cytokine such as TGF-β or CSF-1. While 'marked point patterns' have been studied in ecology [61–63] and astronomy [64, 65], few methods consider the spatial correlation of continuous marks. Arguably the most common is Stoyan's mark correlation function, $k_{mm}(r)$ [61, 64, 66, 67]. It determines whether the marks of two points separated by a distance r are spatially correlated: $k_{mm}(r) > 1$ indicates that the marks of points separated by r are larger than the average mark, while $k_{mm}(r) < 1$ indicates that they have marks smaller than the average [61]. The mark variogram can be used to test the similarity of two marks, separated by a distance r [63, 68].

The above approaches depend only on the distance between point pairs. In biological applications, we often wish to determine whether points with a particular mark, or range of marks, are correlated at distance r. Accordingly, in this paper, we introduce the weighted PCF (or wPCF), which weights the contribution from distinct points according to how close their marks are to a target mark. In this paper, we apply the wPCF to synthetic data generated from our ABM in order to determine the extent to which macrophages at distance r from tumour cells have a target phenotype p. We also use the wPCF to investigate how, for a given simulation, the distribution of macrophage phenotypes changes as the distance r varies.

Paper outline

In this paper, we develop a hybrid ABM which simulates interactions between macrophages and tumour cells. By varying parameters associated with macrophage behaviour, we identify a range of different tumour morphologies including compact growth, fragmented growth and tumour elimination. We use the ABM simulations as a testbed to examine the utility of different statistical descriptions of the data, identifying information that could be extracted from different types of experimental or clinical data (e.g., IHC or IMC). We assess the ability of these statistics to discriminate the different spatial patterns that the ABM can generate, and conclude that care is needed to ensure that the selected statistics can adequately resolve changes of interest in the data.

Materials and methods

Agent-based model

Overview

In this Section, we present a 2D, multiscale, off-lattice ABM which extends an existing model of macrophage [34, 35] infiltration into tumour spheroids. We briefly describe the ABM here; for full details of the implementation and default parameter values, we refer the interested reader to S1 Appendix: Model Description. We then explain how macrophage phenotype and behaviour are incorporated into the ABM. The ABM is implemented within the Chaste (Cancer, Heart and Soft Tissue Environment) modelling environment [30–32].

The ABM distinguishes four cell types: stromal cells, tumour cells, necrotic cells, and macrophages. Their dynamics are influenced by five diffusible species: oxygen (ω), CSF-1 (c), CXCL12 (ξ), TGF-β (g) and EGF (ϵ). In the 2D Cartesian geometry, blood vessels are represented by fixed points which do not compete for space with the cell populations and which act as distributed sources of oxygen [33]. A schematic of the ABM is presented in Fig 1. Following [80], critical oxygen thresholds for hypoxia ($\omega_{\mathrm{H}}^{\mathrm{str}}$ and $\omega_{\mathrm{H}}^{\mathrm{tum}}$) and necrosis ($\omega_{\mathrm{N}}^{\mathrm{str}}$ and $\omega_{\mathrm{N}}^{\mathrm{tum}}$) relate the rates of cell cycle progression of stromal and tumour cells to local oxygen levels (see Fig 1A for details). For example, if, at time $t > 0$, $\omega_{\mathrm{N}}^{\mathrm{tum}} < \omega < \omega_{\mathrm{H}}^{\mathrm{tum}}$, then the cell cycle of a tumour cell at position $\mathbf{x}$ will immediately halt and remain paused until the local oxygen concentration rises above this threshold. However, if the oxygen concentration falls below the necrosis threshold, then the cell becomes necrotic (this switch is irreversible). Necrotic cells occupy space for a finite time period during which they decrease linearly in size, until they reach size 0 and are removed from the simulation. Blood vessels also act as entry points for macrophages, which infiltrate the tissue and alter their phenotype (and, hence, behaviour) at rates which depend on local levels of TGF-β (see Fig 1B).

We represent each cell by the spatial coordinates of its centre of mass and determine its movement by balancing the forces that act on each it. Using an overdamped form of Newton's second law and neglecting inertial terms, we have that for cell i

$$\nu \frac{d\mathbf{x}_i}{dt} = \mathbf{F}_i, \tag{1}$$

where ν is the assumed constant drag coefficient and $\mathbf{F}_i$ denotes the net force acting on cell i at position $\mathbf{x}_i$ and time t. The forces that act on a cell depend on its type (see Fig 1C and S1 Appendix: Model Description). Cells interact via spring forces if their centres are within a distance R_{int} of each other [81]; intercellular adhesion and volume exclusion are represented by attractive and repulsive forces respectively. We also associate with each cell an area, defined as the area of a circle with a radius equal to the cells average interaction radius with neighbouring cells, or its equilibrium radius where neighbouring cells are beyond this distance. Stromal cells which are so mechanically compressed that their area falls below a threshold proportion ${A_i^H}_{\mathrm{str}}$ of their target area pause their cell cycle due to contact-inhibition (see S1 Appendix: Model Description).

Macrophage phenotype

The behaviour of individual macrophages is determined by a continuous subcellular variable which represents their phenotype, $p \in [0, 1]$. We assume that, following extravasation, macrophage i has an M_1-like phenotype which corresponds to $p_i = 0$. Macrophage exposure to TGF-β levels above a threshold value, g_{crit}, causes p_i to increase at a constant rate Δp, per timestep dt, until the maximum value $p_i = 1$ is

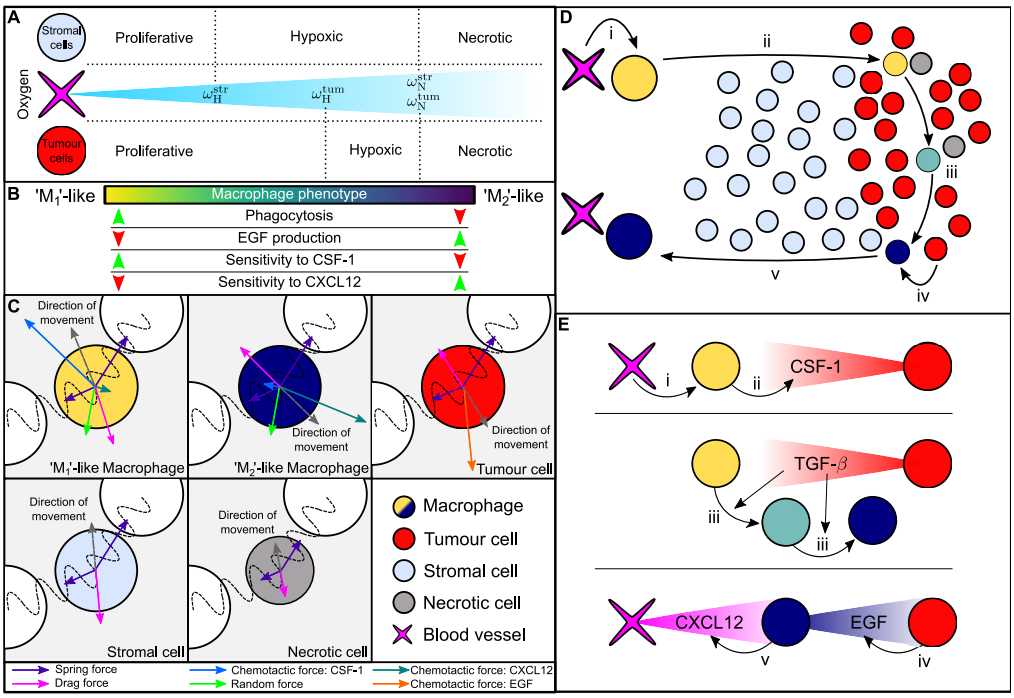

Fig 1. Schematic summarising the key interactions that are included in the agent based model. A: Oxygen is supplied by blood vessels and consumed by stromal cells and tumour cells. Cell-cycle progression is determined by a cell's local oxygen concentration: a cell may be 'proliferative' (and progress through its cell cycle), 'hypoxic' (the cell cycle is temporarily paused until oxygen concentrations return to a sufficiently high level) or 'necrotic' (the cell becomes necrotic cell and degrades). Cell cycles also pause if there is insufficient space available for proliferation.
B: Macrophage behaviour depends on phenotype p, modulating their rates of tumour cell killing, EGF production, and chemotactic sensitivity to gradients of CSF-1 and CXCL12.
C: Forces acting on different cell types. Macrophages are subject to mechanical forces due to interactions with nearby cells, and random forces which simulate their exploration of their environment as highly motile cells. Macrophages also experience chemotactic forces that are directed up spatial gradients of CSF-1 and CXCL12, and whose magnitude depends on p. Tumour cells experience mechanical forces due to interactions with neighbouring cells, and chemotactic forces in the direction of increasing EGF. Stromal cells experience mechanical forces due to interactions with neighbouring cells. Necrotic cells experience these interaction forces, which decrease in magnitude as they decrease in size. All cells experience a drag force.
D: Summary of the phases of macrophage-mediated tumour cell migration in our ABM. i) M_1-like macrophages extravasate from blood vessels in response to CSF-1. ii) M_1 macrophages migrate into the tumour mass in response to CSF-1, where they may kill tumour cells. iii) Exposure to TGF-β causes macrophages to adopt an M_2-like phenotype. iv) M_2-like macrophages produce EGF, which acts as a chemoattractant for tumour cells. v) M_2-like macrophages migrate towards blood vessels, in response to CXCL12 gradients.
E: Schematic summarising the sources of CSF-1, TGF-β, EGF and CXCL12 in our model, and their interactions with cells, as described in steps i-v of panel D.

reached and the macrophage has a fully M_2-like phenotype. Its phenotype remains fixed at $p_i = 1$ for all later times. Thus, we have:

$$\frac{dp_i}{dt} = \mathcal{H}(g(\mathbf{x}_i, t) - g_{\text{crit}})\mathcal{H}(1 - p_i)\Delta p, \tag{2}$$

where $\mathcal{H}$ is the Heaviside function ($\mathcal{H}(x) = 1$ when $x > 0$ and $\mathcal{H}(x) = 0$ otherwise).

We now explain how changes in phenotype p affect macrophage behaviour and function, and how these changes are incorporated into the ABM (see also Fig 1B).

Macrophage chemotactic forces Fig 1C shows the forces which act on different cell types (functional forms for these forces are given in S1 Appendix: Model Description). Here we highlight two, macrophage-specific, forces which describe their directed movement up spatial gradients of CSF-1 and CXCL12, and whose magnitude varies with phenotype p. Noting that M_1-like macrophages are sensitive to CSF-1 and insensitive to CXCL12 (and conversely for M_2-like macrophages), we assume that the chemotactic forces due to CSF-1 and CXCL12 depend linearly on phenotype. In particular, macrophages with $p = 0$ are maximally sensitive to CSF-1 and minimally sensitivity to CXCL12 (and conversely for macrophages with $p = 1$). Under these assumptions, the chemotactic forces acting on macrophage i, at position $\mathbf{x}_i$ with phenotype p_i are:

$$\mathbf{F}_i^{\chi_c} = \chi_c^m (1 - p_i) \frac{\nabla c}{|\nabla c|} \tag{3}$$

and

$$\mathbf{F}_i^{\chi_\xi} = \chi_\xi^m p_i \frac{\nabla \xi}{|\nabla \xi|} \tag{4}$$

respectively, where the non-negative parameters χ_c^m and χ_ξ^m indicate macrophage sensitivity to spatial gradients of CSF-1 and CXCL12, and ∇c and $\nabla \xi$ are evaluated at $\mathbf{x}_i$. The forces $\mathbf{F}_i^{\chi_c}$ and $\mathbf{F}_i^{\chi_\xi}$ contribute to the net force $\mathbf{F}_i$ in Eq (1) (see Fig 1C and S1 Appendix: Model Description).

Macrophage cell killing We assume that when the distance between a macrophage and a tumour cell is less than, or equal to, the interaction radius R_{int}, the macrophage will attempt to kill the tumour cell, with M_1-like macrophages more likely to kill a tumour cell than M_2-like macrophages. We define a probability of cell kill per hour, P_φ, which is a monotonic decreasing function of macrophage phenotype. We suppose further that, after a macrophage has killed a tumour cell, it experiences a 'cooldown' period of t_{cool} hours during which it cannot attempt tumour cell killing. Thus, we associate with macrophage i a subcellular timer $t_{\varphi,i}$ that is updated in real time and set to zero on tumour cell killing. We define $P_{\varphi,i}$ as:

$$P_{\varphi,i} = \begin{cases} P_\varphi^\star \times (1 - \frac{p_i^{10}}{p_i^{10} + 0.5^{10}}) & \text{for } t_{\varphi,i} \geq t_{\text{cool}} \\ 0 & \text{otherwise} \end{cases}, \tag{5}$$

where $P_\varphi^\star$ is the maximum probability of tumour cell killing. If multiple tumour cells are within distance R_{int} of macrophage i, then one is selected at random for cell death. Following cell killing, tumour cells are labelled as 'necrotic' and decay in the same way as other necrotic cells.

Macrophage extravasation Macrophages enter the domain via blood vessels with a probability per hour P_{ex} which is an increasing, saturating function of CSF-1:

$$P_{\text{ex}} = P^\star \times \frac{c}{c + c_{1/2}}, \tag{6}$$

where the non-negative parameter $P^\star$ represents the maximum probability per hour of macrophage extravasation from a vessel, and $c_{1/2}$ is the concentration of CSF-1 at which the probability is half-maximal.

Macrophage production of EGF The diffusible cytokine EGF, ϵ, is produced by M_2-like macrophages and undergoes natural decay. It is also a potent chemoattractant for tumour cells. For simplicity, we assume that macrophage i produces EGF at a rate which is linearly proportional to its phenotype p_i, with constant of proportionality κ_ϵ. Denoting by D_ϵ and λ_ϵ the assumed constant diffusion coefficient and natural decay rate of EGF, we suppose that its evolution can be described by the following reaction diffusion equation:

$$\frac{\partial \epsilon}{\partial t} = D_\epsilon \nabla^2 \epsilon - \lambda_\epsilon \epsilon + \kappa_\epsilon \sum_i p_i \delta(\mathbf{x} - \mathbf{x}_i). \tag{7}$$

where $\delta(\mathbf{x}) = \mathbf{1}$ when $\mathbf{x} = 0$ and $\delta(\mathbf{x}) = \mathbf{0}$ elsewhere. In (7), we sum over all macrophages to determine the net rate of production at spatial position $\mathbf{x}$.

Constructing synthetic datasets from model simulations

We post-process the output from ABM simulations at a fixed timepoint, to generate synthetic datasets which mimic information that can be collected from biological experiments (see Fig 2). We categorise these datasets as follows:

1. Raw cell counts of viable tumour cells (N_T) and macrophages (N_M) are recorded, with spatial and phenotypic information neglected (Fig 2A);
2. Spatial locations and cell types are recorded, with macrophage phenotype is neglected (synthetic IHC data: Fig 2B)
3. Raw cell counts and macrophage phenotype are recorded, with spatial information neglected (synthetic single cell sequencing data: Fig 2C);
4. Spatial location, cell type and macrophage phenotype are recorded (synthetic IMC data: Fig 2D)

Data analysis

Before defining the statistics that we apply to the synthetic datasets, we introduce some notation. Consider an object i (which may be a cell or a blood vessel), whose centre is located at $\mathbf{x}_i = (x_i, y_i)$ at time t. We associate with object i a categorical label $q_i \in \{B, M, S, T, N\}$ which indicates whether it is a blood vessel, macrophage, stromal cell, tumour cell or necrotic cell. Given a target label $Q \in \{B, M, S, T, N\}$, we define a binary target function $\Theta(Q, q_i)$ to indicate whether the label q_i associated with object i matches Q:

$$\Theta(Q, q_i) = \begin{cases} 1 & \text{if } q_i = Q, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

Cell counts

We use the binary target function to determine the number of objects of a particular type at time t. For example, $N_T(t)$, the number of tumour cells at time t, is given by

$$N_T(t) = \sum_i \Theta(T, q_i) \tag{9}$$

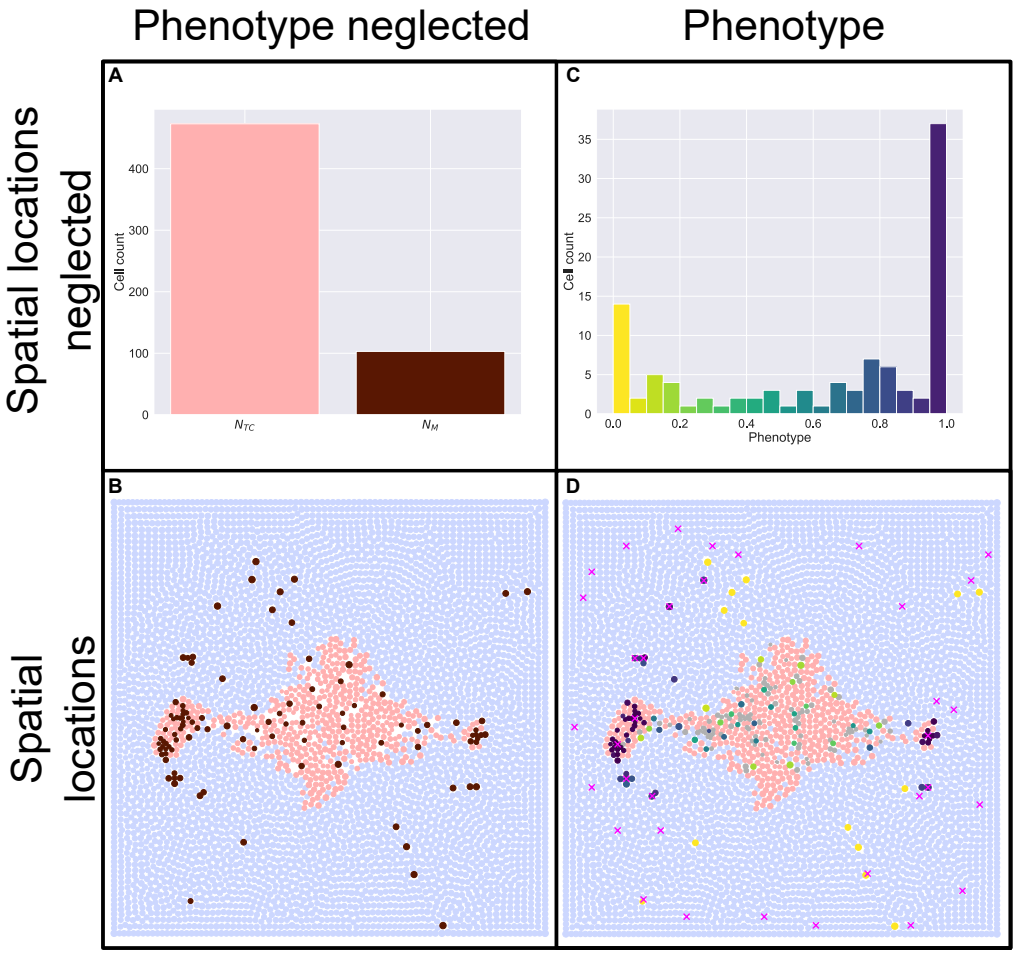

Fig 2. Post-processing ABM output at a fixed time point to emulate data collected from biological experiments
A: Disaggregated cell counts (tumour cells N_T and macrophages N_M).
B: A synthetic IHC slide. Two or three different cell types can be resolved spatially. Since no detailed expression data is available, all macrophages appear identical (brown). Information about some cell types may not be visible (e.g., necrotic cells).
C: Synthetic single-cell sequencing data provides a detailed picture of expression levels across the sample (here represented by the 'phenotype' variable in macrophages), but without spatial resolution.
D: Synthetic IMC provides detailed phenotype information for each macrophage, and spatial locations of all cells.

where the sum is over all objects in the simulation at time t. Similarly, $N_M(t)$, the total number of macrophages at time t, is given by

$$N_M(t) = \sum_i \Theta(M, q_i). \tag{10}$$

Phenotype distributions

Statistics that summarise the phenotypic distribution of macrophages can also be computed. For example, $\bar{p} = \sum_i \Theta(M, q_i) p_i / N_M$ defines the mean macrophage phenotype at time t. If we assume that M_1-like macrophages have phenotype $0 \leq p < 0.5$, then we can define $N_{M_1}(t)$ and $N_{M_2}(t)$, the numbers of M_1-like and M_2-like macrophages, as follows:

$$N_{M_1}(t) = \sum_i \Theta(M, q_i) \Phi(0, 0.5; p_i), \tag{11}$$

$$N_{M_2}(t) = \sum_i \Theta(M, q_i) \Phi(0.5, 1; p_i), \tag{12}$$

where

$$\Phi(P_1, P_2; p_i) = \begin{cases} 1 & \text{if } P_1 \leq p_i \leq P_2, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

and $0 \leq P_1 < P_2 \leq 1$. We also calculate the proportions of macrophages with M_1-like ($\phi_1(t) \equiv N_{M_1}/N_M$) and M_2-like phenotypes: $\phi_1(t) \equiv N_{M_1}/N_M$ and $\phi_2(t) \equiv N_{M_2}/N_M$, respectively.

Cross pair correlation function (cross-PCF)

The cross-PCF identifies spatial correlations between objects with categorical labels that are separated by a distance r. We define a sequence of annuli, of inner radius r_k and outer radius $r_{k+1} = r_k + dr$ where $r_0 = 0$ and $dr > 0$. We denote by $A_{r_k}(\mathbf{x})$ the area of the annulus with inner radius r_k that is centred at the point $\mathbf{x}$. If this annulus lies wholly inside the domain then $A_{r_k}(\mathbf{x}) = \pi((r_k + dr)^2 - r_k^2) = \pi(2r_k + dr)dr$; otherwise, only the area contained within the domain is recorded. We also define the indicator function, $I_k(r)$, as follows:

$$I_k(r) = \begin{cases} 1 & \text{for } r_k \leq r < r_{k+1}, \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

We calculate the cross-PCF for blood vessels and tumour cells by considering a domain of area A. We suppose that, at time t, it contains N_B blood vessels and N_T tumour cells. Then, we define the cross-PCF, $g_{BT}(r)$, by:

$$\begin{aligned} g_{BT}(r) &= \frac{1}{N_B} \sum_{i=1}^{N} \Theta(B, q_i) \left(\sum_{j=1}^{N} \Theta(T, q_j) \frac{I_k(|\mathbf{x}_i - \mathbf{x}_j|)}{A_{r_k}(\mathbf{x}_i)} \Big/ \frac{N_T}{A} \right) \\ \implies g_{BT}(r) &= \frac{1}{N_B N_T} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{A}{A_{r_k}(\mathbf{x}_i)} \Theta(B, q_i) \Theta(T, q_j) I_k(|\mathbf{x}_i - \mathbf{x}_j|). \end{aligned} \tag{15}$$

where $r \in [r_k, r_{k+1})$ and N is the total number of objects in the simulation. For each blood vessel, the cross-PCF compares the density of tumour cells in the annulus that surrounds it to N_T/A, the expected density in the annulus under complete spatial

randomness (CSR). Thus, $g_{BT}(r) > 1$ indicates clustering of tumour cells at distance $r \in [r_k, r_{k+1})$ from blood vessels and $g_{BT}(r) < 1$ indicates anti-correlation, or exclusion, of tumour cells at distance $r \in [r_k, r_{k+1})$ from blood vessels.

Cross-PCFs for other pairs of categorical variables can be defined similarly. In this paper, we focus on g_{TB}, g_{BM} and g_{TM}. We note that the cross-PCF is not necessarily symmetric (i.e., $g_{BT} \neq g_{TB}$ since, for any pair of points, the annulus surrounding one point may intersect with the domain boundary while the annulus surrounding the other may not).

Weighted pair correlation function (wPCF)

The weighted PCF (wPCF) quantifies spatial correlations between objects with discrete categorical labels (here, blood vessels or tumour cells) and those with continuous labels (here, macrophages with phenotype p). Its relationship to the cross-PCF is explored in S3 Appendix: Derivation of PCFs and Cross-PCFs from wPCF.

We calculate the wPCF by replacing $\Theta(Q, q_i)$ in Equation (15) with a weighting function, $0 \leq w_p(P, p_i) \leq 1$, which describes how p_i differs from a target phenotype, P. Multiple functional forms could be used for the weighting function; in S4 Appendix: Comparison of different weighting functions, we show how alternative functional forms affect the wPCF. For simplicity, throughout this paper we use a triangular weighting function of the form:

$$w_p(P, p_i) = \max\left(1 - \frac{|P - p_i|}{\Delta P}, 0\right), \tag{16}$$

and fix $\Delta P = 0.2$. Then, $w_p \approx 1$ for cells whose phenotype p_i is close to the target P and $w_p = 0$ for those with $|P - p_i| > \Delta P$. We note further that $w_p \to \Theta$ as $\Delta P \to 0$.

Replacing $\Theta(T, q_i)$ with $w_p(P, p_i)$ in Eq (15), we define the wPCF for macrophages of target phenotype P and blood vessels B, at lengthscale r, as follows:

$$wPCF(r, P, B) = \frac{1}{W_P N_B} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{A}{A_{r_k}(\mathbf{x}_i)} w_p(P, p_i) \Theta(B, q_j) I_k(|\mathbf{x}_i - \mathbf{x}_j|) \tag{17}$$

where $W_P(P) = \sum_{i=1}^{N} w_p(P, p_i)$ is the total 'weight' associated with the target label P across all macrophages ($W_P(P)$ replaces N_T in Eq (15); non-macrophages have weight $w_p = 0$). Intuitively, the wPCF extends the cross-PCF by weighting the contribution of each macrophage based on how closely its phenotype matches the target phenotype.

In Fig 3 we present two examples showing how the wPCF characterises spatial correlations between objects with a continuous label p (coloured circles, analogous to macrophages with phenotype p) and objects with a categorical label (magenta crosses, analogous to blood vessels). In both examples, 200 crosses are uniformly distributed along the line $y = 1$, and 1000 circles are randomly placed throughout a square domain of length 2. In Fig 3A, the label p_i of a circle at (x_i, y_i) increases linearly with distance from the line $y = 1$ ($p_i = |1 - y_i|$); in Fig 3B, the label increases with the square of this distance ($p_i = |1 - p_i|^2$). The corresponding wPCFs are shown in the middle panels of Fig 3, for a range of target labels P and distances r. By construction, a circle at distance r from the nearest cross has label $P \approx r$ for (A) (or $P \approx r^2$ for (B)). The two wPCFs show strong clustering along these lines and exclusion at shorter distances for points with larger labels (above the dashed lines). We explain the weaker clustering observed below the lines as follows. Consider a cross at position $(x_j, 1)$. In (A), the largest label associated with a circle at distance r from this cross is $p = r$ (if the circle is directly above the cross). Smaller labels can also be recorded, for circles at distance r which are offset from the cross. In the bottom panels of Fig 3, we plot $wPCF(r, P, B)$ for fixed values of the target label P. These curves can be interpreted as cross-PCFs for

points whose labels p_i are "close" to P, and show the strongest clustering at the expected values.

A natural extension to the wPCF considers objects with two continuous labels, P_1 and P_2, say, in order to identify spatial correlations between objects with labels close to P_1 and objects with labels close to P_2 (e.g., colocalisation of macrophages with $p \approx 0$ with those with $p \approx 1$, or colocalisation of a particular macrophage phenotype with a particular concentration of CSF-1). We discuss such extensions in S5 Appendix: wPCF for comparing two continuous labels.

Results

Tumour progression in the presence and absence of macrophages

Fig 4 shows how the presence of macrophages can alter the spatial composition and growth dynamics of a small number of tumour cells initially located at the centre of a two-dimensional square domain that contains stromal cells and a fixed number of static blood vessels. The vessels are randomly distributed in the domain but excluded from a circle of radius R_{B} centred on the initial tumour mass. Panels A and C show that, in the absence of macrophages (i.e., setting the maximum probability of extravasation $P^{\star} = 0$), the tumour increases rapidly in size but remains as a compact mass. At much longer timescales, the tumour evolves to a steady state where the net proliferation rate of oxygen-rich cells on the tumour periphery balances the net death rate of oxygen-starved cells towards the tumour centre (see S2 Appendix: Tumour growth in the absence of macrophages).

When macrophage extravasation is active (Fig 4B and Fig 4D), the ABM reproduces the qualitative behaviours outlined in Fig 1 [23]. At early times ($t \approx 100$), CSF-1 levels are below the threshold for macrophage extravasation and the domain is devoid of macrophages. As the tumour increases in size, more CSF-1 is produced until, eventually, CSF-1 levels at the blood vessels reach the threshold for extravasation of M_1-like macrophages ($t \approx 200$). By $t = 300$, some macrophages have infiltrated the tumour mass. Macrophages that have been exposed to sufficient levels of TGF-β become M_2-like and migrate back towards nearby blood vessels, in response to spatial gradients in CXCL12. The M_2-like macrophages also produce EGF which acts as a chemoattractant for the tumour cells. Thus, at $t = 500$ hours, clusters of M_2-like macrophages and tumour cells surrounding multiple blood vessels are visible in the domain.

Comparison of Fig 4A and Fig 4B, and Fig 4C and Fig 4D, reveals how the presence of macrophages can transform a tumour from a rapidly growing, compact mass to one that is slower growing and more diffuse. While the summary data presented in Fig 4D provide useful information about the tumour's overall growth dynamics and changing cellular composition, detailed information about its morphology and spatial heterogeneity is lacking. In Fig 5 we present additional statistics generated from the spatial data at $t = 500$. Fig 5B shows the cross-PCF $g_{TB}(r)$ (mean and SD of 10 simulations) for the tumour cells and blood vessels in Fig 5A. There is complete exclusion between blood vessels and tumour cells up to a radius of approximately 5 cell diameters, a lengthscale characterising the minimum distance between the blood vessels and the tumour mass. By comparison, the cross-PCF $g_{TB}(r)$ shown in Fig 5D quantifies the short-range clustering of tumour cells and blood vessels in Fig 5C. The peak around $r = 1$ indicates close proximity between some blood vessels and tumour cells. The cross-PCFs provide additional information about the strong short-range clustering of macrophages with tumour cells (Fig 5E) and blood vessels (Fig 5F). Macrophages are strongly correlated with tumour cells, particularly at short lengthscales indicating prominent colocalisation (Fig 5E). There is also strong short-range colocalisation

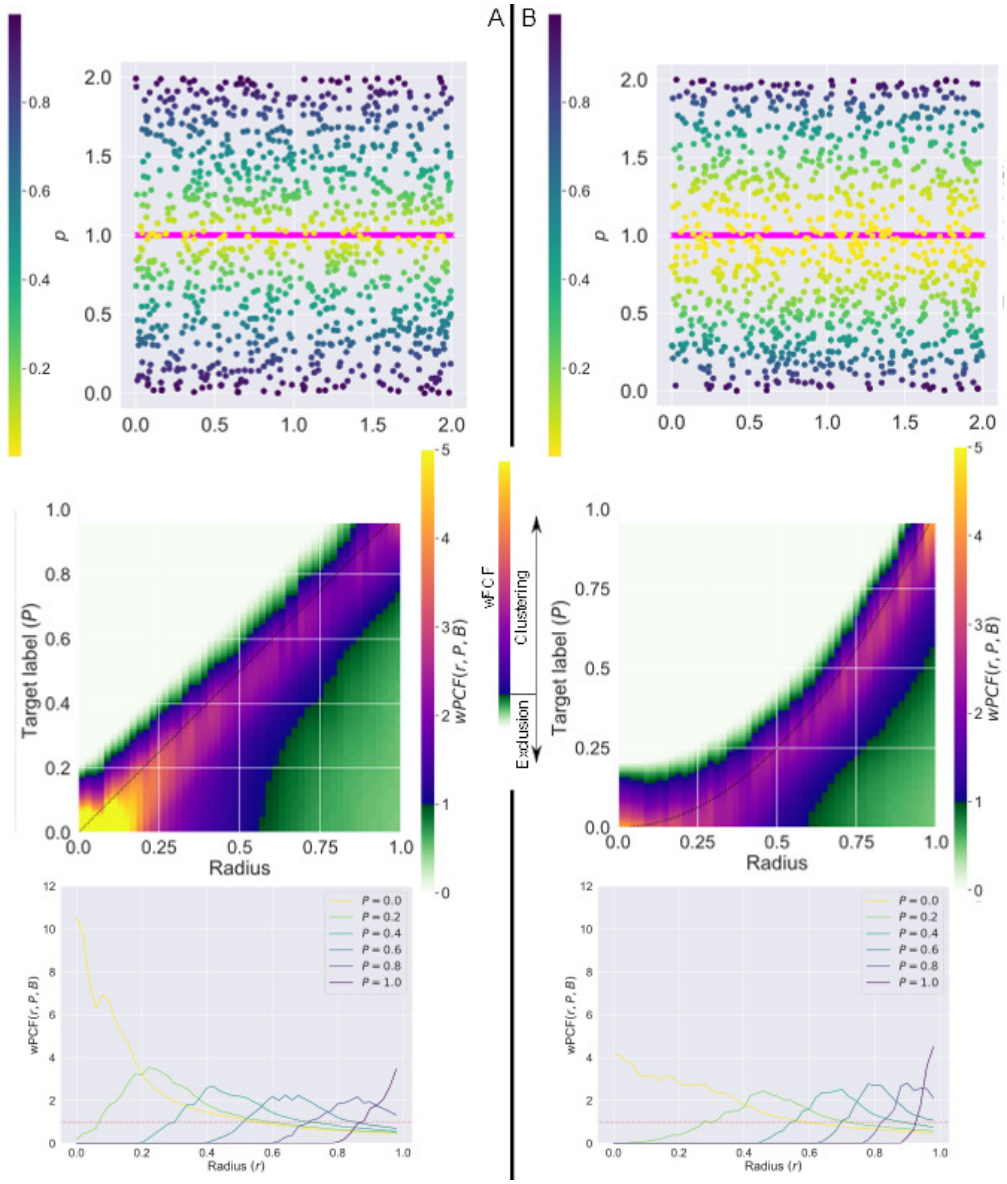

Fig 3. Examples illustrating the wPCF.
Two examples showing how the wPCF can identify spatial correlations between categorical objects (200 pink crosses equally spaced on the line $y = 1$) and objects with real values (1000 randomly placed circles with labels $p \in [0, 1]$).
A: Points are labelled according to the formula $p_i = |1 - y_i|$. B: Points are labelled according to the formula $p_i = |1 - y_i|^2$. Top: Point patterns consisting of equally spaced pink crosses and randomly placed circles with non-random labelling. Middle: wPCFs corresponding to the above point patterns. Dashed black lines show the lines $P = r$ and $P = r^2$, which by construction should show the strongest correlation. Bottom: Horizontal slices through the wPCF at fixed values of P. Such slices can be interpreted as a cross-PCF showing colocalisation between the pink crosses and circles with labels close to P.

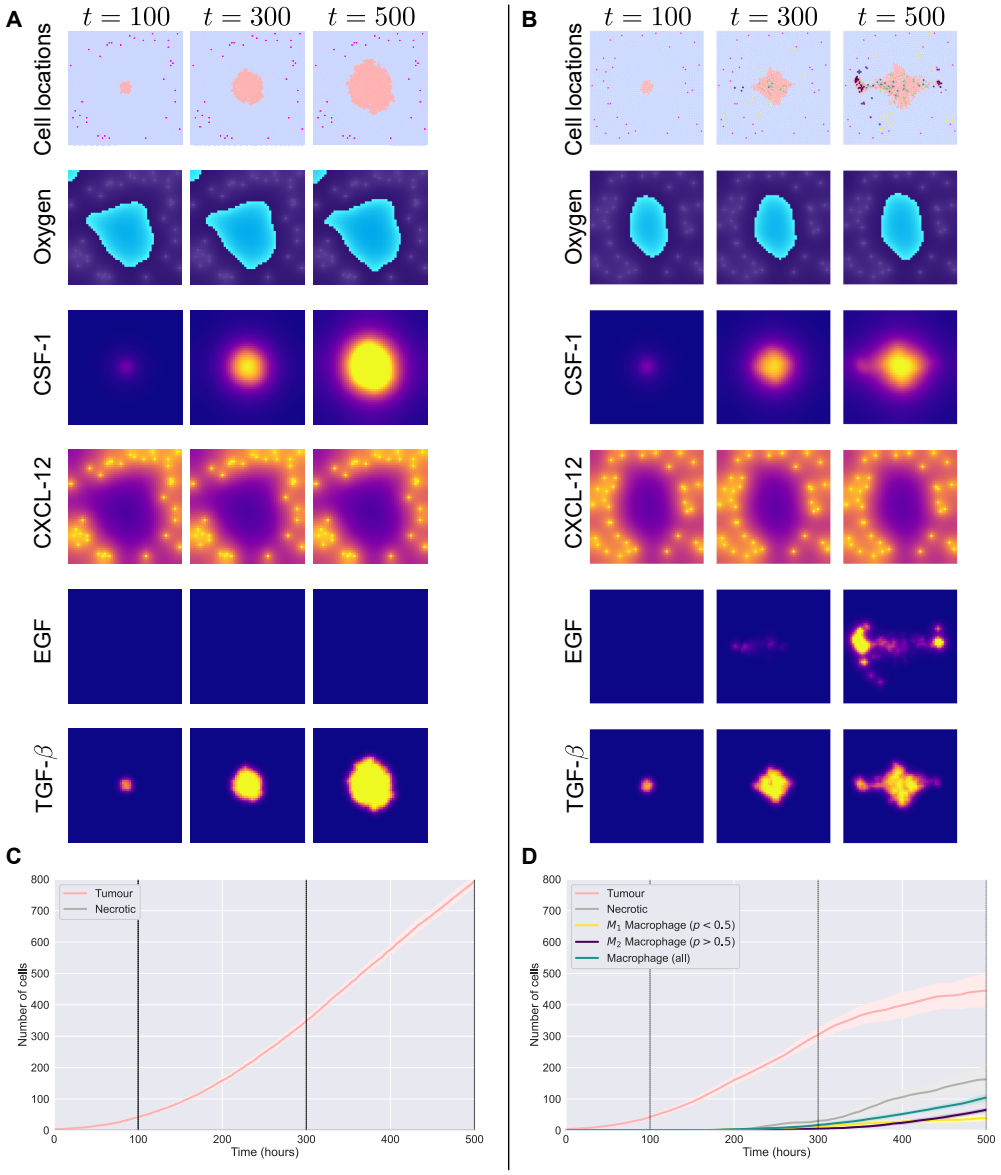

Fig 4. Representative output from model simulations
A, B: Spatial distributions of cells, oxygen, CSF-1, CXCL12, EGF and TGF-β at times $t = 100, 300, 500$ from simulations which neglect (A, $P^\star = 0$) or include (B, $P^\star = 0.075$) macrophage extravasation. Comparison of these plots shows how the tumour's growth rate and spatial composition can change in the presence of macrophages. Parameter values: $\chi_c^m = 2$,$\chi_\xi^m = 3$, $\chi_\epsilon^T = 3$, $c_{1/2} = 0.9$, $g_{\text{crit}} = 0.05$; all other parameters fixed at their default values (see Table 3 in S1 Appendix: Model Description).
C, D: Change in numbers of tumour cells, necrotic cells, M_1-like and M_2-like macrophages, and total number of macrophages over time for the simulations presented in (A) and (B) (mean and SD from 10 realisations).

between macrophages and blood vessels (Fig 5F), suggesting the presence of perivascular niches containing macrophages, tumour cells and blood vessels.

The corresponding weighted PCFs, $wPCF(r, P, B)$ and $wPCF(r, P, T)$ (mean of 10 iterations), highlight differences in the spatial colocalisation of macrophages with different phenotypes (Fig 5G and Fig 5H). In particular, in this simulation macrophages with $p \approx 0$ do not correlate strongly with either tumour cells or blood vessels. By contrast, macrophages with intermediate phenotype ($0.2 < p < 0.7$ are strongly colocalised with tumour cells, and excluded from blood vessels. Finally, M_2-skewed macrophages (with $p > 0.7$) are strongly clustered with blood vessels and tumour cells at short length scales. Taken together, the wPCFs capture the following trends: M_1-macrophages dispersed throughout the domain and changing their phenotype when they encounter a mass of tumour cells; and, M_2-like macrophages localised around blood vessels, where they form clusters with tumour cells of approximately 2-3 cell diameters in size.

Varying macrophage sensitivity to environmental cues generates diverse patterns of tumour growth

The simulation results presented in Fig 4 and Fig 5 show that macrophages can have a marked effect on tumour morphology and growth. We conduct a parameter sweep in which we vary two parameters that influence macrophage sensitivity to two environmental cues: $c_{1/2}$, the concentration of CSF-1 at which macrophage extravasation is half-maximal, and χ_c^m, the chemotactic sensitivity of macrophages to spatial gradients of CSF-1. All other model parameters are fixed at the default values stated in Table 3 in S1 Appendix: Model Description.

We consider $0.5 \leq \chi_c^m \leq 4.5$ and $0.1 \leq c_{1/2} \leq 0.9$. For each pair of values of $\chi_c^m \in \{0.5, 1.5, 2.5, 3.5, 4.5\}$ and $c_{1/2} \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$, we simulate the ABM and output the spatial composition of the tumour at $t = 500$ (10 realisations per parameter set).

Diverse patterns of tumour growth

The results presented in Fig 6 illustrate the range of qualitative behaviours that the ABM generates as χ_c^m and $c_{1/2}$ vary. We summarise these behaviours as follows:

- A compact tumour mass, with macrophages confined to the surrounding stroma. The dominant macrophage phenotype is M_1 (COMP; yellow box in Fig 6).
- Total, or near-total, tumour cell elimination. Some macrophages may be clustered around the last surviving tumour cells and the dominant macrophage phenotype is M_1 (ELIM; blue box in Fig 6).
- The tumour is asymmetric, with a diffuse, fragmented structure. Perivascular niches containing M_2-like macrophages and tumour cells surround blood vessels. The bulk of the tumour is infiltrated with M_1-like and transitioning macrophages, with central tumour necrosis caused by macrophages killing tumour cells (DIFF; red box in Fig 6).

Compact tumour growth (COMP, yellow box) arises for low values of χ_c^m (e.g, $\chi_c^m = 0.5$). Large numbers of macrophages extravasate in response to CSF-1 but, since they are not strongly attracted to the tumour mass, they remain in the stroma. As a result, tumour growth is unaffected by the macrophages, which are predominantly M_1-like. When $c_{1/2}$ is also small (e.g., $c_{1/2} \lessapprox 0.3$), the rate of macrophage extravasation is high, and some macrophages reach the tumour boundary through random exploration

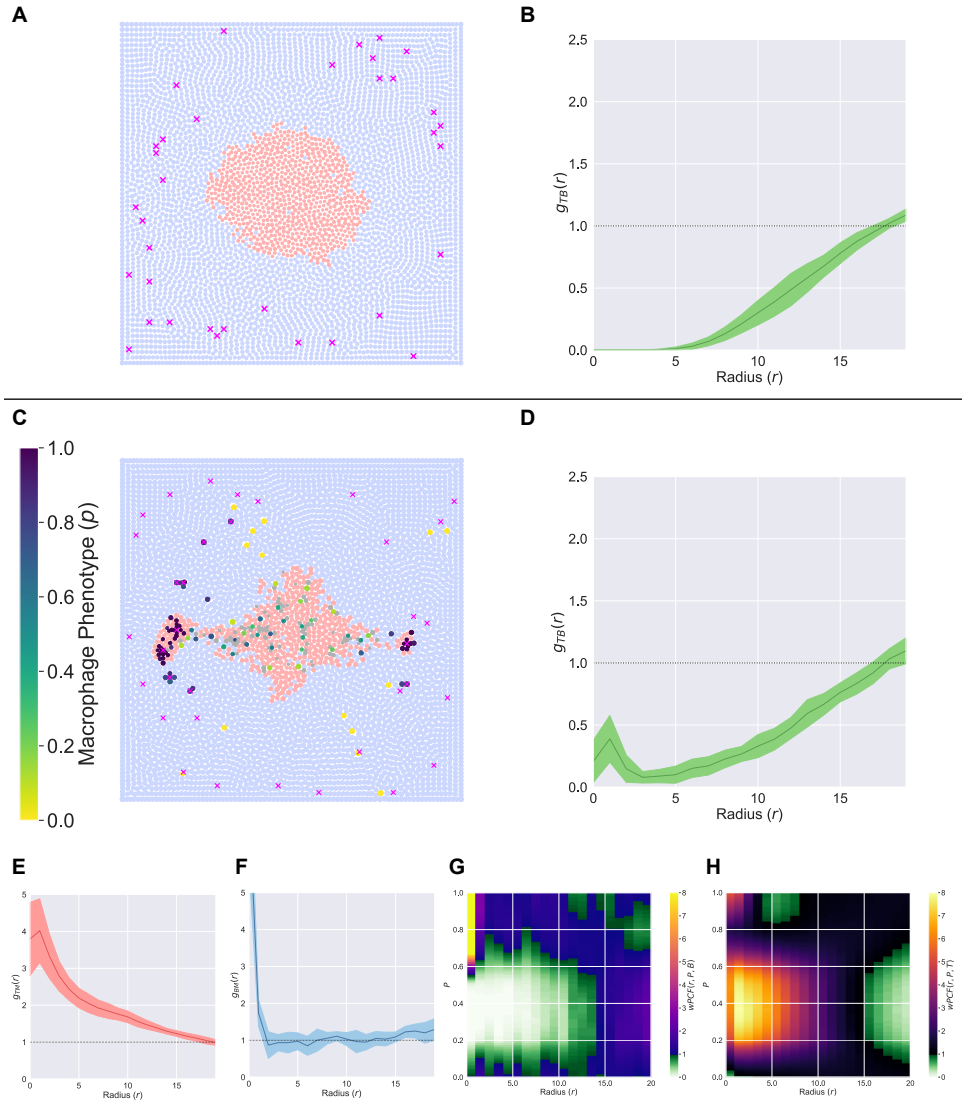

Fig 5. Statistical analysis of the simulation endpoints in Fig 4
A: $t = 500$ for tumour growth without macrophage extravasation.
B: The cross-PCF $g_{TB}(r)$ for tumour cells and blood vessels in (A). No tumour cells are observed within a distance of 5 cell diameters from a blood vessel.
C: $t = 500$ for tumour growth with macrophage extravasation.
D: The cross-PCF $g_{TB}(r)$ for the tumour cells and blood vessels in (C). The cross-PCF reveals short range interactions between tumour cells and blood vessels. Comparison with (B) quantifies how the spatial distribution of tumour cells relative to the blood vessels changes in the presence of macrophages.
E: The cross-PCF $g_{TM}(r)$ associated with (C). The cross-PCF reveals strong short range interactions between macrophages and tumour cells.
F: The cross-PCF $g_{BM}(r)$ associated with (C). Short-range correlations between blood vessels and macrophages are very strong and decay rapidly with distance r.
G: The weighted PCF $wPCF(r, P, B)$ associated with (C). There is strong, short-range colocalisation of macrophages with $p > 0.6$ and blood vessels, while macrophages with $0.2 \lessapprox p \lessapprox 0.6$ are excluded from regions of radius approximately 10-15 cell diameters surrounding blood vessels.
H: The weighted PCF $wPCF(r, P, T)$ associated with (C). Macrophages with $p > 0.6$ are strongly colocalised with tumour cells at distances $0 \leq r \leq 10$, indicating their presence inside the tumour mass. Short-range colocalisation ($r \approx 3$) is also observed for M_2-like macrophages with $p \geq 0.9$.

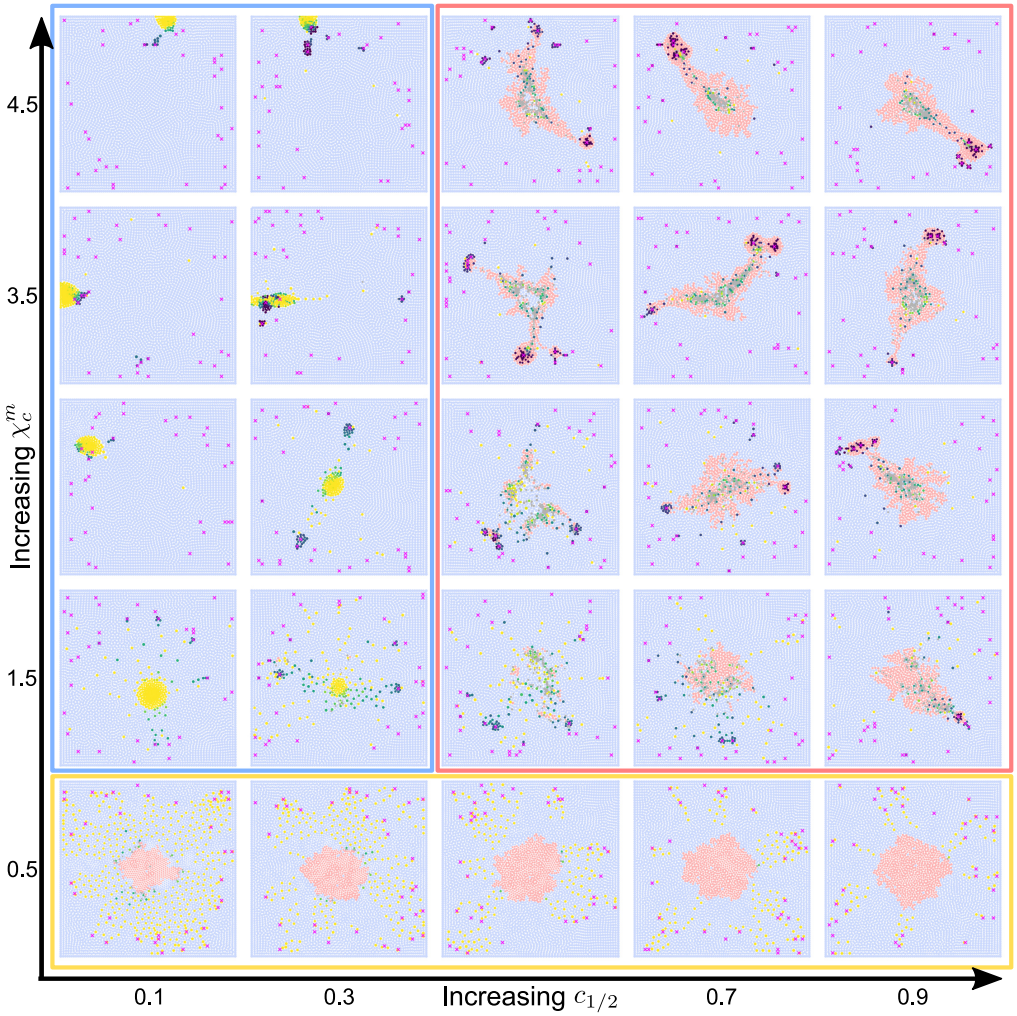

Fig 6. Varying macrophage sensitivity to environmental cues generates diverse patterns of tumour growth

Representative simulation endpoints for combinations of $\chi_c^m = 0.5, 1.5, 2.5, 3.5, 4.5$ and $c_{1/2} = 0.1, 0.3, 0.5, 0.7, 0.9$. We group these into three qualitatively similar behaviours:
COMP - yellow box: compact tumour mass, with predominantly M_1-like macrophages confined to the stroma.
ELIM - blue box: total or near total tumour elimination.
DIFF - red box: establishment of perivascular niches containing M_2-like macrophages, tumour cells and blood vessels. Tumour masses are asymmetrical.

of the tissue. These macrophages kill tumour cells on contact, causing the tumour mass to decrease in size.

Tumour elimination (ELIM, blue box) occurs when the rate of macrophage extravasation is very high (low values of $c_{1/2}$). In contrast to COMP simulations, tumour elimination occurs because the M_1-like macrophages are strongly attracted to the tumour mass and kill tumour cells before they are 'reprogrammed' to an M_2-like phenotype. Strong chemotactic sensitivity to CSF-1 (large values of χ_c^m) causes the macrophages to cluster around the last tumour cells to be eliminated.

Diffuse tumour growth (DIFF, red box) is intermediate between the COMP and DIFF growth regimes. Macrophages are recruited to the tumour slowly enough that they do not overwhelm it. Their exposure to TGF-β is sufficient to 'reprogram' them to an M_2-like phenotype. The M_2-like macrophages migrate towards nearby blood vessels in response to spatial gradients in CXCL12, with the CSF-1/EGF paracrine loop causing tumour cells to trail behind them. DIFF simulations replicate the behaviours reported in Fig 4 and Fig 5.

We now analyse the simulation endpoints in Fig 6 using the statistics described in the methods section. In this way, we aim to identify metrics which distinguish COMP, ELIM and DIFF simulations.

Cell counts analysis

Fig 7 shows how the numbers of tumour cells (N_T) and macrophages (N_M) at $t = 500$ change as χ_c^m and $c_{1/2}$ vary. While the tumour cell counts show clearly the region of parameter space within which tumour elimination occurs (blue box), they are unable to distinguish between COMP and DIFF outcomes. For example, when $c_{1/2} = 0.9$, N_T and N_M are similar for $\chi_c^m = 0.5$ and $\chi_c^m = 4.5$ (Fig 7 bottom right, top right), although the tumour is a compact mass when $\chi_c^m = 0.5$ and diffuse with substantial numbers of perivascular tumour cells when $\chi_c^m = 4.5$. Additionally, standard cell counts cannot distinguish differences in macrophage phenotype or tumour structure.

Single cell sequencing analysis

In Fig 8 we present the macrophage phenotype distributions corresponding to the simulations in Fig 6. Macrophages in COMP tumours (yellow box) are M_1-skewed, while those in DIFF (red box) are M_2-skewed ($p > 0.5$). Since these histograms neglect information about spatial location, we cannot determine whether M_2-skewed macrophages are perivascular.

By combining the tumour cell counts from Fig 7 with the macrophage phenotype distributions from Fig 8, it is possible to detect differences between COMP and ELIM simulations. Although both cases are dominated by M_1-like macrophages (with $p = 0$), ELIM simulations have increased numbers of M_2-like macrophages ($p > 0.5$). Since some macrophages in ELIM simulations have been exposed to high enough levels of TGF-β for a sufficient period to change their phenotype, we deduce that they must have infiltrated the tumour mass. In contrast, since no macrophages have $p > 0$ for COMP simulations, we deduce that none have been exposed to high levels of TGF-β and, hence, that none have infiltrated the tumour.

Immunohistochemistry analysis

We now focus on perivascular niches, using the cross-PCF to describe the relationships between tumour cells and blood vessels ($g_{TB}(r)$) and macrophages and blood vessels ($g_{BM}(r)$). Relationships between tumour cells and macrophages are included in S6 Appendix: cross-PCFs of tumour cells to macrophages.

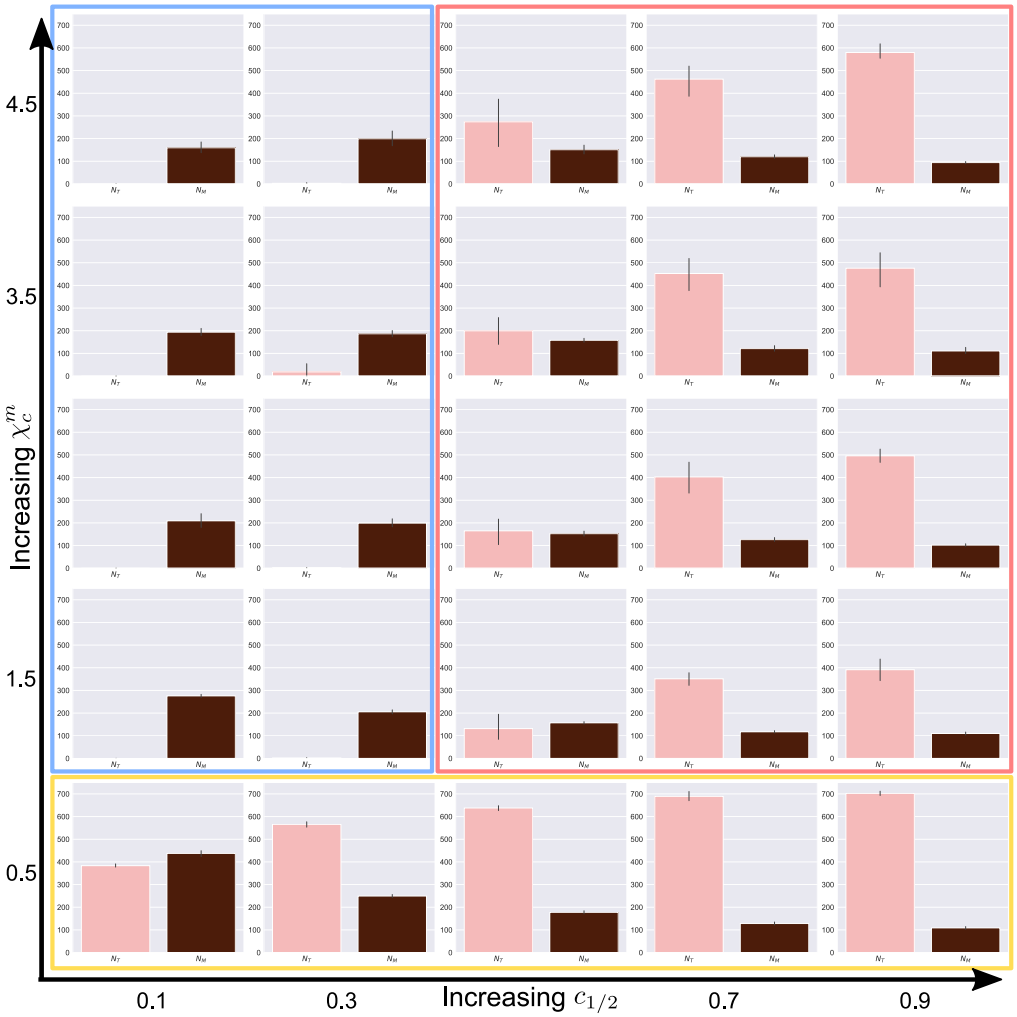

Fig 7. Cell counts for simulations in Fig 6

Mean cell counts (with 95% confidence interval) of tumour cells (N_T) and macrophages (N_M) corresponding to simulations in Fig 6.

ELIM simulations can be distinguished from other cases (blue box) due to differences in N_T. However, cell counts alone cannot easily distinguish between COMP (yellow box) and DIFF (red box) simulations, despite crucial differences in simulation behaviours.

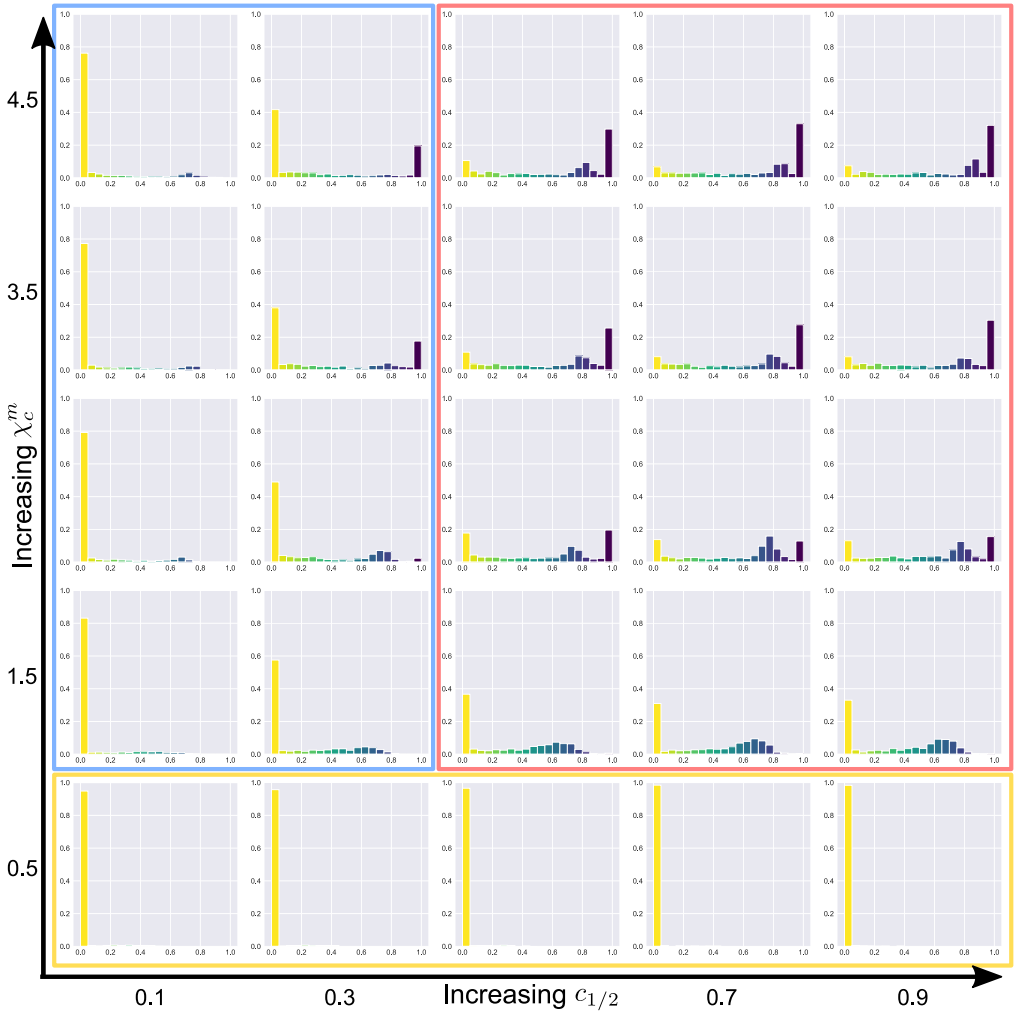

Fig 8. Distributions of macrophage phenotypes
Macrophage phenotype distributions corresponding to the simulations in Fig 6 (averaged over 10 repetitions). Macrophages in COMP simulations (yellow box) overwhelmingly display a phenotype of $p = 0$. ELIM simulations (blue box) contain a majority of M_1 macrophages, although there are some M_2-like macrophages present in each simulation. DIFF simulations (red box) contain macrophage populations which are broadly M_2-like or split between M_1 and M_2 macrophages, with a lower proportion of M_1-like macrophages than ELIM and DIFF simulations.

Fig 9 shows $g_{TB}(r)$ for the simulation endpoints from Fig 6 (mean and SD; cross-PCFs for individual simulations can be seen in the Zegami collection accessible via the github repository associated with this paper). Since $g_{TB}(r)$ is not defined when there are no tumour cells, we leave a blank space for ELIM simulations (blue box in Fig 9). For many DIFF simulations (red box) the cross-PCF shows clustering of tumour cells close to blood vessels ($g_{TB}(r) > 0$ for $r \lessapprox 2$). Non-zero values of $g_{TB}(r)$ indicate that some tumour cells are present at distance r from a blood vessel, and the height of the cross-PCF shows the strength of their colocalisation relative to the overall distribution of tumour cells. The local maximum at $r \approx 1$ indicates stronger colocalisation at one to two cell diameters than at other distances from blood vessels.

The COMP simulations (yellow box in Fig 9), for which no tumour cells reach the blood vessels, show complete exclusion of tumour cells at small distances from blood vessels, with $g_{TB}(r) = 0$ for $0 < r \lessapprox 5 - 10$. The first r value for which $g_{TB}(r) > 0$ indicates the minimum distance between the blood vessels and the tumour boundary. In the COMP simulations, this distance is maximised when $c_{1/2} = 0.1$ (and the tumour mass is smallest) and minimised when $c_{1/2} = 0.9$, which is consistent with the trends for N_T depicted in Fig 7.

DIFF simulations (red box in Fig 9) show stronger colocalisation between blood vessels and tumour cells for larger values of χ_c^m, suggesting that recruitment of anti-tumour macrophages to the tumour mass causes increased movement of tumour cells to the vasculature in some regimes. We note that an increase in $g_{TB}(r)$ at short distances could also be caused by a reduction in the number of tumour cells in the main tumour mass, while the number of perivascular tumour cells remains constant. However, if we combine the cross-PCF with the tumour cell count N_T, then we may infer increased colocalisation between blood vessels and tumour cells as χ_c^m increases.

Fig 10 shows the cross-PCFs describing macrophage proximity to blood vessels, $g_{BM}(r)$, when differences in phenotype are neglected. In almost all simulations there is short-range ($r \approx 0$) correlation between macrophages and blood vessels, although this is most pronounced for DIFF simulations. Since g_{BM} does not distinguish phenotype, we are unable to determine whether macrophages surrounding blood vessels are M_1-like and have recently extravasated (as expected in COMP simulations), or M_2-like and have migrated to the blood vessels from the tumour (as expected in DIFF simulations).

Imaging mass cytometry analysis

In Fig 11 and Fig 12 we present wPCFs showing how the spatial location of macrophages relative to blood vessels ($wPCF(r, p, B)$) and tumour cells ($wPCF(r, p, T)$) varies with macrophage phenotype, p. ELIM, DIFF and COMP simulations have distinct 'fingerprints', with DIFF simulations having wPCFs similar to that of the example in Fig 5. For DIFF simulations (red boxes in Fig 11 and Fig 12), we observe short range clustering ($r \approx 0$) between $p = 1$ macrophages and both blood vessels and tumour cells. On longer lengthscales ($r \lessapprox 10 - 15$) we see exclusion between transitioning macrophages and blood vessels, and clustering of transitioning macrophages and tumour cells. Taken together, these results indicate the formation of perivascular niches containing tumour cells, blood vessels and M_2-macrophages, as reported by Arwert *et al.* [23].

COMP simulations (red box in Fig 11 and Fig 12) exhibit a different signature, with short-range clustering of M_1-like macrophages and blood vessels, and exclusion between blood vessels and other macrophage phenotypes. There is also exclusion between M_1-like macrophages and tumour cells. The wPCFs reveal a stark difference in the spatial localisation of macrophages with $p = 0$ and those with $p > 0$, with transitioning macrophages displaying no exclusion from tumour cells and exhibiting short-range exclusion from blood vessels. These results describe situations in which macrophages

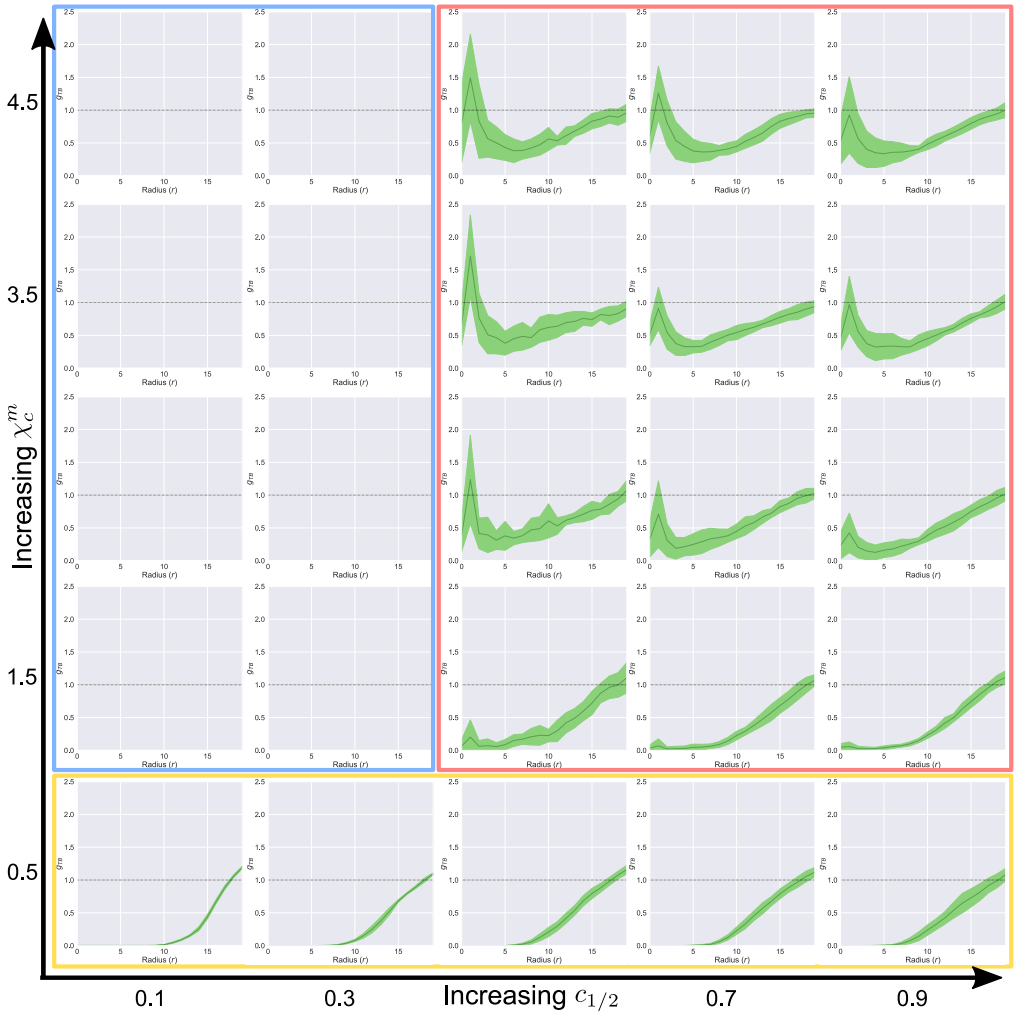

Fig 9. $g_{TB}(r)$ shows the presence of perivascular tumour cells
Mean and SD of $g_{TB}(r)$ for simulations in Fig 6. The presence of perivascular tumour cells can be inferred from $g_{TB}(r) > 0$ for small distances r. Variations in simulation behaviour within the DIFF and COMP groups can also be observed: for instance, DIFF tumours (red box) show stronger colocalisation between blood vessels and tumour cells for larger χ_c^m, and tumour cells in COMP simulations (yellow box) are excluded from a larger region around blood vessels as $c_{1/2}$ decreases.

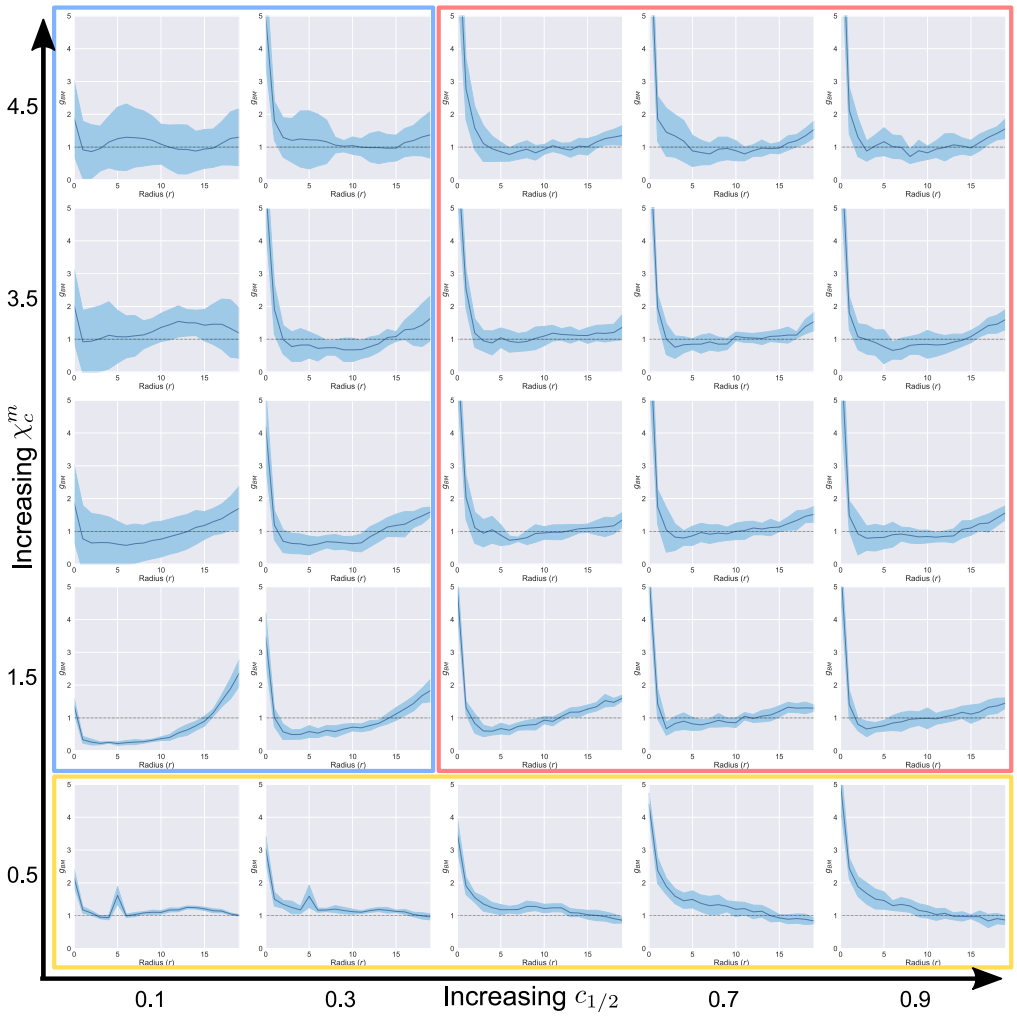

Fig 10. $g_{BM}(r)$ cannot distinguish mechanisms causing blood vessel/macrophage colocalisation Almost all simulations have strong short range colocalisation between macrophages and blood vessels. DIFF simulations (red box) have strong perivascular clusters of M_2 macrophages, while macrophages around blood vessels in COMP simulations (yellow box) are recently-extravasated M_1-like macrophages. The resulting cross-PCFs $g_{BM}(r)$ show similar short-range colocalisation in each case.

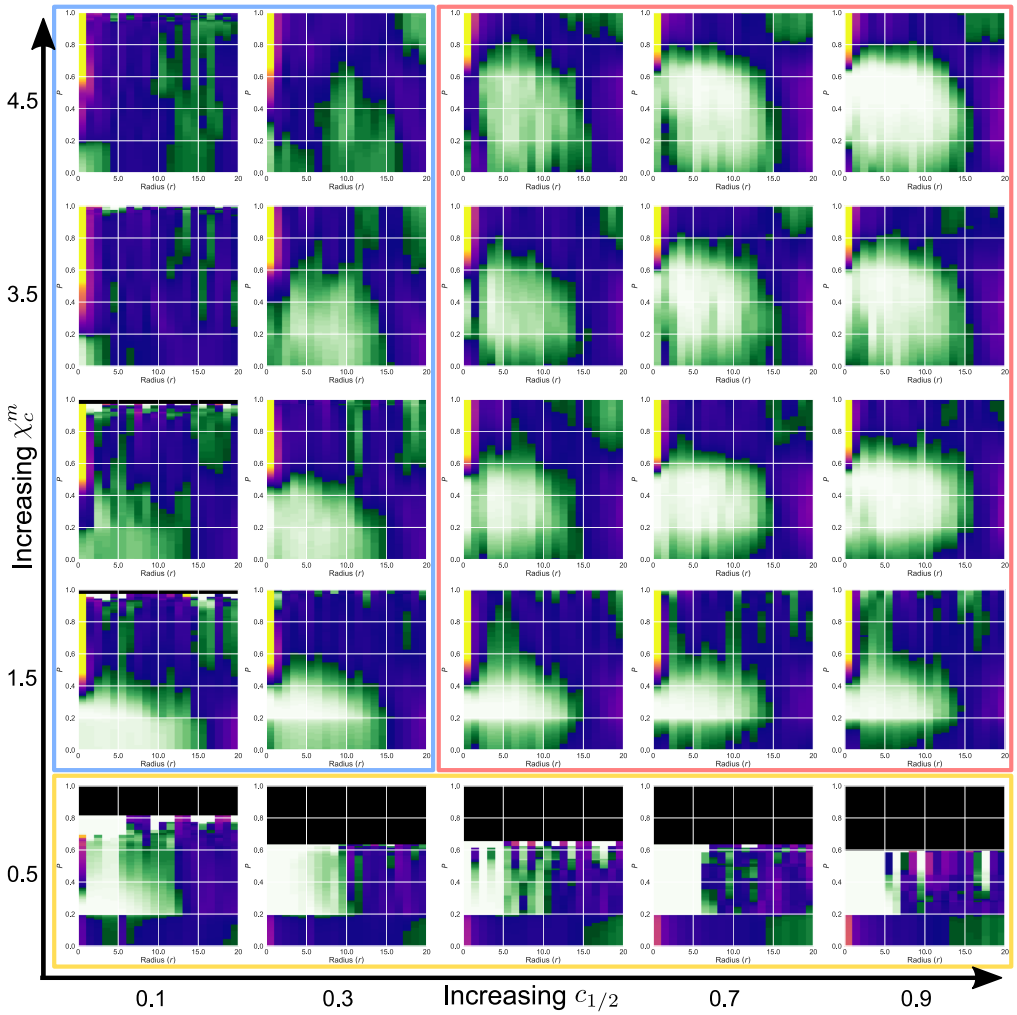

Fig 11. $wPCF(r, p, B)$ shows macrophage localisation by phenotype
The wPCF describing macrophage correlation with blood vessels has a distinct fingerprint for ELIM, DIFF and COMP simulations. COMP simulations (yellow box) show short-range correlation of macrophages with $p = 0$ and blood vessels, while macrophages with higher phenotypes are excluded from blood vessels. DIFF simulations have strong colocalisation between M_2-macrophages and blood vessels, while transitioning macrophages are completely excluded from the proximity of vasculature. ELIM simulations have a similar $wPCF(r, p, B)$ to DIFF simulations, although those with low $c_{1/2}$ and high χ_c^m exhibit weaker exclusion and correlation.

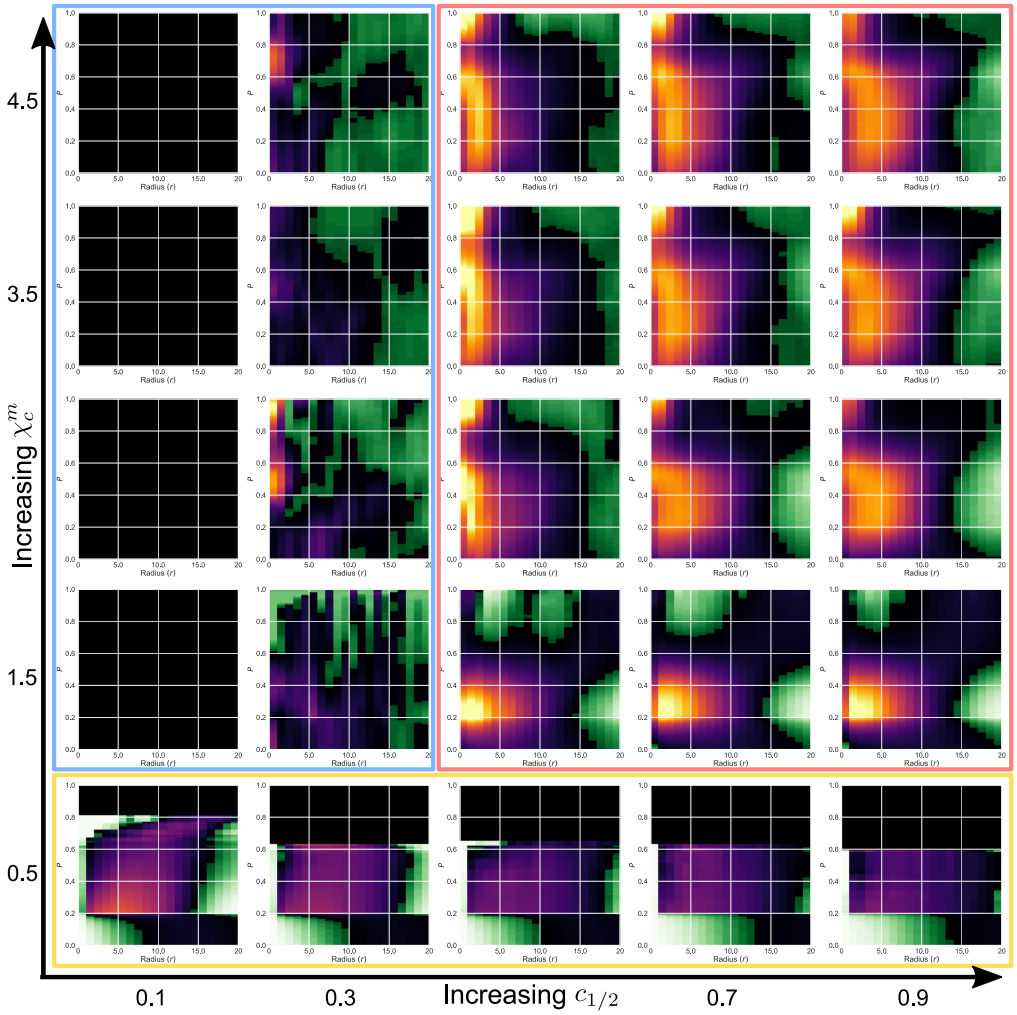

Fig 12. $wPCF(r,p,T)$ shows macrophage localisation by phenotype ELIM, DIFF and COMP simulations have different $wPCF(r,p,T)$ signatures. In particular, when well-defined, ELIM wPCFs do not exhibit strong clustering or exclusion, although some noisy short-range clustering between transitioning or M_2 macrophages and tumour cells can be observed. COMP simulations exhibit strong exclusion on length scales up to $r \approx 10$ between M_1-macrophages and tumour cells, while transitioning macrophages are more likely to be located close to tumour cells. DIFF simulations show spatial colocalisation between tumour cells and transitioning macrophages at length scales up to $r \approx 10$, and between tumour cells and macrophages with $p = 1$ on short length scales ($r \lessapprox 3$). This suggests that transitioning macrophages are in the tumour bulk (which has a radius of approximately 10 cell diameters) while fully M_2-like macrophages are colocalising with smaller groups of tumour cells (2-3 cell diameters wide) elsewhere - in this case, around blood vessels.

with no TGF-β exposure ($p = 0$) are restricted to the stroma while those with slightly increased phenotype cluster around the tumour mass, at distance from the blood vessels.

Finally, for ELIM simulations (blue box in Fig 11 and Fig 12), the wPCFs are generally flatter, with fewer extreme values, indicating weaker correlations. As for the DIFF simulations, M_2-like macrophages are strongly associated with blood vessels, suggesting that these macrophages have followed the same dynamic trajectory as those in DIFF simulations, but that the tumour was eliminated before the selected timepoint.

Wider exploration of parameter space

We have characterised a range of qualitative behaviours, obtained by systematically varying parameters associated with macrophage sensitivity to CSF-1 (namely $c_{1/2}$ and χ_c^m). In this section, we explore a larger, 6-dimensional parameter space that regulates macrophage sensitivity to environmental cues. In addition to χ_c^m, we vary the chemotactic sensitivity parameters χ_ξ^m (macrophage sensitivity to CXCL-12) and χ_ϵ^T (tumour cell sensitivity to EGF). In addition to $c_{1/2}$, we vary $P^\star$ (maximum extravasation probability for saturated CSF-1) and g_{crit} (the critical TGF-β threshold above which macrophage phenotype increases). For each simulation, we sample each parameter uniformly at random from a range covering extreme values of each parameter ($n = 432$ simulations; see Table of ABM parameters in S1 Appendix: Model Description for parameter ranges).

wPCFs characterise diverse model behaviours

In Fig 13 we illustrate some of the ABM outputs generated from the 6D parameter sweep, together with their wPCFs. We highlight how these behaviours differ from those generated by varying only χ_c^m and $c_{1/2}$, and explain how the wPCFs identify subtle changes in the distributions of macrophages and tumour cells.

Panel	χ_c^m	χ_ξ^m	χ_ϵ^T	$c_{1/2}$	$P^\star$	g_{crit}
A	3.494	0.120	2.421	0.409	0.088	0.032
B	2.731	0.190	4.784	0.458	0.011	0.919
C	2.613	0.328	3.442	0.470	0.056	0.955
D	3.448	0.871	3.672	0.497	0.044	0.737
E	2.942	4.074	3.404	0.356	0.011	0.517
F	1.750	3.455	1.892	0.324	0.091	0.032

Table 1. Parameters used for simulations in Fig 13 Values of the six parameters randomly selected for each simulation are shown here; other parameters are fixed at the default values given in S1 Appendix: Model Description.

Fig 13A shows a simulation in which the tumour has been eliminated, but the macrophage distribution differs from those in Fig 6. In those ELIM simulations, the macrophages were M_1-skewed and densely clustered. In Fig 13A, the macrophages exhibit a range of phenotypes, including many with $p = 1$, and they are spread across the centre of the domain, away from the blood vessels. This behaviour is caused by high maximal extravasation rates ($P^\star = 0.088$), high macrophage sensitivity to TGF-β ($g_{\text{crit}} = 0.032$) and weak chemosensitivity to CXCL12 ($\chi_\xi^m = 0.120$). The $wPCF(r, P, B)$ in Fig 13A shows strong exclusion of all macrophage phenotypes with $p > 0$ from blood vessels and contrasts the wPCFs associated with ELIM simulations in Fig 11, where strong clustering between M_2-like macrophage and blood vessels is observed.

In Fig 13B, macrophages of all phenotypes colocalise in the domain centre and, therefore, $wPCF(r, P, B)$ is similar to that in Fig 13A. In contrast to Fig 13A, the low extravasation rate ($P^\star = 0.011$) prevents the macrophages from eliminating the tumour.

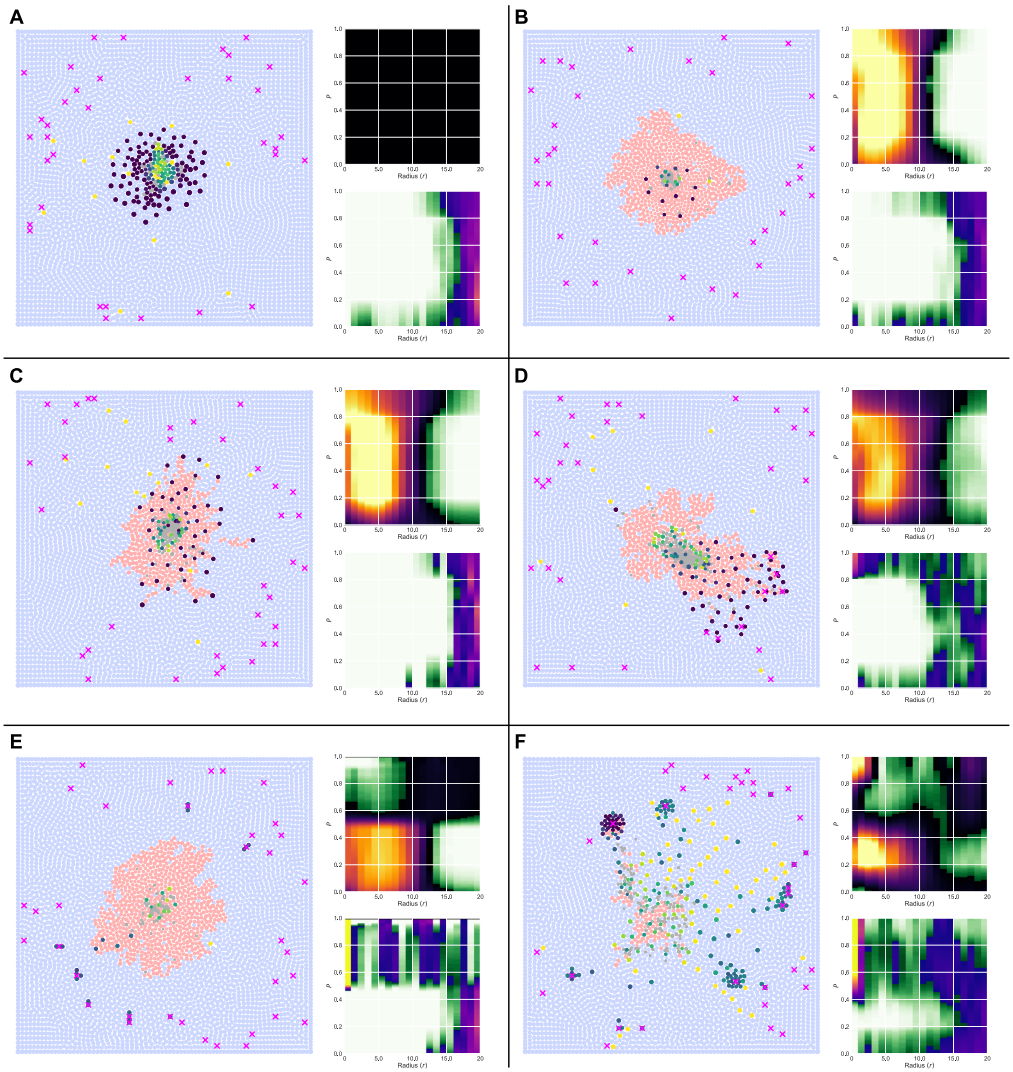

Fig 13. wPCFs distinguish diverse model behaviours
The parameter values used to generate the simulations shown in panels A-F are shown in Table 1. The model generates a wide range of simulation behaviours with distinct wPCF signatures. Panels show cell locations at $t = 500$, with insets of $wPCF(r, P, T)$ (top) and $wPCF(r, P, B)$ (bottom).
A: Tumour elimination in which M_2-like macrophages do not localised around blood vessels.
B: Compact tumour growth in which M_2-like macrophages remain localised within the tumour mass.
C: Diffuse tumour growth in which M_2-like macrophages direct the migration of streams of tumour cells away from the tumour mass.
D: Diffuse tumour growth in which a large cluster of M_2-like macrophages directs tumour cells to localise around blood vessels.
E: M_2-like macrophages migrate towards blood vessels with a low phenotype ($p \approx 0.5$) and hence do not recruit tumour cells.
F: M_2-like macrophages with phenotypes in the range $0.5 < p \leq 1$ localise around blood vessels, and are accompanied by tumour cells only if the perivascular macrophages have $p \approx 1$.

The low value of χ_ξ^m means that these macrophages are unable to migrate towards the blood vessels and remain localised within the tumour long enough to adopt an M_2-like phenotype. In contrast to the wPCFs for the compact tumour growth simulations shown in Fig 12, the wPCFs in Fig 13B show strong colocalisation between all macrophage phenotypes and tumour cells, and exclusion between all macrophage phenotypes and blood vessels. Thus, the wPCFs distinguish behavioural differences between this simulation, that shown in Fig 13A, and the COMP simulations in Fig 6.

Both Fig 13C and Fig 13D show simulations with diffuse tumour growth. In Fig 13C, individual M_2-like macrophages guide thin trails of tumour cells away from the tumour mass. However, since the chemotactic force due to CXCL12 is weak ($\chi_\xi^m = 0.328$), no macrophages have reached the blood vessels at $t = 500$. Hence, while $wPCF(r, P, T)$ is similar to those for the DIFF simulations in Fig 12, $wPCF(r, P, B)$ shows exclusion between all macrophage phenotypes and vasculature. In contrast, Fig 13D shows a diffuse pattern of tumour growth, with a large number of tumour cells guided by M_2-like macrophages to cluster around the blood vessels. This difference is evident from the wPCFs: Fig 13C and Fig 13D have similar $wPCF(r, P, T)$, but $wPCF(r, P, B)$ shows colocalisation between M_2-like macrophages and blood vessels in Fig 13D.

The simulations presented in Fig 13E and Fig 13F are characterised by strong chemotactic sensitivity of M_2-like macrophages to CXCL12 ($\chi_\xi^m = 4.074$ and $\chi_\xi^m = 3.455$ respectively). Both simulations contain perivascular macrophages with phenotypes in the range $0.5 \leq p \leq 1$. This similarity is evident from the $wPCF(r, P, B)$ plots, which show strong colocalisation between vessels and macrophages with $p > 0.5$. In Fig 13E the perivascular macrophages are strongly attracted to CXCL12 ($\chi_\xi^m = 4.074$) and migrate so rapidly that the tumour cells are unable to follow them. As a result, $wPCF(r, P, T)$ shows two distinct behaviours: for $p < 0.5$ macrophages are strongly colocalised with tumour cells, with exclusion at long lengthscales (suggesting that M_1-like macrophages are primarily in the tumour mass). For $p > 0.5$ we see exclusion between macrophages and tumour cells at short lengthscales.

The wPCFs for Fig 13F reveal similar trends, with strong short range correlation between tumour cells and macrophages for $p < 0.5$ and exclusion for $0.5 < p < 0.8$. However, in Fig 13F tumour cells localise around the blood vessels only if macrophages with $p \approx 1$ are present. This is reflected by the strong short range clustering in $wPCF(r, P, T)$ for $p \approx 0.8$. We note that this wPCF distinguishes two populations of perivascular macrophages: those with $0.5 < p < 0.8$ which do not associate with tumour cells, and those with $p > 0.8$ which do.

Partial correlation between summary statistics and model parameters

The simulations in Fig 13 suggest that the six parameters which have been varied affect tumour size and composition and whether tumour cells localise around blood vessels. In order to understand how each parameter affects the spatial distributions of tumour cells and macrophages, we calculate the partial rank correlation coefficient (PRCC) between each parameter and a range of scalar statistics, based on model outputs at $t = 500$ for $n = 432$ simulations (see S7 Appendix: Partial correlations in 6-parameter sweep, for a full list of these statistics and their partial rank correlations with each parameter). In Table 2 we focus on selected statistics and three parameters, associated with macrophage extravasation ($P^\star$), macrophage localisation (χ_ξ^m) and macrophage phenotype (g_{crit}).

The PRCCs in Table 2 (and S7 Appendix: Partial correlations in 6-parameter sweep) show that, while changes in each parameter can significantly alter tumour size and morphology and the spatial and phenotypic distributions of the macrophages, these changes are not captured by every statistic: the PRCCs indicate which statistics resolve which model behaviours. For example, changes in the maximal extravasation rate ($P^\star$) are strongly positively correlated with changes in the number of macrophages, N_M at

Modality	Scalar statistic	$P^\star$	χ_ξ^m	g_{crit}
Cell counts	N_M	0.87	-0.03	-0.14
	N_T	-0.78	0.02	-0.12
Single-cell	$\bar{p}$	-0.47	-0.01	-0.64
	ϕ_2	-0.43	-0.00	-0.64
	N_{M_2}	-0.16	-0.04	-0.70
IHC	$\max(g_{BM}(r))$	-0.32	0.34	-0.31
IMC	$\int\int wPCF(r,p,B)dpdr$ (above wPCF = 1)	0.30	0.63	0.13

Table 2. Partial rank correlation coefficients (PRCCs) PRCCs corresponding to selected scalar statistics and parameters, based on simulations in which 6 parameters were randomly varied ($n = 432$). The full table, corresponding to variation of all 6 parameters and additional scalar statistics, can be found in S7 Appendix: Partial correlations in 6-parameter sweep.

$t = 500$, and negatively correlated with the number of tumour cells, N_T (PRCC = 0.87 and -0.78 respectively). However, since neither χ_ξ^m nor g_{crit} affects macrophage extravasation, changes in these parameters cannot be detected from cell counts. To establish how changes in g_{crit} affect ABM simulations, statistics which describe the distribution of macrophage phenotypes must be used (e.g., $\bar{p}, \phi_2$, or N_{M_2}). Each of these statistics strongly negatively correlates with g_{crit}.

A similar trend is evident when we consider how macrophage localisation is influenced by the parameter χ_ξ^m. This parameter does not correlate strongly with most scalar statistics, but there is some correlation between χ_ξ^m and the maximum value of $g_{BM}(r)$ (PRCC = 0.34), or the total volume of $wPCF(r, P, B)$ which is greater than $wPCF = 1$ (i.e., the total amount of clustering observed across all phenotypes and spatial scales; PRCC = 0.63). Importantly, reducing spatial statistics to a scalar metric ignores most spatial information, either by focussing on a particular length scale (as in $\max g_{BM}(r)$) or by averaging over space and phenotype (as in the volume of the wPCF).

Discussion

In this paper, we have presented a new ABM which shows how varying macrophage sensitivity to environmental cues can generate diverse patterns of tumour growth. We have further shown how hybrid ABMs can be used as a testbed for developing statistical descriptors of spatially resolved imaging data. Existing methods for analysing data typically resolve either spatial position without phenotype (e.g., IHC), or cell subtype without spatial localisation (e.g., single cell sequencing). State-of-the-art multiplex imaging provides information about the spatial position and phenotype, or subtype, of individual cells. These technological advances are, in turn, driving the development of new analytical methods that can summarise and quantify multiplex imaging data. In this paper, we have tested new methods for quantifying and interpreting relationships between cell phenotype and spatial localisation for future application to multiplex images using a hybrid ABM of tumour-macrophage interactions as a source of comparable synthetic data.

The ABM extends existing models in two important ways: the scope of the biology it describes and the introduction of a continuous and dynamic variable representing macrophage phenotype. The ABM incorporates multiple tumour-macrophage interactions, including paracrine signalling, migration of macrophages towards and away from the tumour, and rates of tumour cell killing which depend on macrophage phenotype. The model describes macrophage infiltration into a growing avascular tumour that is embedded within a healthy tissue, whose blood vessels act as point

sources of oxygen, sites of macrophage extravasation and niches for M_2-like macrophages which facilitate tumour cell intravasation and subsequent metastasis. ABM simulations reveal how careful coordination of the above processes is needed to simulate localisation of M_1- and M_2-like macrophages in different regions and the consequences of this spatial segregation on tumour cell proliferation and migration. In this way, the model shows how the spatial and phenotypic distribution of macrophages within tumour tissue influences tumour size and structure, while in turn being altered in response to tumour-derived microenvironmental cues. We used the ABM to show how changing macrophage sensitivity to specific environmental cues can generate diverse patterns of tumour growth, including growth as a compact mass, tumour elimination, diffuse or fragmented growth towards the vasculature, and the formation of perivascular niches containing M_2-like macrophages and tumour cells.

We proposed a series of statistics which summarise the spatial and phenotypic data that our ABM generates, and which could be routinely extracted from biological imaging data. We assessed the extent to which each statistic can distinguish between different simulation behaviours. Metrics such as cell counts can describe gross features including tumour elimination and tumour size. Statistics which resolve phenotypic distribution (and neglect spatial heterogeneity), such as the proportion of M_2-like macrophages, can distinguish some simulations that have similar cell counts based on macrophage behaviour. By contrast, cross-PCFs can identify differences in the spatial localisation of different cell populations across ABM simulations, as exemplified by the presence or absence of perivascular tumour cells and macrophages. We have also introduced the weighted PCF (wPCF) to describe relationships between cell phenotype and spatial localisation. By applying the wPCF to ABM data, we have shown how it can distinguish between different qualitative behaviours (and associated spatial patterns) that arise as a result of tumour-macrophage interactions, as described by Arwert *et al*, such as macrophage phenotype plasticity in the tumour bulk or migration of M_2-like macrophages towards blood vessels.

By combining the above statistics, we generate a multidimensional signature, or 'fingerprint', describing the output from ABM simulations. In particular, we can identify qualitative and quantitative differences in simulation behaviours, such as compact and diffuse tumour growth. Each statistic identifies different features of the simulated data. Although the wPCF provides the most detailed description, simpler statistics resolve differences in the simulation behaviours that do not rely on spatial and/or phenotypic cues.

There are many ways in which we could extend the ABM presented in this paper. Its extension to three spatial dimensions would yield more realistic data, but would be more computationally challenging to generate and analyse. The model could also be initialised from medical imaging data showing cell locations [93]. We could include a more detailed description of the vasculature, that accounts for tumour blood flow, angiogenesis and vascular remodelling [94–99]. This would enable us to account for tumour cell intravasation and metastasis, and to investigate how distinct tumour regions are initiated and evolve over time. Since macrophage phenotype represents just one facet of the immune response to cancer, a natural extension would be to include multiple immune cell populations, such as T cells, in our ABM. We could also analyse data from multiple time points to determine how the different statistics evolve over time and to understand when and how differences in the patterns of tumour growth, caused by varying specific parameter values, can be detected from the statistical descriptors. In future work, we plan to use the ABM to study tumour responses to treatments such as radiotherapy and chemotherapy, and to investigate the prognostic power of the statistical descriptors, individually and in combination, to identify not only those tumours that will benefit from such treatments but also those that would benefit from

immunotherapy. In the longer term, the insight gained from such studies could inform the development of new biomarkers for identifying those patients who will respond to a particular treatment.

Future work will involve applying these statistics to multiplex imaging data, in order to validate their use in biological and clinical settings. Applying these statistics to medical images would enable their high-throughput, automated quantification and comparison in a manner that goes beyond expert visual inspection and is arguably more interpretable than machine learning approaches [90, 91]. While some relationships between cells described by the wPCF may be apparent from visual inspection, it is increasingly difficult to identify such relationships by eye as the number of multiplex markers increases; typically only 5 or 6 distinct cell types may be visualised simultaneously (from a potential pool of up to 40 markers).

The 'statistical fingerprint' described in this paper could be augmented with other features. For example, topological data analysis can describe spatial features such as immune deserts that exist in noisy data [34], or changes in tumour and vascular architecture in response to radiotherapy [92]. Multiple spatial statistics can be combined to obtain more detailed descriptions of 2D data [22], or new statistics can be derived from networks of cell contact [21] or observations of immune cell locations [100, 101].

In concluding, we note that the statistical methods described in this paper are not limited to the analysis of synthetic data generated by off-lattice ABMs. For example, the PCF, the cross-PCF, and the wPCF can be applied to data at a range of length scales including coarser-grained data (e.g., meso-scale images from PET scans or MRIs). In this study, macrophage phenotype p is a continuous label for the point patterns, but the wPCF can be applied to other continuous variables, including subcellular labels (e.g., cell cycle position or expression levels of a particular gene) and/or local levels of diffusible species (e.g., oxygen, glucose or a chemotherapeutic drug), as shown by examples in S5 Appendix: wPCF for comparing two continuous labels. In particular, we note that before cell segmentation and classification, multiplex images consist of continuous labels describing pixel intensities for a range of markers. These intensities could be viewed as continuous variables for the wPCF, permitting comparison of pixels with 'high' and 'low' levels of each marker without requiring the definition of threshold intensities.

Taken together, the ABM and statistical approach presented in this paper represent a methodology by which multidimensional spatially resolved imaging data can be examined. We used the ABM to generate multidimensional spatial data of the type produced by multiplex imaging modalities, and computed statistical descriptions of this data within a framework where interactions between different cell types are well understood. This enabled us to compare a range of existing and new statistical descriptors, which can be used to better understand multidimensional spatial data from multiplex imaging.

Supporting information

S1 Appendix: Model Description

Our ABM extends existing ABMs developed to describe infiltration of microbeads [35] and macrophages [34] into tumour spheroids growing *in vitro*. Here, we consider an *in vivo* scenario, where tumour cells are embedded within a tissue containing stromal cells and where oxygen is supplied by blood vessels. We use an overlapping-spheres model, representing each cell as a point whose movement and position are determined by balancing the forces that act on the cell. We distinguish four types of agent: **macrophages**, **tumour cells**, **stromal cells** and **necrotic cells**. We also consider five diffusible species: **oxygen** ($\omega(\mathbf{x},t)$), **colony stimulating factor-1** (CSF-1, $c(\mathbf{x},t)$), **transforming growth factor-β** (TGF-β, $g(\mathbf{x},t)$), **epidermal growth factor** (EGF, $\epsilon(\mathbf{x},t)$), and **C-X-C motif chemokine 12** (CXCL12, $\xi(\mathbf{x},t)$). Their dynamics are defined by reaction-diffusion equations (RDEs). Following [34,35], we use the Chaste (Cancer, Heart and Soft Tissue Environment) modelling environment to solve the model equations [30–32].

Diffusible species

The distribution of the diffusible species is modelled using RDEs, with each equation solved numerically on a triangular finite element mesh large enough to span the domain. For the simulations in this manuscript, the domain Ω is taken to be a square with height and width equal to 50 cell diameters (1 mm in dimensional units). In this Section we briefly describe the role of each diffusible species, the factors that influence rates of production and depletion, and the RDE that defines its evolution. Fig 1 in the main text summarises how the diffusible species interact with tumour cells and macrophages.

Oxygen (ω) Oxygen diffuses through the domain from blood vessels which are represented by static point sources, and is consumed by tumour cells and stromal cells. We assume that the oxygen concentration at each blood vessel is constant, ω_0.

Following previous models [34,35] we assume that since the timescale of diffusion for an oxygen molecule is faster than the timesteps in our model for describing cell movement, the distribution of oxygen can be approximated using a steady state solution. For simplicity, we rescale concentrations with a factor of ω_0, so that $\omega = 1$ at each blood vessel. The governing equation is then given by

$$0 = D_\omega \nabla^2 \omega - \kappa_\omega \omega \sum_i \delta(\mathbf{x} - \mathbf{x}_i) \tag{18}$$

for $\mathbf{x} \in \Omega$, where $\mathbf{x}_i$ is the location of stromal or tumour cell i, the parameter D_ω is the assumed constant diffusion coefficient of oxygen, κ_ω is the oxygen consumption rate and δ is the delta function ($\delta(\mathbf{x}) = 1$ when $\mathbf{x} = 0$, $\delta(\mathbf{x}) = 0$ otherwise).

CSF-1 (Colony stimulating factor-1, c) CSF-1 acts as a chemoattractant for macrophages and is produced by tumour cells. It acts with EGF as part of a paracrine loop [69,70] and is a factor which can recruit macrophages from vasculature [23]. We assume that CSF-1 is produced at a constant rate, κ_c, by each tumour cell, and that CSF-1 in the tumour microenvironment decays at a constant rate, λ_c. The distribution of CSF-1 is therefore described by the equation

$$\frac{\partial c}{\partial t} = D_c \nabla^2 c - \lambda_c c + \kappa_c \sum_i \delta(\mathbf{x} - \mathbf{x}_i) \tag{19}$$

where D_c is the diffusion coefficient of CSF-1. We apply Neumann boundary conditions ($\frac{\partial c}{\partial \mathbf{x}} = 0$) on the boundaries of the domain Ω, and assume that initially $c(\mathbf{x}, t = 0) = 0$ for all $\mathbf{x}$.

TGF-β (Transforming growth factor-β, g) TGF-β causes macrophages to alter their phenotype, and is produced by tumour cells. We assume that TGF-β decays at a constant rate, λ_g, and is produced by tumour cells at a constant rate, κ_g. Combining these effects, we arrive at the following RDE for $g(\mathbf{x}, t)$:

$$\frac{\partial g}{\partial t} = D_g \nabla^2 g - \lambda_g + \kappa_g \sum_i \delta(\mathbf{x} - \mathbf{x}_i) \tag{20}$$

where D_g is the diffusion coefficient of TGF-β and the sum i is over all tumour cells. As for CSF-1, we impose Neumann boundary conditions ($\frac{\partial g}{\partial \mathbf{x}} = 0$ on the boundaries of the domain Ω), and assume that $g(\mathbf{x}, t = 0) = 0 \; \forall \mathbf{x} \in \Omega$.

EGF (Epidermal growth factor, ϵ) EGF is a diffusible chemoattractant for tumour cells that is produced by tumour associated macrophages. We assume that EGF is produced by macrophages at a rate that is linearly dependent on their phenotype p, so that macrophages with an extreme M_1-like phenotype ($p = 0$) produce no EGF and macrophages with an extreme M_2-like phenotype ($p = 1$) produce κ_ϵ. We assume that EGF naturally decays at a constant rate, λ_ϵ. Combining these effects, we obtain the following RDE for EGF, $\epsilon(\mathbf{x}, t)$:

$$\frac{\partial \epsilon}{\partial t} = D_\epsilon \nabla^2 \epsilon - \lambda_\epsilon \epsilon + \kappa_\epsilon \sum_i p_i \delta(\mathbf{x} - \mathbf{x}_i) \tag{21}$$

where D_ϵ is the constant diffusion coefficient of EGF, κ_ϵ is the maximum rate of production of EGF, the sum is over all macrophages i, and p_i is the phenotype of macrophage i ($p_i \in [0, 1]$, and is fully defined below).

CXCL12 (C-X-C motif chemokine 12, ξ) CXCL12 is a diffusible chemoattractant for M_2-like macrophages which is expressed by perivascular cancer-associated fibroblasts (CAFs) [23]. We suppose that CAFs are localised close to blood vessels, and, hence, make the simplifying assumption that blood vessels act as constant sources of CXCL12. We therefore assume that the concentration of CXCL12, ξ, is maintained at a fixed value $\xi = \xi_0$ at all blood vessels. We assume that CXCL12 decays naturally at a constant rate, λ_ξ. The distribution of CXCL12 in the domain is then given by

$$0 = D_\xi \nabla^2 \xi - \lambda_\xi \tag{22}$$

where $D_\xi$is the diffusion coefficient for CXCL12.

Agents

Our model distinguishes four types of agents: **macrophages**, **tumour cells**, **stromal cells** and **necrotic cells**. In addition, a number of **blood vessels** are located at fixed spatial positions throughout each simulation. The location of each agent is represented by its cell centre.

We use Newton's second law to derive the equations of motion for macrophages, tumour cells, stromal cells and necrotic cells. Neglecting inertial effects in the overdamped limit, the force balance for cell i can be written as:

$$\nu \frac{d\mathbf{x}_i}{dt} = \mathbf{F}_i, \tag{23}$$

where ν is the drag coefficient (assuming that the drag on a cell is proportional to its velocity), and $\mathbf{F}_i$ denotes the net force acting on cell i. The forces that act on each cell type are indicated in Fig 1.

All cells are subject to mechanical forces due to cell-cell interactions (incorporating repulsion due to volume exclusion and attraction due to intercellular adhesion). Here we adopt the overlapping spheres approach [82, 84], assuming that two cells interact if the distance between their cell centres is less than a fixed radius of interaction R_{int}. Specifically, for cells at locations $\mathbf{x}_i$ and $\mathbf{x}_j$, if $|\mathbf{x}_i - \mathbf{x}_j| < R_{\text{int}}$ then the interaction force between cells i and j is parallel to the vector $(\mathbf{x_i} - \mathbf{x_j})$ connecting their centres. The resting spring length between the cell centres, $s_{i,j}$, is the sum of the equilibrium spring lengths associated with each cell ($s_{i,j} = s_i + s_j$). For most cells i, the resting spring length is equal to the approximate radius of a cell ($s_i = R_{\text{Cell}} \equiv 0.5$ in non-dimensional units). For newly divided and necrotic cells, s_i changes over time to account for cell growth and shrinkage (see below).

The net mechanical force acting on each cell is the sum of all interaction forces due to cells within the interaction radius:

$$\mathbf{F}_i^m = \sum_{\{j \,:\, |\mathbf{x}_i - \mathbf{x}_j| \leq R_{\text{int}}\}} \mathbf{F}_{i,j}^m. \tag{24}$$

where $\mathbf{F}_{i,j}^m$ is the mechanical force between cells i and j. This force always points in the direction of the vector connecting the cell centres and has magnitude:

$$|\mathbf{F}_{i,j}^m| = \begin{cases} \mu s_{i,j} \log\left(1 + \frac{x}{s_{i,j}}\right) \text{ if } x < 0 \text{ (Repulsive)} \\ \mu x s_{i,j} \exp\left(-\alpha \frac{x}{s_{i,j}}\right) \text{ if } x \geq 0 \text{ (Adhesive)} \end{cases} \tag{25}$$

where $x = |\mathbf{x}_i - \mathbf{x}_j| - s_{i,j}$ is the overlap between cells i and j, μ is a parameter describing the spring stiffness and α determines the strength of intercellular adhesion between neighbouring cells.

We normalise lengths according to the lengthscale R_{Cell}, assuming that 1 cell diameter $= 2R_{\text{Cell}} = 20\mu$m. Following cell division, the radius of the daughter cells is set to $s_i = \frac{R_{\text{Cell}}}{2}$ and increases linearly over one hour until $s_i = R_{\text{Cell}}$. For necrotic cells, s_i decreases linearly to zero over $\bar{\tau}_i$ hours (see below), and then the cell is removed from the simulation. The associated spring stiffness μ of springs attached to a necrotic cell is reduced linearly at the same rate.

Macrophages

Extravasation Macrophages extravasate from blood vessels in response to increasing concentrations of CSF-1. At each timestep, the concentration of CSF-1 at each blood vessel is converted into a probability of extravasation such that the probability of a macrophage extravasating in a one hour time period is P_{ex}, where:

$$P_{\text{ex}} = P^\star \times \frac{c}{c + c_{1/2}} \tag{26}$$

where $P^\star$ is a parameter controlling the maximum possible probability of macrophage extravasation per hour from each vessel, and $c_{1/2}$ is the concentration of CSF-1 at which this probability is half-maximal.

Phenotype The diffusible species included in our model interact with receptors on macrophages and tumour cells to induce behaviours such as chemotaxis or upregulation of different receptors. While we do not explicitly model surface receptors, the receptor CXCR4 (C-X-C motif chemokine receptor type 4) plays an important role in this

system: when exposed to TGF-β, macrophages increase expression of CXCR4 and become sensitive to gradients of CXCL12 [23]. We model this by associating each macrophage with a phenotype, $p \in [0, 1]$, which determines the extent to which it exhibits M_1- or M_2-like behaviour. When first entering the simulation, macrophages have phenotype $p = 0$. Exposure to TGF-β causes macrophage phenotype to increase irreversibly. The change in phenotype for a macrophage i at location $\mathbf{x}_i$ is:

$$\frac{dp}{dt} = \mathcal{H}(g(\mathbf{x}_i, t) - g_{\text{crit}}) \times \mathcal{H}(1 - p)\Delta p \tag{27}$$

where $g(\mathbf{x}_i, t)$ is the concentration of TGF-β at $\mathbf{x}_i$ at time t, $\mathcal{H}(x)$ is the Heaviside function ($\mathcal{H}(x) = 0$ if $x < 0$, $\mathcal{H}(x) = 1$ otherwise), the parameter Δp determines the rate of phenotype change and g_{crit} is a critical threshold of TGF-β above which macrophage phenotype increases. When $p = 1$ the macrophage has a fully M_2 phenotype, and p no longer changes.

Phenotype affects three key macrophage behaviours:

- Production of EGF
- Rate of tumour cell killing
- Response to gradients of chemokines (CSF-1 and CXCL12)

The dependence of EGF production on p is described above in Equation (21); we now describe how cell killing and phenotype-dependent chemotaxis are implemented.

Cell killing Macrophage phenotype determines the likelihood of a macrophage killing a tumour cell in any given timestep, characteristic of M_1-like macrophages. The probability of a macrophage i killing a tumour cell in one hour is

$$P_\varphi = \begin{cases} P_\varphi^\star \times (1 - \frac{p_i^{10}}{p_i^{10} + 0.5^{10}}) & \text{for } t_\varphi \geq t_{\text{cool}} \\ 0 & \text{otherwise} \end{cases} \tag{28}$$

where $P_\varphi^\star$ is the maximum probability of a given macrophage killing a tumour cell in one hour, and p_i is the macrophage phenotype. t_φ is the time since the macrophage last killed a tumour cell, and t_{cool} is a parameter which determines a 'cooldown' period within which the macrophage cannot kill another tumour cell. At each timestep the probability of each macrophage attempting cell killing is evaluated. If a macrophage attempts killing, any cells within distance 1 of the macrophage are identified and, if there is at least one tumour cell, one of the tumour cells is selected at random to be killed. The tumour cell is labelled as necrotic, and t_φ is reset to 0 for that macrophage.

Force laws for macrophages In addition to mechanical forces, macrophages are also subject to forces which model random movement through their neighbourhood and chemotactic forces due to spatial gradients of CXCL12 and CSF-1. We model random macrophage movement by applying a force $\mathbf{F}_i^r = (F_x^r, F_y^r)$ at each timestep, given by:

$$\mathbf{F}_i^r = \sqrt{2Ddt}\,\boldsymbol{n} \tag{29}$$

where the coefficient D describes the strength of the random force and $\boldsymbol{n} = (n_x, n_y)$, where n_x and n_y are random variables drawn from a standard normal distribution.

The chemotactic force applied due to the gradients of chemoattractants for a macrophage i depends on its phenotype, p_i, and is:

$$\mathbf{F}_i^\chi = \chi_c^m (1 - p_i) \frac{\nabla c}{|\nabla c|} + \chi_\xi^m p_i \frac{\nabla \xi}{|\nabla \xi|} \tag{30}$$

where χ_c^m and χ_ξ^m are parameters controlling macrophage sensitivity to gradients of CSF-1 and CXCL12 respectively. This means that macrophage sensitivity to both CSF-1 and CXCL12 gradients scales linearly with macrophage phenotype. The full form of the forces applied to a macrophage i in Equation (23) is therefore:

$$\mathbf{F}_i = \mathbf{F}_i^m + \mathbf{F}_i^r + \mathbf{F}_i^\chi. \tag{31}$$

Tumour cells

Cell cycle / proliferation Each tumour cell has an internal cell cycle which progresses at a rate that depends on the local oxygen concentration. We introduce two oxygen thresholds, $\omega_{\mathrm{H}}^{\mathrm{tum}} < 1$ and $\omega_{\mathrm{N}}^{\mathrm{tum}} \leq \omega_{\mathrm{H}}^{\mathrm{tum}}$. When $\omega_{\mathrm{N}}^{\mathrm{tum}} < \omega \leq \omega_{\mathrm{H}}^{\mathrm{tum}}$, the cell becomes hypoxic and cell cycle progression pauses until ω returns above this threshold. If $0 \leq \omega \leq \omega_{\mathrm{N}}^{\mathrm{tum}}$, then the tumour cell dies and becomes a necrotic cell. Additionally, we account for contact inhibition in our model. If a cell's area drops to a proportion ${A_i^H}_{\mathrm{tum}}$ of its target area, then its cell cycle pauses until space is available for proliferation. Each tumour cell i has a subcellular variable denoted T_i which tracks how far through the cell cycle cell i is. Mathematically, T_i progresses according to the equation

$$\frac{dT_i}{dt} = \min\left(\mathcal{H}(\omega(\mathbf{x}_i, t) - \omega_{\mathrm{H}}^{\mathrm{tum}}), \mathcal{H}(A_i - {A_i^H}_{\mathrm{tum}} \pi R_{\mathrm{Cell}}^2)\right). \tag{32}$$

where A_i is the area of the cell calculated as $A_i = \pi r_i^2$ and r_i is the estimated cell radius calculated via the average separation between cells within the interaction radius of i.

Each cell has a target cell cycle duration τ_i, drawn uniformly at random from a distribution of $U\left(0.75\tau_{\mathrm{tum}}, 1.25\tau_{\mathrm{tum}}\right)$ with average cell cycle duration τ_{tum}. When $T_i = \tau_i$ for a given cell, cell division occurs and a new cell is placed half a cell diameter away in a randomly chosen direction. Both cells are assigned new cell cycle durations. T_i is set to 0 for each cell, and then evolves according to Eq (32). The equilibrium spring length associated with the new cells is set to $s_i = 0.5R_{\mathrm{Cell}}$ to account for the reduced size of the cells, and increased linearly over the course of 1 hour until it reaches $s_i = R_{\mathrm{Cell}}$.

Force laws for tumour cells Tumour cells move according to the force balance described in Equation (23). The force term incorporates the intercellular interactions described above, but also accounts for chemotaxis between tumour cells and EGF. This force, denoted $\mathbf{F}_i^{\chi_\epsilon}$, has the form

$$\mathbf{F}_i^{\chi_\epsilon} = \chi_\epsilon^T \frac{\nabla \epsilon}{|\nabla \epsilon|} \tag{33}$$

where χ_ϵ^T is parameter determining the sensitivity of tumour cells to the EGF gradient. The force term used in Equation (23) is therefore

$$\mathbf{F}_i = \mathbf{F}_i^m + \mathbf{F}_i^{\chi_\epsilon}. \tag{34}$$

Stromal cells

Cell cycle / proliferation Stromal cells follow the same cell cycle model as tumour cells, but are parameterised differently to account for tumour cells increased ability to survive in adverse conditions such as lower oxygen environments or under increased pressure from neighbouring cells. Stromal cells possess the same subcellular age variable T_i, which progresses according to the equation

$$\frac{dT_i}{dt} = \min\left(\mathcal{H}(\omega(\mathbf{x}_i, t) - \omega_{\mathrm{H}}^{\mathrm{str}}), \mathcal{H}(A_i - A_{i\ \mathrm{str}}^{H} \pi R_{\mathrm{Cell}}^2)\right) \tag{35}$$

where $\omega_{\mathrm{H}}^{\mathrm{str}}$ is a parameter determining the oxygen threshold below which stromal cells become hypoxic and $A_{i\ \mathrm{str}}^{H}$is a parameter defining the proportion of a stromal cells target area below which size the cell cycle stops.

As with tumour cells, we define a second oxygen threshold $\omega_{\mathrm{N}}^{\mathrm{str}}$, below which stromal cells become necrotic. In our model, we take $\omega_{\mathrm{H}}^{\mathrm{str}} < \omega_{\mathrm{H}}^{\mathrm{tum}}$ and $A_{i\ \mathrm{str}}^{H} < A_{i\ \mathrm{tum}}^{H}$ to account for the ability of tumour cells to proliferate in more adverse environments than stromal cells.

Force laws Stromal cells are subject to the same intercellular forces as macrophages and tumour cells, defined in Equations (24)-(25). The force balance for stromal cells used in Equation (23) is therefore simply

$$\mathbf{F}_i = \mathbf{F}_i^m. \tag{36}$$

Necrotic cells

When a stromal cell or tumour cell is marked for cell death, either through oxygen starvation or being killed by a macrophage, it irreversibly becomes necrotic. A necrotic cell i occupies space for $\bar{\tau}_i$ hours, where $\bar{\tau}_i$ is drawn from a uniform distribution $U(0.75\,\bar{\tau}, 1.25\,\bar{\tau})$ and $\bar{\tau}$ is the average duration of necrosis. Over this time period, the necrotic cell shrinks in size, by reducing its equilibrium spring constant, s_i, at a constant rate until $s_i = 0$. The cell is then removed from the simulation. While s_i is being reduced, the spring constant μ associated with the cell is also reduced to 0 at a constant rate over the same time period, accounting for weakening intercellular forces between degrading necrotic cells and other cells.

Initial and boundary conditions for cells

RDEs describing the diffusible species are solved numerically on a regular triangular mesh with edge length 1 cell diameter. We initialise the simulation by selecting lattice sites to act as point vessels, which do not occupy space or interact directly with cells in our model. All lattice sites more than R_{B} cell diameters from the centre of the domain are possible blood vessel sites, and N_{B} of these sites are chosen, at random, to be blood vessels. This ensures that the centre of the domain is at least R_{B} cell diameters from the nearest blood vessel and, hence, that there is sufficient space to observe macrophage movement between blood vessels and the tumour, while also ensuring that cells near the domain boundaries are well-oxygenated.

The domain is initialised with stromal cells filling the domain in rows that are 0.75 cell diameters apart, with alternating rows offset by 0.375 cell diameters (i.e., forming a hexagonal lattice). We place four tumour cells in a cluster at the centre of the domain approximately 0.5 cell diameters apart. Stromal and tumour cells are assigned cell cycle durations τ_i from the relevant distributions, and cell cycle progression times T_i from a uniform distribution $U(0, \tau_i)$. All cells are constrained to remain within the domain by imposing reflective boundary conditions.

Schematic

Fig 14 is a schematic showing the order in which the above processes occur. After initialisation, at each timestep the concentrations of diffusible species are updated and macrophage extravasation occurs. Individual cell cycles, target spring lengths,

phenotype changes, and proliferation are then updated. Finally, cells are moved to their new locations.

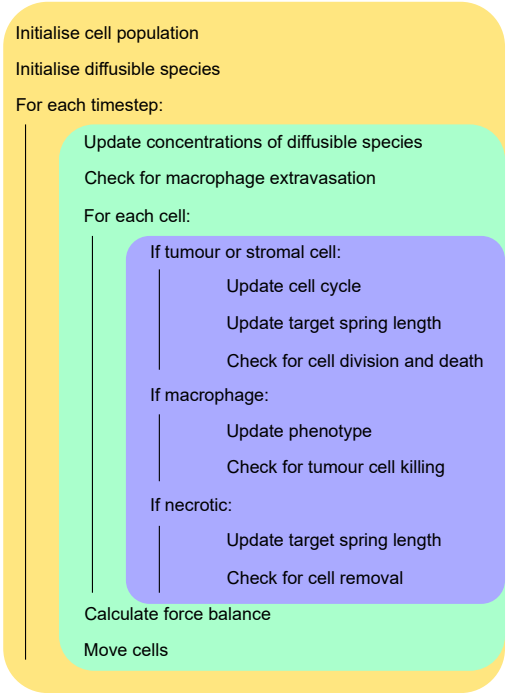

Fig 14. Model overview
Schematic showing an overview of the order in which the model is initialised, updated at each timestep, and updated at a cell level.

Table of ABM parameters

In Table 3 we list the model parameters used, their default dimensional and dimensionless values or ranges, and supporting references where these are available. Value of some parameters, indicated with $*$, have been estimated based on model behaviour.

Our model is implemented such that typical scales are given by:

- Length: 1 cell diameter (taken as 20μm) is 1 unit of length.
- Time: 1 hour is 1 unit of time.
- Concentration: the boundary concentration of oxygen is 1.

Table 3. Table of parameters

Parameter	Symbol	Value	Units	Dimensionless value	Reference
Timestep	dt	1/120	hours	1/120	[35]
Damping coefficient	ν	0.4	$N\,\mathrm{s}^{-1}\mathrm{m}^{-1}$	1	[35, 81, 83]
Blood vessel exclusion radius	R_{B}	340	μm	17	$*$
Number of blood vessels	N_{B}	35	-	35	$*$
Radius of interaction	R_{int}	21 - 36	μm	1.5	[83]
Cell radius	R_{Cell}	7 - 12	μm	0.5	[85]
Spring constant	μ	3 - 50	μg Cell diameter^{-1} hours^{-2}	5	[83, 86]
Intercellular adhesion scaling coefficient	α	-	-	5	[35, 83]
Random force coefficient	D	Assumed	Cell diameter^2 hours^{-1}	0.01	[35]
Concentration of oxygen at blood vessels	ω_0	100 - 150	mm Hg	1.0	[87, 88]
Oxygen diffusion coefficient	D_ω	1,750	$\mu\mathrm{m}^2\,\mathrm{s}^{-1}$	1.0	[35, 89]
Oxygen consumption rate	κ_ω	20×10^{-18}	mol cell^{-1} s^{-1}	0.03	[35, 89]
Tumour cell hypoxia threshold	$\omega_{\mathrm{H}}^{\mathrm{tum}}$	30 - 70	mm Hg	0.01	[35], $*$
Tumour cell necrosis threshold	$\omega_{\mathrm{N}}^{\mathrm{tum}}$	10	mm Hg	0.01	[35], $*$
Stromal cell hypoxia threshold	$\omega_{\mathrm{H}}^{\mathrm{str}}$	30 - 70	mm Hg	0.1	[35], $*$
Stromal cell necrosis threshold	$\omega_{\mathrm{N}}^{\mathrm{str}}$	10	mm Hg	0.01	[35], $*$
CSF-1 diffusion coefficient	D_c	160	$\mu\mathrm{m}^2\,\mathrm{s}^{-1}$	1.0	[51], $*$
CSF-1 production rate	κ_c	1.7×10^{-23}	mol m^{-3} s^{-1}	0.25	[51], $*$
CSF-1 decay rate	λ_c	1.9×10^{-4}	s^{-1}	0.02	[51], $*$
TGF-β diffusion coefficient	D_g	21.3	$\mu\mathrm{m}^2\,\mathrm{s}^{-1}$	0.1	[103], $*$
TGF-β production rate	κ_g	0.01-0.11	ng (million cells $\times$ day)$^{-1}$	0.1	[103], $*$
TGF-β decay rate	λ_g	0.23-0.34	min^{-1}	0.1	[104], $*$
EGF diffusion coefficient	D_ϵ	160	$\mu\mathrm{m}^2\,\mathrm{s}^{-1}$	0.2	[51], $*$
EGF production rate	κ_ϵ	1.7×10^{-23}	mol m^{-3} s^{-1}	0.2	[51], $*$
EGF decay rate	λ_ϵ	1.9×10^{-4}	s^{-1}	0.1	[51], $*$
Concentration of CXCL12 at blood vessels	ξ_0	Assumed	mm Hg	1.0	$*$
CXCL12 diffusion coefficient	D_ξ	150×10^{-6}	$\mu\mathrm{m}^2\,\mathrm{s}^{-1}$	1.0	[102], $*$
CXCL12 decay rate	λ_ξ	2×10^{-5}	s^{-1}	0.02	[102], $*$
Chemotaxis sensitivity coefficient (macrophage to CSF-1)	χ_c^m	Assumed	Cell diameter^2 hour^{-1} nM^{-1}	0 - 5	$*$
Chemotaxis sensitivity coefficient (macrophage to CXCL-12)	χ_ξ^m	Assumed	Cell diameter^2 hour^{-1} nM^{-1}	0 - 5	$*$
Chemotaxis sensitivity coefficient (tumour cell to EGF)	χ_ϵ^T	Assumed	Cell diameter^2 hour^{-1} nM^{-1}	0 - 5	$*$
Half-maximal macrophage extravasation CSF-1 concentration	$c_{1/2}$	Assumed	mol Cell diameter^{-3}	0.1 - 0.5	$*$
Maximum probability of macrophage extravasation	$P^\star$	Assumed	-	0.01 - 0.1	$*$
Critical TGF-β threshold	g_{crit}	Assumed	mol Cell diameter^{-3}	0 - 2	$*$
Average duration of tumour cell cycle	τ_{tum}	24	hours	24	[35], $*$
Average duration of stromal cell cycle	τ_{str}	32	hours	32	$*$
Average duration of necrosis	$\bar{\tau}$	48	hours	48	[35], $*$
Proportion of equilibrium area for contact inhibition (tumour cell)	$A_i^H{}_{\mathrm{tum}}$	Assumed	-	0.6	$*$
Proportion of equilibrium area for contact inhibition (stromal cell)	$A_i^H{}_{\mathrm{str}}$	Assumed	-	0.75	$*$
Macrophage phenotype increment	Δp	Assumed	hours^{-1}	0.01	$*$
Maximum killing probability per hour	$P_\varphi^\star$	Assumed	-	0.2	$*$
Killing cooldown duration	t_{cool}	Assumed	hours	4	$*$

S2 Appendix: Tumour growth in the absence of macrophages

We use the same parameter values as were used to generate Fig 4 but extend the simulation time to $t = 2000$ hours.

The simulation results presented in Fig 15A show that at long times the tumour evolves to a steady state for which the rate at which oxygen-rich cells on the outer tumour boundary proliferate balances the rate at which necrotic cells in the oxygen-starved core are degraded. We have positioned the blood vessels and fixed the properties of the tumour cells (e.g., the thresholds for hypoxia and necrosis, $\omega_{\mathrm{H}}^{\mathrm{tum}}$ and $\omega_{\mathrm{N}}^{\mathrm{tum}}$, and the cell cycle duration τ_i) so that a tumour initially located at the centre of the domain grows as a compact mass. At long times, the total number of tumour cells and necrotic cells remains approximately constant (see Figures 15B and C). Further, the tumour attains its steady state before it can spread to the surrounding blood vessels.

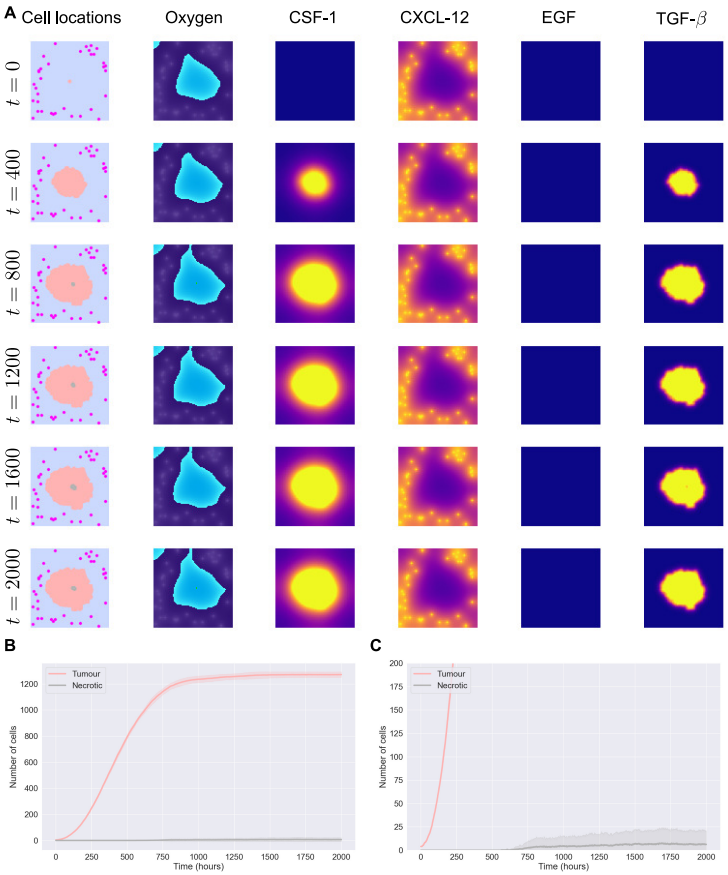

Fig 15. Tumour growth in the absence of macrophages
A: When no macrophages enter the simulation, the tumour grows as a compact mass in response to oxygen supplied from blood vessels. At long times, the tumour attains a steady state with a central necrotic core. At steady state, the proliferation rate of cells on the oxygen rich outer tumour boundary balances the death rate of cells in the central, oxygen-starved necrotic core. B/C: Tumour cell counts reach a steady state at approximately $t = 1000$. After approximately 750 a necrotic core forms due to hypoxia at the tumour centre. (Panel C shows a magnified view of panel B).

S3 Appendix: Derivation of PCFs and Cross-PCFs from wPCF

In the main text, we show how the wPCF can identify correlations between a continuous label (phenotype) and a categorical one (cell type). We now explain how we can extend the wPCF to study correlations between two continuous labels. Consider two labels u and v associated with each cell, which may be continuous or discrete. For two target marks U and V, the wPCF in its most general form can be written as:

$$wPCF(r, U, V) = \frac{1}{W_U W_V} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{A}{A_{r_k}(\mathbf{x}_i)} w_u(U, u_i) w_v(V, v_j) I_{[r_k, r_{k+1})}(|\mathbf{x}_i - \mathbf{x}_j|) \quad (37)$$

where w_u and w_v are weighting functions, and $W_U = \sum_i w_u(U, u_i)$ and $W_V = \sum_i w_v(V, v_i)$ are the total weights associated with each mark. The functional forms of w_u and w_v depend on several factors:

- Are u and v discrete or continuous marks?
- What are the ranges of u and v?
- At what resolution do we wish to identify correlations?

Through appropriate choices of w_U and w_V, Equation (37) reduces to the ordinary PCF, the cross-PCF, or the discrete-continuous form of the wPCF that was introduced in the main text. Table 4 illustrates weighting functions which achieve these different cases.

Table 4. Choices of w which simplify the wPCF. By appropriately choosing w_u and w_v, Equation (37) reduces to the wPCF presented in Equation (17), to the cross-PCF in Equation (15), or to the original definition of the PCF.

Function name	Notation	u: Discrete?	v: Discrete?	w_u	w_v
PCF	$g(r)$	Discrete	Discrete	$w_u(U, u_i) = \Theta(U, u_i)$	$w_v(V, v_i) = \Theta(V, v_i)$ with $u = v$
Cross-PCF	$g_{UV}(r)$	Discrete	Discrete	$w_u(U, u_i) = \Theta(U, u_i)$	$w_v(V, v_i) = \Theta(V, v_i)$
wPCF	$wPCF(r, U, V)$	Continuous	Discrete	$w_u(U, u_i)$	$w_v(V, v_i) = \Theta(V, v_i)$
wPCF	$wPCF(r, U, V)$	Continuous	Continuous	$w_u(U, u_i)$	$w_v(V, v_i)$

S4 Appendix: Comparison of different weighting functions

In this appendix we consider different choices of the weighting, or kernel, function w_p, and apply them to synthetic data. In Fig 16 we place 50 pink crosses uniformly along the line $y = 1$, and 1000 circles according to complete spatial randomness throughout the (2×2) square domain. The circles are labelled with a 'phenotype' p based on their distance from $y = 1$, such that for circle i the label p_i is given by:

$$p_i = \begin{cases} 1 - y_i \text{ if } y_i < 1 \\ 2 - y_i \text{ if } y_i \geq 1 \end{cases} . \tag{38}$$

By construction, this results in two prominent values of p which are correlated with pink crosses at distance r: $p = r$ (below $y = 1$) and $p = 1 - r$ (above $y = 1$).

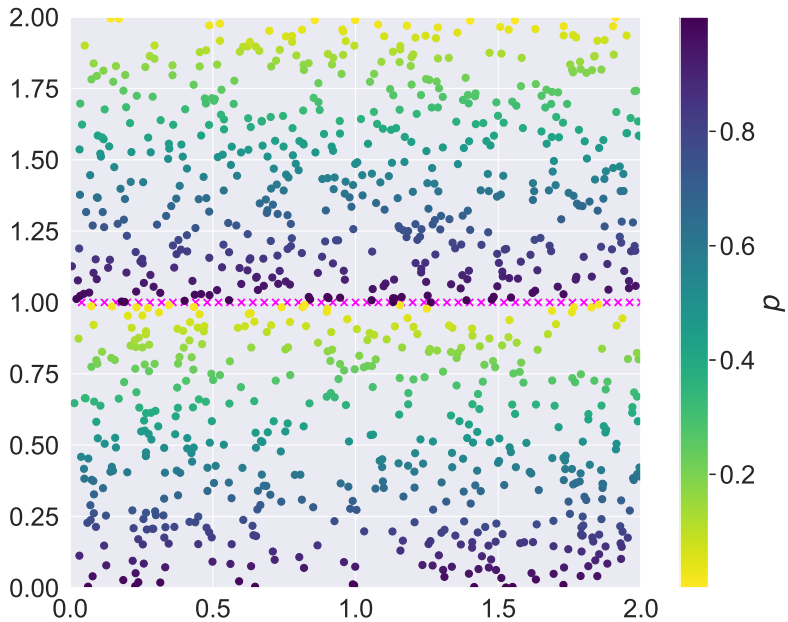

Fig 16. Synthetic data generated to demonstrate different weighting functions
The test data consists of 50 pink crosses ('blood vessels' uniformly positioned along the line $y = 1$) and 1000 circles randomly positioned in the domain ('macrophages' with different phenotypes placed according to complete spatial randomness). Circles are labelled according to their distance from the line $y = 1$.

In Fig 17 we present wPCFs for this point cloud for different choices of the weighting function $w_p(P, p_i)$ in Equation (17). We consider here two choices of functional form for w_p. Firstly, wPCFs in Fig 17A use the kernel discussed in the main text in Equation (16), which has the form $w_p(P, p_i) = \max\left(1 - \frac{|P - p_i|}{\Delta P}, 0\right)$, for different values of ΔP. In the main text, we use this weighting function with $\Delta P = 0.2$, which we choose as it leads to a sufficiently compact kernel to identify variations in phenotype while being broad enough to reduce noise. Weighting functions in Fig 17B have the form $w_p(P, p_i) = \max\left(1 - \frac{(P - p_i)^2}{\Delta P}, 0\right)$, which leads to a smoother kernel and hence a

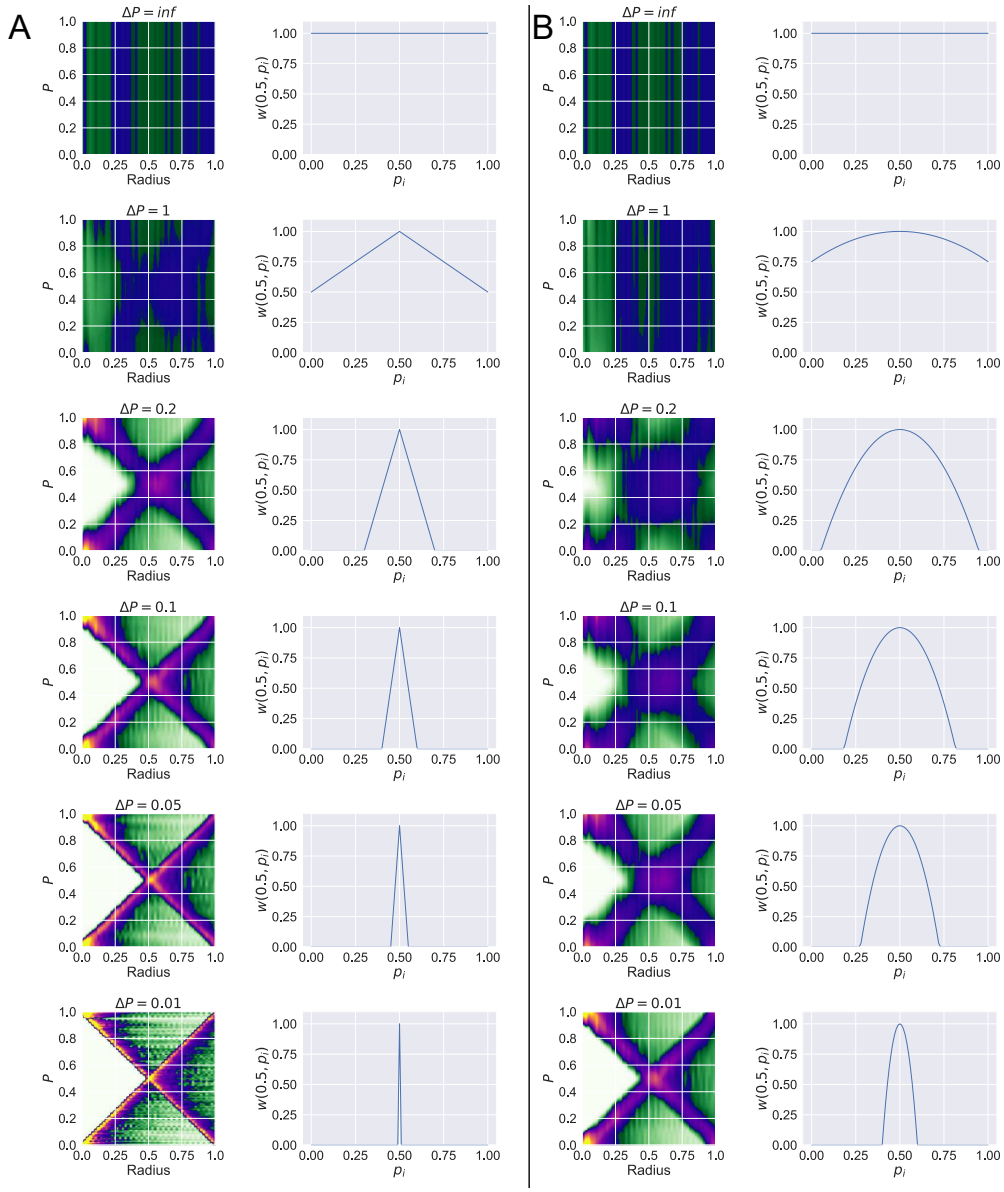

Fig 17. wPCFs generated from different weighting functions
$wPCF(r, P, B)$ for the point pattern in Fig 16. Changing the shape of the weighting function adjusts the balance between signal and noise in the wPCF: the narrower the support of w, the more clearly the relationship between label and distance can be discerned. Using a weighting function with extremely narrow support relative to the range of labels results in more noise in the wPCF, most evident in the triangular weighting functions with $\Delta P = 0.05$ and $\Delta P = 0.01$.
A: wPCFs generated using weighting functions of the form $w(P, p) = 1 - m|P - p|$, together with $w(0.5, p)$
B: wPCFs generated using weighting functions of the form $w(P, p) = 1 - m(P - p)^2$, together with $w(0.5, p)$

wPCF which varies more smoothly with P. In all cases, Fig 17 shows the resulting wPCF alongside $w_p(0.5, p_i)$.

Fig 17 shows that the shape of the weighting function plays a key role in

determining the balance between signal and noise in the wPCF. When the support of the weighting function is broad (e.g., $\Delta P \geq 1$), the wPCF does not identify correlations between r and the target phenotype P. On the other hand, when the non-zero part of the weighting function is very compact (e.g., $\Delta P = 0.01$ for the triangular weighting function) the resulting wPCF identifies correlation well but contains a lot of noise. This suggests that as long as the support of w_p is chosen appropriately then correlations will be correctly identified through the wPCF, regardless of the precise shape of the kernel.

We conclude that the choice of weighting function can have a strong effect on the resulting wPCF, and should be chosen with care. Selecting a weighting function which is too 'narrow' will result in noisy wPCFs, while an unsuitably 'broad' choice will not produce wPCFs with a high enough resolution to identify key features. The appropriate choice is likely to depend on both the distribution of labels in the data and the number of points available, in the same way that the selection of an appropriate annulus radius for the PCF must be tailored to the dataset in question.

S5 Appendix: wPCF for comparing two continuous labels

As discussed in S3 Appendix: Derivation of PCFs and Cross-PCFs from wPCF, the wPCF can be written in a more general form to account for relationships between points with two continuous marks (rather than a continuous mark and discrete mark, as considered in the main text). In its most general form, the PCF is given by

$$wPCF(r, U, V) = \frac{1}{W_U W_V} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{A}{A_{r_k}(\mathbf{x}_i)} w_u(U, u_i) w_v(V, v_j) I_{[r_k, r_{k+1})}(|\mathbf{x}_i - \mathbf{x}_j|) \quad (39)$$

where w_u and w_v are appropriately chosen weighting functions, and $W_U = \sum_i w_u(U, u_i)$ and $W_V = \sum_i w_v(V, v_i)$ are the total weights associated with each mark.

To demonstrate this, we consider the synthetic point pattern in Fig 18A. There are two different types of points here: 200 circles with a continuous label $p \in [0, 1]$, and 200 triangles with a continuous label $\psi \in [0, 10]$. These ranges are chosen to demonstrate the ability of the wPCF to analyse marks which vary over different ranges. The location (x_i, y_i) of each point is chosen randomly, but the labels are assigned according to the formulae $p_i = y_i$ and $\psi_i = 10(1 - y_i)$ for each point type.

We therefore expect to see colocalisation between triangles with label ψ and circles with label $p = 1 - 0.1\psi$. More generally, we expect to observe strong correlation at distance r between triangles with a target label Ψ and circles with a target label $P = 1 - 0.1\Psi \pm r$.

Since we must now specify two target labels, $p = P$ and $\psi = \Psi$, and a radius, r, $wPCF(r, P, \Psi)$ is a three dimensional statistic. This can be most easily visualised by considering fixed values of r, and observing which values of P and Ψ lead to higher or lower values of the wPCF. Figs 18B-D show $wPCF(r, P, \Psi)$ for $r = 0, 0.25, 0.5$. The dotted black lines in each panel represent the lines $P = 1 - 0.1\Psi \pm r$, and show that the wPCF successfully describes the relationships between points with label P and Ψ separated by distance r.

In this example, we have chosen $w_p(P, p_i) = \max\left(1 - \frac{|P - p_i|}{\Delta P}, 0\right)$ with $\Delta P = 0.1$, and $w_\psi(\Psi, \psi_i) = \max\left(1 - \frac{|\Psi - \psi_i|}{\Delta \Psi}, 0\right)$ with $\Delta \Psi = 1$, to reflect that the range of ψ is 10 times larger than that of p.

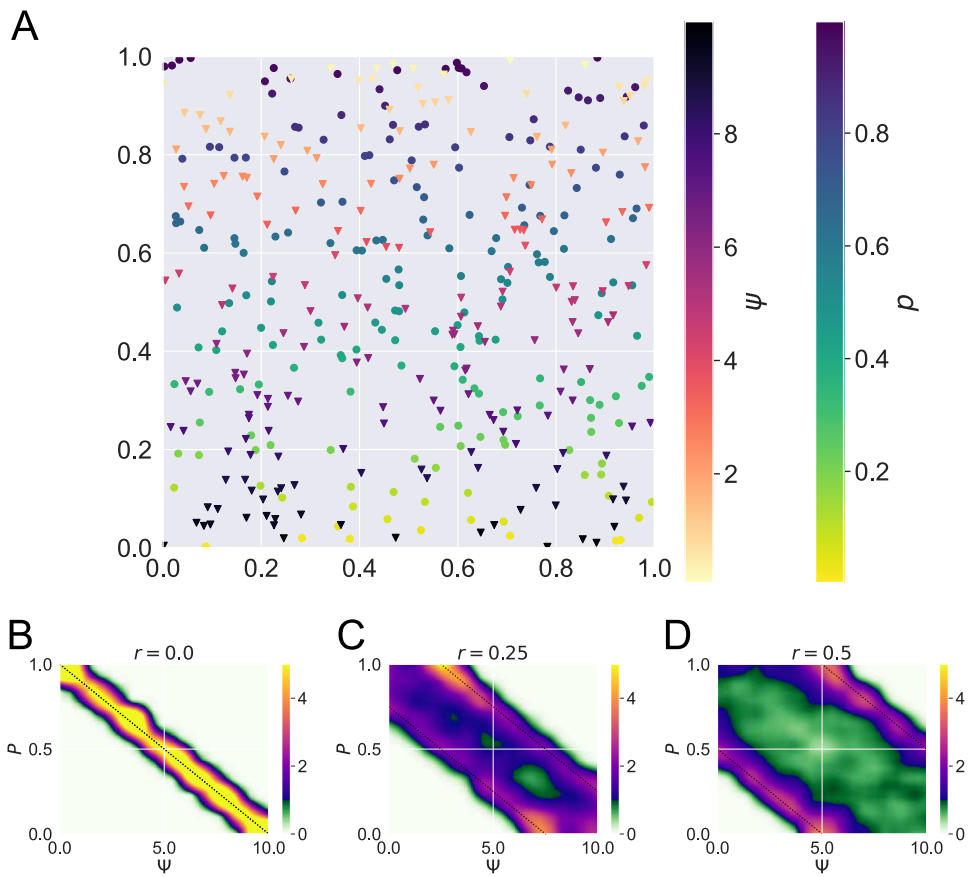

Fig 18. wPCF comparing two continuous labels
A: Two types of points (circles with label p and triangles with label ψ) distributed at random. Points are labelled according to their y coordinate, with point i having $p_i = y_i$ or $\psi_i = 10(1 - y_i)$.
B-D: $wPCF(r, P, \Psi)$ for B) $r = 0$, C) $r = 0.25$, D) $r = 0.5$. Dotted black lines show the expected maximal values of the wPCF according to the construction of the point pattern.

S6 Appendix: cross-PCFs of tumour cells to macrophages

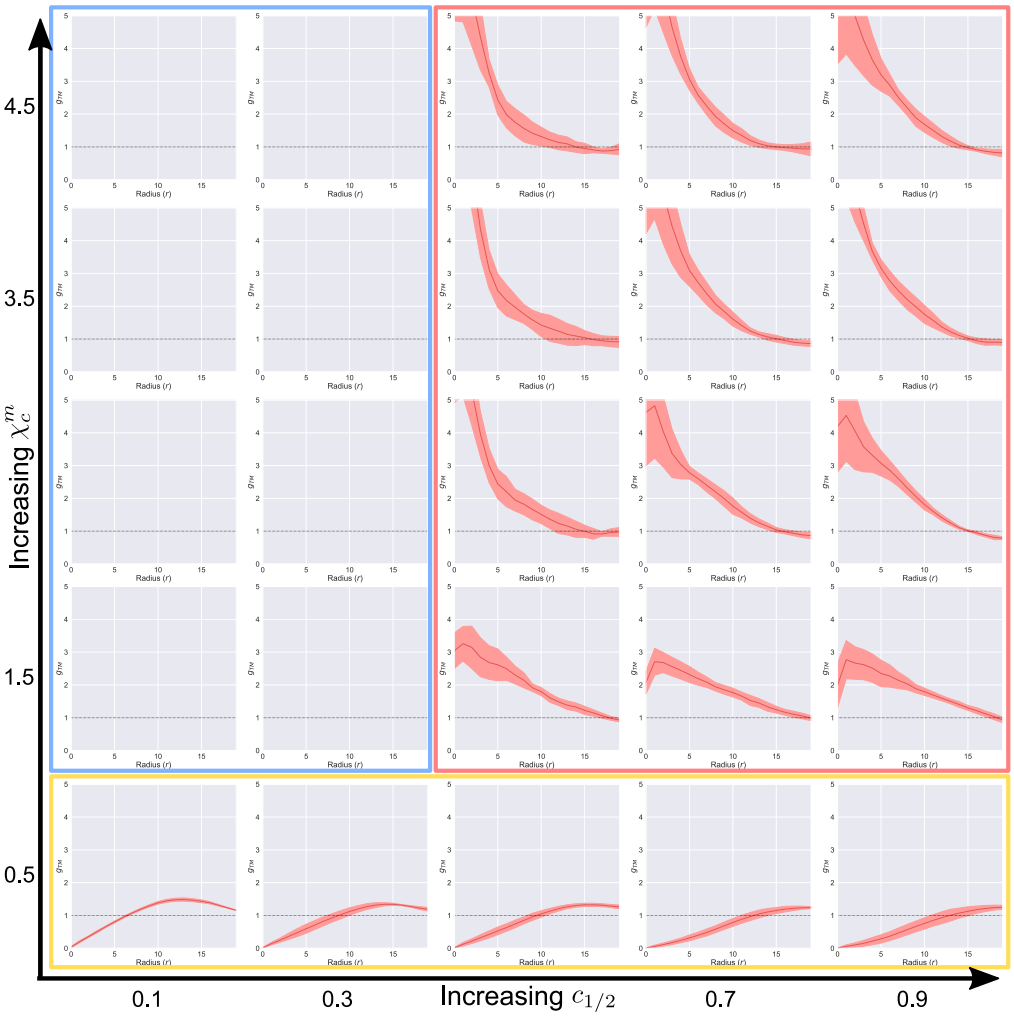

Fig 19. g_{TM} for simulations from the 2 parameter sweep
$g_{TM}(r)$ corresponding to the simulations in Fig 6 (mean and SD of 10 stochastic repetitions). g_{TM} is undefined for ELIM simulations (blue box) where the tumour is eliminated. In COMP simulations (yellow box) there is short range exclusion of macrophages from tumour cells. DIFF simulations (red box) show correlation between macrophages and tumour cells at most lengthscales, especially short range.

Fig 19 shows $g_{TM}(r)$ for the simulations discussed in Fig 6. While the three different qualitative behaviours (DIFF - red box; COMP - yellow box; ELIM - blue box) have different signatures, $g_{TM}(r)$ does not distinguish between different macrophage phenotypes and therefore cannot be used to identify variations in localisation of macrophages with different phenotypes.

S7 Appendix: Partial correlations in 6-parameter sweep

In Table 5 we list all of the partial rank correlation coefficients used to generate Table 2 in the main text. Interpretations of each statistic are included in Table 6.

Modality	Scalar statistic	$P_\star$	$c_{1/2}$	χ_ξ^m	χ_c^m	χ_ϵ^T	g_{crit}
Cell counts	N_M	0.87	-0.32	-0.03	-0.26	-0.09	-0.14
	N_{TC}	-0.78	0.40	0.02	-0.37	0.03	-0.12
Single-cell	$\bar{p}$	-0.47	0.25	-0.01	0.48	-0.01	-0.64
	ϕ_1	0.43	-0.21	0.00	-0.48	0.01	0.64
	ϕ_2	-0.43	0.21	-0.00	0.48	-0.01	-0.64
	N_{M_1}	0.86	-0.35	-0.03	-0.37	-0.07	0.16
	N_{M_2}	-0.16	0.14	-0.04	0.49	-0.04	-0.70
IHC	$\int g_{TM} dr$	0.52	-0.39	-0.10	0.62	0.03	0.11
	$\max(g_{TM})$	0.58	-0.40	-0.02	0.64	0.03	0.07
	$\min(g_{TM})$	0.54	-0.37	0.06	0.40	-0.03	-0.05
	$\int g_{TB} dr$	0.56	-0.39	0.04	0.57	-0.00	-0.01
	$\max(g_{TB})$	0.65	-0.38	-0.01	0.17	0.06	0.08
	$\min(g_{TB})$	0.72	-0.44	-0.01	0.53	0.03	0.02
	$\int g_{BM} dr$	-0.30	0.17	0.24	-0.37	0.04	-0.43
	$\max(g_{BM})$	-0.32	0.13	0.34	0.05	-0.01	-0.31
	$\min(g_{BM})$	-0.12	0.16	0.18	-0.59	0.06	-0.34
	mean M-T	0.43	-0.00	0.13	-0.81	0.01	0.17
	mean T-M	-0.58	0.09	0.09	-0.44	0.11	0.04
	mean T-BV	-0.11	0.12	-0.02	-0.68	0.10	0.13
	min T-BV	0.29	-0.04	0.07	-0.61	-0.03	0.46
	mean BV-T	0.56	-0.01	0.03	-0.39	-0.07	0.33
	min BV-T	0.29	-0.04	0.07	-0.61	-0.03	0.46
IMC	mean M_1-T	0.49	0.04	-0.03	-0.83	-0.03	0.18
	mean M_2-T	0.27	-0.18	0.42	-0.32	0.01	0.11
	mean T-M_1	-0.52	0.11	-0.03	-0.31	0.08	-0.15
	mean T-M_2	-0.16	0.06	0.58	-0.30	0.02	0.48
	$\int wPCF(r=0,p,B)dp$ (above 1)	-0.03	0.01	0.53	0.05	-0.03	-0.32
	$\int wPCF(r=0,p,B)dp$ (below 1)	-0.16	0.05	-0.54	0.36	-0.06	-0.25
	$\int wPCF(r=0,p,T)dp$ (above 1)	-0.46	0.34	-0.09	0.15	0.01	-0.17
	$\int wPCF(r=0,p,T)dp$ (below 1)	-0.23	0.24	0.12	-0.54	-0.02	0.05
	$\int\int wPCF(r,p,B)dpdr$ (above 1)	0.30	-0.25	0.63	0.23	-0.08	0.13
	$\int\int wPCF(r,p,B)dpdr$ (below 1)	0.45	-0.19	0.56	0.03	-0.02	-0.24
	$\int\int wPCF(r,p,T)dpdr$ (above 1)	-0.67	0.42	-0.26	-0.01	0.05	-0.29
	$\int\int wPCF(r,p,T)dpdr$ (below 1)	0.60	-0.47	0.19	0.10	0.02	0.24

Table 5. Partial rank correlation coefficients (PRCCs) Full table of PRCCs corresponding to selected scalar statistics and parameters, based on simulations in which 6 parameters were randomly varied ($n = 432$). Selected PRCCs are presented in the main text in Table 2.

Modality	Scalar statistic	Interpretation
Cell counts	N_M	Number of macrophages
	N_{TC}	Number of tumour cells
Single-cell	$\bar{p}$	Mean macrophage phenotype
	ϕ_1	Proportion of macrophages with $p \leq 0.5$
	ϕ_2	Proportion of macrophages with $p > 0.5$
	N_{M_1}	Number of macrophages with $p \leq 0.5$
	N_{M_2}	Number of macrophages with $p > 0.5$
IHC	$\int g_{TM} dr$	Integral of g_{TM}
	$\max(g_{TM})$	Maximum value of g_{TM}
	$\min(g_{TM})$	Minimum value of g_{TM}
	$\int g_{TB} dr$	Integral of g_{TB}
	$\max(g_{TB})$	Maximum value of g_{TB}
	$\min(g_{TB})$	Minimum value of g_{TB}
	$\int g_{BM} dr$	Integral of g_{BM}
	$\max(g_{BM})$	Maximum value of g_{BM}
	$\min(g_{BM})$	Minimum value of g_{BM}
	mean M-T	Mean distance from macrophage to closest tumour cell
	mean T-M	Mean distance from tumour cell to closest macrophage
	mean T-BV	Mean distance from tumour cell to closest vessel
	min T-BV	Min distance from tumour cell to closest vessel
	mean BV-T	Mean distance from vessel to closest tumour cell
	min BV-T	Min distance from vessel to closest tumour cell
IMC	mean M_1-T	Mean distance from M_1 macrophage to closest tumour cell
	mean M_2-T	Mean distance from M_2 macrophage to closest tumour cell
	mean T-M_1	Mean distance from tumour cell to closest M_1 macrophage
	mean T-M_2	Mean distance from tumour cell to closest M_2 macrophage
	$\int wPCF(r=0,p,B)dp$ (above 1)	Area between $\max(wPCF(r=0,p,B),1)$ and $wPCF=1$
	$\int wPCF(r=0,p,B)dp$ (below 1)	Area between $\min(wPCF(r=0,p,B),1)$ and $wPCF=0$
	$\int wPCF(r=0,p,T)dp$ (above 1)	Area between $\max(wPCF(r=0,p,T),1)$ and $wPCF=1$
	$\int wPCF(r=0,p,T)dp$ (below 1)	Area between $\min(wPCF(r=0,p,T),1)$ and $wPCF=0$
	$\int\int wPCF(r,p,B)dpdr$ (above 1)	Volume between $\max(wPCF(r,p,B),1)$ and $wPCF=1$
	$\int\int wPCF(r,p,B)dpdr$ (below 1)	Volume between $\min(wPCF(r,p,B),1)$ and $wPCF=0$
	$\int\int wPCF(r,p,T)dpdr$ (above 1)	Volume between $\max(wPCF(r,p,T),1)$ and $wPCF=1$
	$\int\int wPCF(r,p,T)dpdr$ (below 1)	Volume between $\min(wPCF(r,p,T),1)$ and $wPCF=0$

Table 6. Interpretations of statistics Interpretation of the statistics featured in Table 5.

References

1. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, Angell H, Fredriksen T, Lafontaine L, Berger A, Bruneval P, Fridman WH, Becker C, Pagès F, Speicher MR, Trajanoski Z, Galon J. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity. 2013;39(4):782–795. doi:10.1016/j.immuni.2013.10.003.

2. Singleton DC, Macann A, Wilson WR. Therapeutic targeting of the hypoxic tumour microenvironment. Nature Reviews Clinical Oncology. 2021;18(12):751–772. doi:10.1038/s41571-021-00539-4.

3. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, Nair VS, Xu Y, Khuong A, Hoang CD, Diehn M, West RB, Plevritis SK, Alizadeh AA. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nature Medicine. 2015;21(8):938–945. doi:10.1038/nm.3909.

4. Cortese N, Carriero R, Laghi L, Mantovani A, Marchesi F. Prognostic significance of tumor-associated macrophages: past, present and future. Seminars in Immunology. 2020;48:101408. doi:10.1016/j.smim.2020.101408.

5. Mantovani A, Marchesi F, Malesci A, Laghi L, Allaven P. Tumour-associated macrophages as treatment targets in oncology. Nature Reviews Clinical Oncology. 2017;14(7):399–416. doi:10.1038/nrclinonc.2016.217.

6. Li Z, Maeda D, Yoshida M, Umakoshi M, Nanjo H, Shiraishi K, et al. The intratumoral distribution influences the prognostic impact of CD68- and CD204-positive macrophages in non-small cell lung cancer. Lung Cancer. 2018;123:127–135. doi:10.1016/j.lungcan.2018.07.015.

7. Rakaee M, Rasmussen Busund L-T, Jamaly S, Paulsen E-E, Richardsen E, et al. Prognostic Value of Macrophage Phenotypes in Resectable Non-Small Cell Lung Cancer Assessed By Multiplex Immunohistochemistry. Neoplasia. 2019;21(3):282–293. doi:10.1016/j.neo.2019.01.005.

8. Pinto ML, Rios E, Durães C, Ribeiro R, Machado JC, Mantovani A, Barbosa MA, Carneiro F, Oliveira MJ. The Two Faces of Tumor-Associated Macrophages and Their Clinical Significance in Colorectal Cancer. Frontiers in Immunology. 2019;10:1–12. doi:10.3389/fimmu.2019.01875.

9. López-Janeiro Á, Padilla-Ansala C, de Andrea CE, Hardisson D, Melero I. Prognostic value of macrophage polarization markers in epithelial neoplasms and melanoma. A systematic review and meta-analysis. Modern Pathology. 2020;33(8):1458–1465. doi:10.1038/s41379-020-0534-z.

10. Jackute J, Zemaitis M, Pranys D, Sitkauskiene B, Miliauskas S, Vaitkiene S, Sakalauskas R. Distribution of M1 and M2 macrophages in tumor islets and stroma in relation to prognosis of non-small cell lung cancer. BMC Immunology. 2018;19(3):1–13. doi:10.1186/s12865-018-0241-4.

11. Murray PJ, Allen JE, Biswas SK, Fisher EA, Gilroy DW, Goerdt S, Gordon S, Hamilton JA, Ivashkiv LB, Lawrence T, Locati M, Mantovani A, Martinez FO, Mege J-L, Mosser DM, Natoli G, Saeij JP, Schultze JL, Shirey KA, Sica A, Suttles J, Udalova I, van Ginderachter JA, Vogel SN, Wynn TA. Macrophage Activation and Polarization: Nomenclature and Experimental Guidelines. Immunity. 2014;41(1):14–20. doi:10.1016/j.immuni.2014.06.008.

12. Huang YK, Wang M, Sun Y, Di Costanzo N, Mitchell C, Achuthan A, Hamilton JA, Busuttil RA, Boussioutas A. Macrophage spatial heterogeneity in gastric cancer defined by multiplex immunohistochemistry. Nature Communications. 2019;10(1):1–15. doi:10.1038/s41467-019-11788-4.

13. Giesen C, Wang HAO, Schapiro D, Zivanovic N, Jacobs A, Hattendorf B, Schüffler PJ, Golimund D, Buhmann JM, Brandt S, Varga Z, Wild PJ, Günther D, Bodenmiller B Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. Nature Methods. 2014;11(4):417–422. doi:10.1038/nmeth.2869.

14. Kuett L, Catena R, Özcan A, Plüss A, Cancer Grand Challenges IMAXT Consortium, Schraml P, Moch H, de Souza N, Bodenmiller B. Three-dimensional imaging mass cytometry for highly multiplexed molecular and cellular mapping of tissues and the tumor microenvironment. Nature Cancer. 2022;3(1):122–133. doi:10.1038/s43018-021-00301-w.

15. Nagl S, Haas M, Lahmer G, Büttner-Herold M, Grabenbauer GG, Fietkau R, Distel LV. Cell-to-cell distances between tumor-infiltrating inflammatory cells have the potential to distinguish functionally active from suppressed inflammatory cells. Oncoimmunology. 2016;5(5):e1127494. doi:10.1080/2162402X.2015.1127494.

16. Nawaz S, Trahearn NA, Heindl A, Banerjee S, Maley CC, Sottoriva A, Yuan Y. Analysis of tumour ecological balance reveals resource-dependent adaptive strategies of ovarian cancer. EBioMedicine. 2019;48:224–235. doi:10.1016/j.ebiom.2019.10.001.

17. Yuan Y. Spatial Heterogeneity in the Tumor Microenvironment. Cold Spring Harbor Perspectives in Medicine. 2016;6(8):1–17. doi:10.1101/cshperspect.a026583.

18. Yuan Y. Modelling the spatial heterogeneity and molecular correlates of lymphocytic infiltration in triple-negative breast cancer. Journal of the Royal Society Interface. 2015;12:20141153. doi:10.1098/rsif.2014.1153.

19. Khan AM, Yuan Y. Biopsy variability of lymphocytic infiltration in breast cancer subtypes and the ImmunoSkew score. Scientific Reports. 2016;6:36231. doi:10.1038/srep36231.

20. Sailem HZ, Al Haj Zen A. Morphological landscape of endothelial cell networks reveals a functional role of glutamate receptors in angiogenesis. Scientific Reports. 2020;10:13829. doi:10.1038/s41598-020-70440-0.

21. Failmezger H, Muralidhar S, Rullan A, de Andrea CE, Sahai E, Yuan Y. Topological Tumor Graphs: A Graph-Based Spatial Model to Infer Stromal Recruitment for Immunosuppression in Melanoma Histology. Cancer Research. 2020;256(4):1199–1209. doi:10.1158/0008-5472.CAN-19-2268.

22. Bull JA, Macklin PS, Quaiser T, Braun F, Waters SL, Pugh CW, Byrne HM. Combining multiple spatial statistics enhances the description of immune cell localisation within tumours. Scientific Reports. 2020;10:18624. doi:10.1038/s41598-020-75180-9.

23. Arwert EN, Harney AS, Entenberg D, Wang Y, Sahai E, Pollard JW, et al. A Unidirectional Transition from Migratory to Perivascular Macrophage Is Required for Tumor Cell Intravasation. Cell Reports. 2018;23(5):1239–1248. doi:10.1016/j.celrep.2018.04.007.

24. Harney AS, Arwert EN, Entenberg D, Wang Y, Guo P, Qian BZ, et al. Real-Time Imaging Reveals Local, Transient Vascular Permeability, and Tumor Cell Intravasation Stimulated by TIE2hi Macrophage-Derived VEGFA. Cancer Discovery. 2015;5(9):932–943. doi:10.1158/2159-8290.CD-15-0012.

25. Klowss JJ, Browning AP, Murphy RJ, Carr EJ, Plank MJ, Gunasingh G, Haass NK, Simpson MJ. A stochastic mathematical model of 4D tumour spheroids with real-time fluorescent cell cycle labelling. Journal of the Royal Society Interface. 2022;19:20210903. doi:10.1098/rsif.2021.0903.

26. Ghaffarizadeh A, Heiland R, Friedman SH, Shannon MM, Macklin P. PhysiCell: An open source physics-based cell simulator for 3-D multicellular systems. PLoS Computational Biology. 2018;14(2):1–34. doi:10.1371/journal.pcbi.1005991.

27. Bravo RR, Baratchart E, West J, Schenck RO, Miller AK, Gallaher J, Gatenbee CD, Basanta D, Robertson-Tessi R, Anderson ARA. Hybrid Automata Library: A flexible platform for hybrid modeling with real-time visualization. PLoS Computational Biology. 2020;16(3):1–28. doi:10.1371/journal.pcbi.1007635.

28. Leschiera E, Lorenzi T, Shen S, Almeida L, Audebert C. A mathematical model to study the impact of intra-tumour heterogeneity on anti-tumour CD8+ T cell immune response. Journal of Theoretical Biology. 2022;538:111028. doi:10.1016/j.jtbi.2022.111028.

29. Swat MH, Thomas GL, Belmonte JM, Shirinifard A, Hmeljak D, Glazier JA. Multi-Scale Modeling of Tissues Using CompuCell3D. Methods in Cell Biology. 2012;110:325–366. doi:10.1016/B978-0-12-388403-9.00013-8.

30. Cooper F, Baker R, Bernabeu M, Bordas R, Bowler L, Bueno-Orovio A, et al. Chaste: Cancer, Heart and Soft Tissue Environment. Journal of Open Source Software. 2020;5(47):1848. doi:10.21105/joss.01848.

31. Mirams GR, Arthurs CJ, Bernabeu MO, Bordas R, Cooper J, Corrias A, et al. Chaste: An Open Source C++ Library for Computational Physiology and Biology. PLoS Computational Biology. 2013;9(3). doi:10.1371/journal.pcbi.1002970.

32. Pitt-Francis J, Pathmanathan P, Bernabeu MO, Bordas R, Cooper J, Fletcher AG, et al. Chaste: A test-driven approach to software development for biological modelling. Computer Physics Communications. 2009;180(12):2452–2471. doi:10.1016/j.cpc.2009.07.019.

33. Scott JG, Fletcher AG, Anderson ARA, Maini PK. Spatial Metrics of Tumour Vascular Organisation Predict Radiation Efficacy in a Computational Model. PLoS Computational Biology. 2016;12(1). doi:10.1371/journal.pcbi.1004712.

34. Vipond O, Bull JA, Macklin PS, Tillmann U, Pugh CW, Byrne HM, Harrington HA. Multiparameter persistent homology landscapes identify immune cell spatial patterns in tumors. Proceedings of the National Academy of Sciences. 2021;118(41):e2102166118. doi:10.1073/pnas.2102166118.

35. Bull JA, Mech F, Quaiser T, Waters SL, Byrne HM. Mathematical modelling reveals cellular dynamics within tumour spheroids. PLOS Computational Biology. 2020;16(8):e1007961. doi:10.1371/journal.pcbi.1007961.

36. Rockne RC, Hawkins-Daruud A, Swanson KR, Sluka JP, Glazier JA, Macklin P, Hormuth II DA, Jarrett AM, Lima EABF, Oden JT, Biros G, Yankeelov TE, Curtius K, Al Bakir I, Wodarz D, Komarova N, Aparicio L, Bordyuh M, Rabadan R, Finley SD, Enderling H, Caudell J, Moros EG, Anderson ARA, Gatenby RA, Kaznatcheev A, Jeavons P, Krishnan N, Pelesko J, Wadhwa RR, Yoon N, Nichol D, Marusyk A, Hinczewski M, Scott JG. The 2019 mathematical oncology roadmap. Physical Biology. 2019;16:041005. doi:10.1088/1478-3975/ab1a09.

37. Bull JA, Byrne HM. The Hallmarks of Mathematical Oncology. Proceedings of the IEEE. 2022. In press. doi:10.1109/JPROC.2021.3136715.

38. Owen MR, Sherratt JA. Mathematical modelling of macrophage dynamics in tumours. Mathematical Models and Methods in Applied Sciences. 1999;9(4):513–539. doi:10.1142/S0218202599000270.

39. Kelly CE, Leek RD, Byrne HM, Cox SM, Harris AL, Lewis CE. Modelling Macrophage Infiltration into Avascular Tumours. Journal of Theoretical Medicine. 2002;4(1):21–38. doi:10.1080/10273660290015242.

40. Owen MR, Byrne HM, Lewis CE. Mathematical modelling of the use of macrophages as vehicles for drug delivery to hypoxic tumour sites. Journal of Theoretical Biology. 2004;226(4):377–391. doi:10.1016/j.jtbi.2003.09.004.

41. Norton KA, Jin K, Popel AS. Modeling triple-negative breast cancer heterogeneity: Effects of stromal macrophages, fibroblasts and tumor vasculature. Journal of Theoretical Biology. 2018;452:56–68. doi:10.1016/j.jtbi.2018.05.003.

42. Suveges S, Eftimie R, Trucu D. Directionality of Macrophages Movement in Tumour Invasion: A Multiscale Moving-Boundary Approach. Bulletin of Mathematical Biology. 2020;82(12):1–50. doi:10.1007/s11538-020-00819-7.

43. Li X, Jolly MK, George JT, Pienta KJ, Levine H. Computational modeling of the crosstalk between macrophage polarization and tumor cell plasticity in the tumor microenvironment. Frontiers in Oncology. 2019;9(JAN):1–12. doi:10.3389/fonc.2019.00010.

44. Mahlbacher G, Curtis LT, Lowengrub J, Frieboes HB. Mathematical modeling of tumor-associated macrophage interactions with the cancer microenvironment. Journal for ImmunoTherapy of Cancer. 2018;6(1):1–17. doi:10.1186/s40425-017-0313-7.

45. Webb SD, Owen MR, Byrne HM, Murdoch C, Lewis CE. Macrophage-based anti-cancer therapy: Modelling different modes of tumour targeting. Bulletin of Mathematical Biology. 2007;69(5):1747–1776. doi:10.1007/s11538-006-9189-2.

46. den Breems NY, Eftimie R. The re-polarisation of M2 and M1 macrophages and its role on cancer outcomes. Journal of Theoretical Biology. 2016;390:23–39. doi:10.1016/j.jtbi.2015.10.034.

47. Cess CG, Finley SD. Multi-scale modeling of macrophage—T cell interactions within the tumor microenvironment. vol. 16; 2020. Available from: `http://dx.doi.org/10.1371/journal.pcbi.1008519`.

48. Curtis LT, Sebens S, Frieboes HB. Modeling of tumor response to macrophage and T lymphocyte interactions in the liver metastatic microenvironment. Cancer Immunology, Immunotherapy. 2021;70(5):1475–1488. doi:10.1007/s00262-020-02785-4.

49. Eftimie R. Investigation into the role of macrophages heterogeneity on solid tumour aggregations. Mathematical Biosciences. 2020;322(March 2019):108325. doi:10.1016/j.mbs.2020.108325.

50. El-Kenawi A, Gatenbee C, Robertson-Tessi M, Bravo R, Dhillon J, et al. Acidity promotes tumour progression by altering macrophage phenotype in prostate cancer. British Journal of Cancer. 2019;121(7):556–566. doi:10.1038/s41416-019-0542-2.

51. Elitas M, Zeinali S. Modeling and Simulation of EGF-CSF-1 pathway to Investigate Glioma - Macrophage Interaction in Brain Tumors. International Journal of Cancer Studies & Research (IJCR). 2016; p. 1–8.

52. Knútsdóttir H, Pálsson E, Edelstein-Keshet L. Mathematical model of macrophage-facilitated breast cancer cells invasion. Journal of Theoretical Biology. 2014;357:184–199. doi:10.1016/j.jtbi.2014.04.031.

53. Knútsdóttir H, Condeelis JS, Pálsson E. 3-D individual cell based computational modeling of tumor cell–macrophage paracrine signaling mediated by EGF and CSF-1 gradients. Integrative Biology. 2016;8(1):104–119. doi:10.1039/C5IB00201J.

54. Binder BJ, Simpson MJ. Quantifying spatial structure in experimental observations and agent-based simulations using pair-correlation functions. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics. 2013;88(2):1–10. doi:10.1103/PhysRevE.88.022705.

55. Agnew DJG, Green JEF, Brown TM, Simpson MJ, Binder BJ. Distinguishing between mechanisms of cell aggregation using pair-correlation functions. Journal of Theoretical Biology. 2014;352:16–23. doi:10.1016/j.jtbi.2014.02.033.

56. Gavagnin E, Owen JP, Yates CA. Pair correlation functions for identifying spatial correlation in discrete domains. Physical Review E. 2018;97(6). doi:10.1103/PhysRevE.97.062104.

57. Johnston ST, Crampin EJ. Corrected pair correlation functions for environments with obstacles. Physical Review E. 2019;99(3):1–19. doi:10.1103/PhysRevE.99.032124.

58. Browning AP, McCue SW, Binny RN, Plank MJ, Shah ET, Simpson MJ. Inferring parameters for a lattice-free model of cell migration and proliferation using experimental data. Journal of Theoretical Biology. 2018;437:251–260. doi:10.1016/j.jtbi.2017.10.032.

59. Fozard JA, Kirkham GR, Buttery LDK, King JR, Jensen OE, Byrne HM. Techniques for analysing pattern formation in populations of stem cells and their progeny. BMC Bioinformatics. 2011;12:396. doi:10.1186/1471-2105-12-396.

60. Dini S, Binder BJ, Green JEF. Understanding interactions between populations: Individual based modelling and quantification using pair correlation functions. Journal of Theoretical Biology. 2018;439:50–64. doi:S0022519317305210.

61. Ben-Said M. Spatial point-pattern analysis as a powerful tool in identifying pattern-process relationships in plant ecology: an updated review. Ecological Processes. 2021;10(56). doi:10.1186/s13717-021-00314-4.

62. Szmyt J. Spatial statistics in ecological analysis: from indices to functions. Silva Fennica. 2014;48(1). doi:10.14214/sf.1008.

63. Wälder O, Stoyan D. On Variograms in Point Process Statistics. Biometrical Journal. 1996;38(8):895–905. doi:10.1002/bimj.4710380802.

64. Beisbart C, Kerscher M, Mecke K. Mark Correlations: Relating Physical Properties to Spatial Distributions. In: Mecke K, Stoyan D, editors. Morphology of Condensed Matter. Springer, Berlin, Heidelberg; 2002. pp. 358-390. doi:10.1007/3-540-45782-8_15.

65. Beisbart C, Kerscher M. Luminosity- and morphology-depdendent clustering of galaxies. The Astrophysical Journal. 2000;545:6–25. doi:10.1086/317788.

66. Stoyan D, Stoyan H. Fractals, Random Shapes and Point Fields: Methods of Geometrical Statistics. Chichester, Wiley; 1994.

67. Gavrikov V, Stoyan D. The use of marked point processes in ecological and environmental forest studies. Environmental and Ecological Statistics. 1995;2(4):331–344. doi:10.1007/BF00569362.

68. Stoyan D, Wälder O. On Variograms in Point Process Statistics, II: Models of Markings and Ecological Interpretation. Biometrical Journal. 2000;42(2):171–187. doi:10.1002/(SICI)1521-4036(200005)42:2<171::AID-BIMJ171>3.0.CO;2-L.

69. Wyckoff J, Wang W, Lin EY, Wang Y, Pixley F, Stanley ER, et al. A paracrine loop between tumor cells and macrophages is required for tumor cell migration in mammary tumors. Cancer Research. 2004; p. 7022–7029. doi:10.1158/0008-5472.CAN-04-1449.

70. Wyckoff JB, Wang Y, Lin EY, Li JF, Goswami S, Stanley ER, et al. Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors. Cancer Research. 2007;67(6):2649–2656. doi:10.1158/0008-5472.CAN-06-1823.

71. Taylor S, Noble R. HTML5 PivotViewer: high-throughput visualization and querying of image data on the web. Bioimage informatics. 2014;30(18):2691–2692. doi:10.1093/bioinformatics/btu349.

72. Marino S, Hogue IB, Ray CJ, Kirschner DE. A methodology for performing global uncertainty and sensitivity analysis in systems biology. Journal of Theoretical Biology. 2008;254(1):178–196. doi:10.1016/j.jtbi.2008.04.011.

73. Cassetta L, Pollard JW. Targeting macrophages: Therapeutic approaches in cancer. Nature Reviews Drug Discovery. 2018;17(12):887–904. doi:10.1038/nrd.2018.169.

74. Zhou J, Tang Z, Gao S, Li C, Feng Y, Zhou X. Tumor-Associated Macrophages: Recent Insights and Therapies. Frontiers in Oncology. 2020;10(February):1–13. doi:10.3389/fonc.2020.00188.

75. Jia XH, Du Y, Mao D, Wang ZL, He ZQ, Qiu JD, et al. Zoledronic acid prevents the tumor-promoting effects of mesenchymal stem cells via MCP-1 dependent recruitment of macrophages. Oncotarget. 2015;6(28):26018–26028. doi:10.18632/oncotarget.4658.

76. Comito G, Segura CP, Taddei ML, Lanciotti M, Serni S, Morandi A, et al. Zoledronic acid impairs stromal reactivity by inhibiting M2-macrophages polarization and prostate cancer-associated fibroblasts. Oncotarget. 2017;8(1):118–132. doi:10.18632/oncotarget.9497.

77. Borghese C, Cattaruzza L, Pivetta E, Normanno N, De Luca A, Mazzucato M, et al. Gefitinib inhibits the cross-talk between mesenchymal stem cells and prostate cancer cells leading to tumor cell proliferation and inhibition of docetaxel activity. Journal of Cellular Biochemistry. 2013;114(5):1135–1144. doi:10.1002/jcb.24456.

78. Cannarile MA, Ries CH, Hoves S, Rüttinger D. Targeting tumor-associated macrophages in cancer therapy and understanding their complexity. Oncoimmunology. 2014;3(9):e955356. doi:10.4161/21624011.2014.955356.

79. Ries CH, Hoves S, Cannarile MA, Rüttinger D. CSF-1/CSF-1R targeting agents in clinical development for cancer therapy. Current Opinion in Pharmacology. 2015;23:45–51. doi:10.1016/j.coph.2015.05.008.

80. Greenspan HP. Models for the Growth of a Solid Tumor by Diffusion. Studies in Applied Mathematics. 1972;51(4):317–340. doi:10.1002/sapm1972514317.

81. Pathmanathan P, Cooper J, Fletcher A, Mirams G, Murray P, Osborne JM, et al. A computational study of discrete mechanical tissue models. Physical biology. 2009;6(3):036001. doi:10.1088/1478-3975/6/3/036001.

82. Meineke FA, Potten CS, Loeffler M. Cell migration and organization in the intestinal crypt using a lattice-free model. Cell Proliferation. 2001;34(4):253–266. doi:10.1046/j.0960-7722.2001.00216.x.

83. Osborne JM, Fletcher AG, Pitt-Francis JM, Maini PK, Gavaghan DJ. Comparing individual-based approaches to modelling the self-organization of multicellular tissues. PLoS Computational Biology. 2017;13(2):1–34. doi:10.1371/journal.pcbi.1005387.

84. Drasdo D, Höhme S. A single-cell-based model of tumor growth in vitro: monolayers and spheroids. Physical Biology. 2005;2(3):133–147. doi:10.1088/1478-3975/2/3/001.

85. Laget S, Broncy L, Hormigos K, Dhingra DM, BenMohamed F, Capiod T, et al. Technical insights into highly sensitive isolation and molecular characterization of fixed and live circulating tumor cells for early detection of tumor invasion. vol. 12; 2017.

86. Dunn SJ, Näthke IS, Osborne JM. Computational models reveal a passive mechanism for cell migration in the crypt. PLoS ONE. 2013;8(11):1–18. doi:10.1371/journal.pone.0080516.

87. Mueller-Klieser WF, Sutherland RM. Oxygen tensions in multicell spheroids of two cell lines. British Journal of Cancer. 1982;45(2):256–264. doi:10.1038/bjc.1982.41.

88. Grimes DR, Kelly C, Bloch K, Partridge M. A method for estimating the oxygen consumption rate in multicellular tumour spheroids. Journal of the Royal Society Interface. 2014;11(92). doi:10.1098/rsif.2013.1124.

89. Schaller G, Meyer-Hermann M. Multicellular tumor spheroid in an off-lattice Voronoi-Delaunay cell model. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics. 2005;71(5):1–16. doi:10.1103/PhysRevE.71.051910.

90. Colling R, Pitman H, Oien K, Rajpoot N, Macklin P, CM-Path AI in Histopathology Working Group, Snead D, Sackville T, Verrill C. Artificial intelligence in digital pathology: a roadmap to routine use in clinical practice. Journal of Pathology. 2019;249:143–150. doi:10.1002/path.5310.

91. Hakkoum H, Abnane I, Idri A. Interpretability in the medical field: A systematic mapping and review study. Applied Soft Computing. 2022;117:108391. doi:10.1016/j.asoc.2021.108391.

92. Stolz BJ, Kaeppler J, Markelc B, Mech F, Lipsmeier F, Muschel RJ, Byrne HM, Harrington HA. Multiscale Topology Characterises Dynamic Tumour Vascular Networks. arXiv preprint. 2020:2008.08667 doi:10.48550/arXiv.2008.08667.

93. Kingsley JL, Costello JR, Raghunand N, Rejniak KA. Bridging cell-scale simulations and radiologic images to explain short-time intratumoral oxygen fluctuations. PLoS Computational Biology. 2021;17(7):e1009206 doi:10.1371/journal.pcbi.1009206.

94. Vaghi C, Fanciullino R, Benzekry S, Poignard C. Macro-scale models for fluid flow in tumour tissues: impact of microstructure properties. Journal of Mathematical Biology. 2022;84:27. doi:10.1007/s00285-022-01719-1.

95. Bernabeu MO, Köry J, Grogan JA, Markelc B, Beardo A, et al. Abnormal morphology biases hematocrit distribution in tumor vasculature and contributes to heterogeneity in tissue oxygenation. PNAS. 2020;117(45):27811–27819. doi:10.1073/pnas.2007770117.

96. Grogan JA, Connor AJ, Markelc B, Muschel RJ, Maini PK, Byrne HM, Pitt-Francis JM. Microvessel Chaste: An Open Library for Spatial Modeling of Vascularized Tissues. Biophysical Journal. 2017;112(9):1767–1772. doi:10.1016/j.bpj.2017.03.036.

97. Anderson ARA, Chaplain MAJ. Continuous and Discrete Mathematical Models of Tumor-induced Angiogenesis. Bulletin of Mathematical Biology. 1998;60:857–900. doi:10.1006/bulm.1998.0042.

98. Owen MR, Alarcón T, Maini PK, Byrne HM. Angiogenesis and vascular remodelling in normal and cancerous tissues. Journal of Mathematical Biology. 2009;58:689–721. doi:10.1007/s00285-008-0213-z.

99. Alarcón T, Byrne HM, Maini PK. A cellular automaton model for tumour growth in inhomogeneous environment. Journal of Theoretical Biology. 2003;225(2):257–274. doi:10.1016/S0022-5193(03)00244-3.

100. Galon J, Mlecnik B, Bindea G, Andell HK, Berger A, et al. Towards the introduction of the 'Immunoscore' in the classification of malignant tumours. Journal of Pathology. 2014;232(2):199–209. doi:10.1002/path.4287.

101. AbdulJabbar K, Ahmed Raza SE, Rosenthal R, Jamal-Hanjani M, Veeriah S, et al. Geospatial immune variability illuminates differential evolution of lung adenocarcinoma. Nature Medicine. 2020;26(7):1054–1062. doi:10.1038/s41591-020-0900-x.

102. Chang SL, Cavnar SP, Luker KE, Takayama S, Luker GD, Linderman JJ. Cell, Isoform, and Environment Factors Shape Gradients and Modulate Chemotaxis. PLoS ONE. 2015;10(4):e0123450. doi:10.1371/journal.pone.0123450.

103. Albro MB, Nims RJ, Cigan AD, Yeroushalmi KJ, Alliston T, Hung CT, Ateshian GA. Accumulation of Exogenous Activated TGF-β in the Superficial Zone of Articular Cartilage. Biophysical Journal. 2013;104(8):1794–1804. doi:10.1016/j.bpj.2013.02.052.

104. Wakefield LM, Winokur TS, Hollands RS, Christopherson K, Levinson AD, Sporn MB. Recombinant Latent Transforming Growth Factor β1 Has a Longer Plasma Half-Life in Rats than Active Transforming Growth Factor β1, and a Different Tissue Distribution. The Journal of Clinical Investigation. 1990;86(6):1976–1984. doi:10.1172/JCI114932.