Patient-calibrated agent-based modelling of ductal carcinoma in situ (DCIS): Supplementary Material

Paul Macklin [1,2,3,4], Mary E. Edgerton [5],
Alastair M. Thompson [5,6], Vittorio Cristini [4,7]

Abstract

We provide here the following supplementary materials for Macklin et al. (2011):

(1) Sampling of significant agent-based modelling beyond DCIS

(2) Agent model generalisations

(3) Additional mathematical notes and analyses

(4) Expanded detail for patient-independent parameter estimations

(5) Numerical methods

(6) `MultiCellXML`, sample data sets, and C++ source

(7) Additional numerical results

(8) Videos

Extensive resources, including this document, C++ source code, benchmark datasets, and animations are mirrored and maintained at MathCancer.org. [8] In particular, we have posted an introduction to the pathobiology of DCIS (Macklin, 2012), including numerous biological and clinical references.

Email address: `Paul.Macklin@usc.edu` (Paul Macklin).
URL: `http://www.MathCancer.org` (Paul Macklin).
[1] Corresponding author
[2] Center for Applied Molecular Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA
[3] Formerly of: Division of Mathematics, University of Dundee (UoD), Dundee, UK
[4] Formerly of: SBMI, Univ. of Texas Health Science Center, Houston, TX, USA
[5] M.D. Anderson Cancer Center, Houston, TX, USA
[6] Department of Surgery and Molecular Oncology, UoD, Dundee, UK
[7] Depts. of Pathology & Chemical Eng., U. of New Mexico, Albuquerque, NM, USA
[8] `http://www.mathcancer.org/JTB_DCIS_2012/`

1 A sampling of significant agent-based modelling beyond DCIS

We briefly sample significant relevant prior agent-based models, beyond the scope of DCIS modelling emphasised in the main text. For a broader and deeper review, please see Lowengrub et al. (2010); Macklin et al. (2010b); Deisboeck et al. (2011) and the references therein.

Cellular automata methods–which arrange cells on a regular (e.g., Cartesian or hexagonal) lattice with probabilities governing state changes and jumps between lattice points–are efficient for linking molecular- and cellular-scale biology in large numbers of virtual cells. However, they cannot accurately model cell and tissue mechanics due to the limitations they place upon cell arrangement (must be grid-aligned), size (all cells have equal size), velocity (cells move one cell diameter per time step), and interactions (can only interact with up to 8 neighbours in 2D). In particular, proliferation is disallowed in cells that are surrounded by cells in the adjacent computational mesh points; in actual tissue, interior cells can proliferate by deforming and pushing neighbouring cells into non-lattice configurations. We use agent-based modelling (ABM), which eliminates the computational lattice and instead assigns each cell a position that evolves under the influence of forces acting upon it. Note that ABMs are sometimes referred to as individual-based models or particle methods. Alternative approaches include the lattice-gas method (Dormann and Deutsch, 2002), off-lattice cellular automata methods such as Voronoi-Delaunay models (Schaller and Meyer-Hermann, 2005), the immersed boundary cell model (Rejniak, 2007; Rejniak and Dillon, 2007; Rejniak and Anderson, 2008a,b), and the cellular potts technique (a.k.a. Graner-Glazier-Hogeweg model) (Graner and Glazier, 1992; Glazier and Garner, 1993).

An excellent agent-based model was developed by Drasdo, Höhme and coworkers (Drasdo et al., 1995; Drasdo and Höhme, 2003, 2005; Drasdo, 2005). Cells are modelled as roughly spherical, slightly compressible, and capable of migration, growth and division. Cell adhesion and repulsion (from limitations on cell deformation and compressibility) are modelled by introducing an interaction energy; cells respond to proliferation and apoptosis in their neighbours by moving to reduce the total interaction energy using a stochastic algorithm. Ramis-Conde et al. (2008a,b) used a similar agent model, but instead used interaction potential functions to simulate cell-cell mechanics: cells move down the gradient of the potential, analogous to minimizing the interaction energy. Their work included a basic accounting for the cell-cell surface contact area, and related the strength of cell-cell adhesion to the concentration of E-cadherin/β-catenin complexes in the contact regions. Others have modelled cells as deformable viscoelastic ellipsoids (e.g., Palsson and Othmer (2000); Dallon and Othmer (2004)).

Drasdo et al. (1995) initially developed their agent model to study epithelial cell-fibroblast-fibrocyte aggregations in connective tissue. More recently, they applied it to avascular tumour growth (Drasdo and Höhme, 2003), with parameters drawn from experimental literature (Drasdo and Höhme, 2005). Byrne and Drasdo (2009) upscaled a discrete model to calibrate a continuum tumour growth model, in part by using a cell velocity-based approximation of the proliferative pressure to calibrate the continuum-scale mechanics. Drasdo and co-workers were able to mechanistically model biomechanical growth limitations and the epithelial-to-mesenchymal transition in tumour cells, and they made testable hypotheses on the links between tumour hypoglycaemia and the size of the necrotic core. Galle et al. (2005, 2009) extended the approach to include cell-BM adhesion, and its impact on cell differentiation and tumour monolayer progression. Ramis-Conde et al. (2008a,b) investigated the links between a sophisticated subcellular model of E-cadherin/β-catenin signalling, intercellular signalling, and tissue morphology.

The very recent agent model of Norton et al. (2010) represented cell-cell adhesion and repulsion using a linear damped spring model, incorporated both apoptosis and necrosis, duct wall adhesion (through adhesion to myoepithelial cells), asymmetric progenitor cell division, and a simplified model of intraductal fluid pressure. The model recapitulated solid-type, comedo-type, micropapillary, and cribriform DCIS, illustrating the great potential in an agent-based modelling approach. However, the model lacked substrate transport, necrosis was modelled by imposing the viable rim thickness *a priori* rather than through a combination of cell energetics and transport limitations, and proliferating cells were randomly distributed across the viable rim with uniform distribution; this contradicts immunohistochemical observations of the distribution of proliferating DCIS cells within the duct. The authors did not treat necrotic core mechanics, which has a great impact on the overall tumour morphology and rate of tumour advance in the duct. The observed microstructures were only partly mechanistic because the model enforced polarised cell-cell adhesion and "microlumens" algorithmically; in a mechanistic model, the tumour microstructure should not be imposed, but rather emerge naturally from the model's biophysics and population dynamics. Nonetheless, their work demonstrates the great potential in using individual-based models to formulate new hypotheses on the biophysical underpinnings of cancer; based upon their polarisation model, they hypothesise that DCIS tumours progress from micropapillary to cribriform to solid-type because overproliferation collapses the "microlumens."

We are drawn to agent-based modelling due to its great potential for calibration to and comparison with *in vitro* and clinical data. After rigorously calibrating their biomechanical models, Galle et al. (2005, 2009) produced quantitatively accurate predictions of *in vitro* monolayer growth in several cell lines. They also tested competing hypotheses and compared the simulations to

additional experiments to investigate the interrelated roles of cell-cell and cell-BM contact inhibition. And in a novel *inverse mapping* approach, Engelberg et al. (2008) used an agent-based framework on a hexagonal mesh to iteratively determine a minimal set of "axiomatic operating principles" that could reproduce *in vitro* measurements of EMT6 (a mammary tumour cell line) growth characteristics in high- and low-nutrient environments. As we see in the main text, by rigorously calibrating our model to clinical DCIS data, were are able to make quantitative, testable predictions on emergent macroscopic DCIS behaviour, as well as new hypotheses on how necrotic core biomechanics affect mammography, diagnostic pathology, and clinical progression.

2 Agent model generalisations

Cell-ECM adhesion ($\mathbf{F}_{\mathbf{cma}}$): Integrins $\mathcal{I}_{\mathrm{E}}$ on the cell surface bond with suitable ligands $\mathcal{L}_{\mathrm{E}}$ in the ECM. We assume that $\mathcal{L}_{\mathrm{E}}$ is distributed proportionally to the (nondimensional) ECM density E. If $\mathcal{I}_{\mathrm{E}}$ is distributed uniformly across the cell surface and E varies slowly relative to the spatial size of a single cell, then cells at rest encounter a uniform pull from $\mathbf{F}_{\mathrm{cma}}$ in all directions, resulting in zero net cell-ECM force. For cells in motion, $\mathbf{F}_{\mathrm{cma}}$ resists that motion similarly to drag due to the energy required to overcome $\mathcal{I}-\mathcal{L}$ bonds:

$$\mathbf{F}_{\mathrm{cma}} = -c_{\mathrm{cma}}\mathcal{I}_{\mathrm{E},i}E\mathbf{v}_i. \tag{1}$$

Here, c_{cma} is a constant. If E or $\mathcal{L}_{\mathrm{E}}$ varies with a higher spatial frequency, or if $\mathcal{I}_{\mathrm{E}}$ is not uniformly distributed, then the finite half-life of $\mathcal{I}_{\mathrm{E}}-\mathcal{L}_{\mathrm{E}}$ bonds will lead to net haptotactic-type migration up gradients of E (Macklin et al., 2010b). We model this effect as part of the net locomotive force $\mathbf{F}_{\mathrm{loc}}$.

For cells in a lumen where E is zero, $\mathbf{F}_{\mathrm{cma}} = \mathbf{0}$. However, cells encounter ECM ($E \neq 0$) when invading the stroma, when pushed into the stroma through breaks in the BM (following an inadequate surgical resection or after a phenotypic change that causes MMP secretions), or following deposition of ECM by other cell species. Inclusion of this term facilitates future investigations of microinvasion and regrowth following inadequate surgical resection. As we shall see below, $\mathbf{F}_{\mathrm{cma}}$ is important for understanding Darcy's law formulations of tissue mechanics. We note that Preziosi and Tosin (2009) discussed a generalisation of Eq. 1 to develop biologically-justified tissue-scale biomechanics models; their work was supported and driven by an extensive review of the experimental literature on cell adhesion molecules.

Generalised hypoxia ($\mathcal{H}$): Cells enter $\mathcal{H}$ at any time that $\sigma < \sigma_{\mathrm{H}}$. Hypoxic cells have an exposure time-dependent probability of becoming necrotic:

$$\Pr\left(\mathcal{S}(t+\Delta t)=\mathcal{N}|\mathcal{S}(t)=\mathcal{H}\right)=1-\exp\left(-\int_t^{t+\Delta t}\beta_{\mathrm{H}}(\sigma)(s)\,ds\right)\,ds$$
$$\approx 1-\exp\left(-\beta_{\mathrm{H}}\left(\sigma\right)(t)\Delta t\right); \tag{2}$$

$\beta_{\mathrm{H}}(\sigma)(t)$ should depend upon σ to explicitly model energy depletion, as in Smallbone et al. (2007); Silva and Gatenby (2010). If $\sigma > \sigma_{\mathrm{H}}$ (normoxia is restored) at time $t+\Delta t$ and the cell has not become necrotic, it returns to its former state and resumes its activity. For example, if the cell transitioned from $\mathcal{P}$ to $\mathcal{H}$ after spending τ time in the cell cycle, and normoxic conditions are restored, then it returns to $\mathcal{P}$ with τ time having elapsed in its cell cycle progression. This is a simplified model of the persistence of the cell's proteomic state during short periods of hypoxia; indeed, restoration of normoxia can sometimes "rescue" hypoxic cells to resume cell cycling (DiGregorio et al., 2001; Green et al., 2001; Gilliland et al., 2009), and hypoxic cells can re-enter an apoptotic state upon restoration of normoxia (Tatsumi et al., 2003). Model refinements may be necessary to fully capture normoxic cell rescue, given the known multiple arrest points, the important role played by glucose in these processes, and a "startup" time to resume transcription, protein synthesis and ATP production prior to cycle resumption.

Notice that by Eq. 2, the probability that a cell succumbs to hypoxia increases with Δt whenever $\mathcal{S}=\mathcal{H}$, independently of previous states. Hence, this probability scales (nonlinearly) with its cumulative exposure time to hypoxia. This construct could model cell response to other stressors (e.g., chemotherapy), similarly to "area under the curve" models (e.g., El-Kareh and Secomb (2005)).

3 Additional mathematical notes and analyses

3.1 Relationship of the inertialess assumption to Darcy's law

Recall the inertialess formulation of the agent velocity:

$$\mathbf{v}_i=\frac{1}{\nu+c_{\mathrm{cma}}\mathcal{I}_{\mathrm{E},i}E}\left(\sum_{\substack{j=1\\ j\neq i}}^{N(t)}\left(\mathbf{F}^{ij}_{\mathrm{cca}}+\mathbf{F}^{ij}_{\mathrm{ccr}}\right)+\mathbf{F}^{i}_{\mathrm{cba}}+\mathbf{F}^{i}_{\mathrm{cbr}}+\mathbf{F}^{i}_{\mathrm{loc}}\right). \tag{3}$$

It is interesting to compare Eq. 3 with Darcy's law, the basis of many continuum-scale tumour models such as Cristini et al. (2003); Macklin and Lowengrub (2005, 2006, 2007, 2008); Macklin et al. (2009b), where tumour growth is considered as incompressible flow in a porous medium (the ECM). A mechanical pressure P models tissue mechanics as a balance of proliferation-induced

stresses, adhesion, and tissue relaxation. If $\mathbf{u}(\mathbf{x}, t)$ is the mean tissue velocity at $\mathbf{x}$, then the Darcy's law formulation of the tissue mechanics is

$$\mathbf{u} = -\mu \nabla P. \tag{4}$$

See the extensive review, discussion, and references in Lowengrub et al. (2010).

The *mobility* coefficient μ models the ability of cells to mechanically respond to pressure gradients by overcoming cell-cell and cell-ECM adhesive bonds, or by deforming the ECM (Macklin and Lowengrub, 2007). In Frieboes et al. (2007); Macklin et al. (2009b), we introduced a functional relationship between the mobility μ and the ECM density E of the form

$$\mu = \frac{1}{\alpha + \beta E + \frac{1}{\epsilon} S}, \tag{5}$$

where S is a "structure variable" that models the presence ($S = 1$) or absence ($S = 0$) of rigid barriers, $\epsilon \approx 0$, and α and β are constants. When $S = 0$, Eq. 5 is identical to the coefficient in Eq. 3. While Eq. 5 was initially chosen as the simplest possible with biologically-reasonable qualitative behaviour (mobility decreases as the ECM density increases, rendering the tissue less "permeable" to cells), it is fully consistent with the cell-scale biophysics presented above.

3.2 Relationship between the exponential random variables and nonhomogeneous Poisson processes

To date, stochastic processes have primarily been applied to understanding cell evolution (e.g., differentiation and mutation networks), but have not been commonly used to model and analyse phenotypic state transitions. Instead, phenotypic state changes are generally modelled by constant probabilities (e.g., see Anderson (2005)) which must be adjusted whenever time step sizes are altered. Modelling phenotypic state transitions as exponentially-distributed random variables is a natural generalisation of this trend, which allows us to rigorously vary the transition probabilities with variable time step sizes, such as those necessary to enforce numerical stability and/or accuracy conditions. If a transition from the quiescent state $\mathcal{Q}$ to some state $\mathcal{X}$ (e.g., $\mathcal{P}$) is governed by an exponential random process with (time-variant) parameter $\alpha(t)$, then

$$\begin{aligned} \Pr\left(\mathcal{S}(t+\Delta t) = \mathcal{X} | \mathcal{S}(t) = \mathcal{Q}\right) &= 1 - \exp\left(-\int_t^{t+\Delta t} \alpha(s)\, ds\right) \\ &\approx 1 - e^{-\alpha(t)\,\Delta t} \\ &= \alpha(t)\Delta t + \mathcal{O}\left(\Delta t^2\right). \end{aligned} \tag{6}$$

When α is constant, we recover (to second order) the commonly-used constant transition probabilities for fixed step sizes Δt; these may be regarded as approximations to our more general model here.

By classical stochastic processes theory, Eq. 6 can be regarded as arising from a nonhomogeneous Poisson process. For any phenotypic state $\mathcal{X}$ and time t, let X_t denote the number of times a cell (including its ancestors and progeny) makes a $\mathcal{Q} \to \mathcal{X}$ phenotypic transition by time t, and set $\mathcal{X}_0 = 0$. Then X_t is a counting process. If we further assume that the process has independent (but not necessarily stationary) increments and that

Property 1: $\Pr\left(X_{t+\Delta t} - X_t \geq 2\right) = \mathcal{O}\left(\Delta t^2\right)$, and
Property 2: $\Pr\left(X_{t+\Delta t} - X_t = 1\right) = \alpha(t)\Delta t + \mathcal{O}\left(\Delta t^2\right)$,

then X_t is a nonhomogeneous (due to the time variation of α) Poisson process with intensity function α. Such processes can be viewed as the originators of the exponential phenotypic transition probabilities used in our model, and the time between $\mathcal{Q} \to \mathcal{X}$ transitions are exponentially-distributed interarrival times. In particular, Eq. 6 gives the probability that there is at least one $\mathcal{Q} \to \mathcal{X}$ transition in $(t, t + \Delta t]$.

In actuality, we wish to model the probability of there being *precisely* one $\mathcal{Q} \to \mathcal{X}$ transition in $(t, t + \Delta t]$. This probability can be calculated by

$$\begin{aligned}\Pr\left(\mathcal{S}(t+\Delta t) = \mathcal{X} | \mathcal{S}(t) = \mathcal{Q}\right) &= \Pr\left(X_{t+\Delta t} - X_t = 1\right) \\ &= \exp\left(-\int_t^{t+\Delta t} \alpha(s)\, ds\right) \int_t^{t+\Delta t} \alpha(s)\, ds \\ &\approx e^{-\alpha(t)\,\Delta t} \alpha(t)\Delta t. \end{aligned} \tag{7}$$

However, by Property 1, Eq. 6 can be regarded as a second-order approximation of Eq. 7. Furthermore, by construction, the cells remain in the $\mathcal{X}$ state for a nonzero length of time τ_X, and so if $X_{t+\Delta t} - X_t \geq 2$ and $\Delta t < \tau_X$, then only the first $\mathcal{Q} \to \mathcal{X}$ transition has physical meaning. Indeed, we generally construct $\alpha(t)$ to satisfy $\alpha(t) = 0$ when $\mathcal{S} = \mathcal{X}$, and so the model precludes the possibility of two $\mathcal{Q} \to \mathcal{X}$ phenotypic transitions in a short time duration. Hence, the exponential interarrival approximation in Eq. 6 is justifiable.

Lastly, note that the simple relationship between the exponential random variables and the parameters is useful for model calibration: for a homogeneous Poisson process with intensity α, the mean time between successive events is $1/\alpha$. For non-homogeneous processes, we use the mean intensity $\langle\alpha\rangle$ to estimate the mean time between events $1/\langle\alpha\rangle$. These times could be measured *in vitro* (e.g., the mean time spent in G_0 between cell cycles), making this formulation potentially valuable for quantitative modelling.

3.3 Volume-averaged model behaviour in the viable rim

For each time t, let $\Omega(t)$ denote the (non-hypoxic) viable rim. Let $P(t)$, $A(t)$, and $Q(t)$ denote the number of proliferating, apoptosing, and quiescent cells in Ω at time t, respectively. Let $N(t) = P + A + Q$. If $\langle \alpha_{\mathrm{P}} \rangle(t) = \frac{1}{|\Omega|} \int_\Omega \alpha_{\mathrm{P}} \, dV$ is the mean value of α_{P} at time t throughout Ω, then the net number of cells entering state $\mathcal{P}$ in the time interval $[t, t + \Delta t)$ is approximately

$$\begin{aligned} P(t+\Delta t) &= P(t) + \Pr\left(\mathcal{S}(t+\Delta t) = \mathcal{P} | \mathcal{S}(t) = \mathcal{Q}\right) Q(t) - \frac{1}{\tau_{\mathrm{P}}} P(t) \Delta t \\ &\approx P(t) + \left(1 - e^{-\langle \alpha_{\mathrm{P}} \rangle \Delta t}\right) Q(t) - \frac{1}{\tau_{\mathrm{P}}} P(t) \Delta t, \end{aligned} \tag{8}$$

whose limit as $\Delta t \downarrow 0$ (after some rearrangement) is

$$\dot{P} = \langle \alpha_{\mathrm{P}} \rangle Q - \frac{1}{\tau_{\mathrm{P}}} P. \tag{9}$$

Similarly,

$$\dot{A} = \alpha_{\mathrm{A}} Q - \frac{1}{\tau_{\mathrm{A}}} A \tag{10}$$

$$\dot{Q} = 2 \frac{1}{\tau_{\mathrm{P}}} P - \left(\langle \alpha_{\mathrm{P}} \rangle + \alpha_{\mathrm{A}}\right) Q. \tag{11}$$

Summing these, we obtain

$$\dot{N} = \frac{1}{\tau_{\mathrm{P}}} P - \frac{1}{\tau_{\mathrm{A}}} A. \tag{12}$$

Next, define $\mathrm{PI} = P/N$ and $\mathrm{AI} = A/N$ to be the *proliferative* and *apoptotic indices*, respectively. We can express the equations above in terms of AI and PI by dividing by N and using Eq. 12 to properly treat $\frac{d}{dt}(P/N)$ and $\frac{d}{dt}(A/N)$. After simplifying, we obtain a nonlinear system of ODEs for PI and AI:

$$\dot{\mathrm{PI}} = \langle \alpha_{\mathrm{P}} \rangle (1 - \mathrm{AI} - \mathrm{PI}) - \frac{1}{\tau_{\mathrm{P}}} \left(\mathrm{PI} + \mathrm{PI}^2\right) + \frac{1}{\tau_{\mathrm{A}}} \mathrm{AI} \cdot \mathrm{PI} \tag{13}$$

$$\dot{\mathrm{AI}} = \alpha_{\mathrm{A}} (1 - \mathrm{AI} - \mathrm{PI}) - \frac{1}{\tau_{\mathrm{A}}} \left(\mathrm{AI} - \mathrm{AI}^2\right) - \frac{1}{\tau_{\mathrm{P}}} \mathrm{AI} \cdot \mathrm{PI}. \tag{14}$$

These equations are far simpler to compare to immunohistochemical measurements, which are generally given in terms of AI and PI.

Lastly, note that Eqs. 13-14 admit two natural time scales τ_{A} and τ_{P}. We later show that both of these scales are on the order of 1 day; see Section 4. Suppose then that we nondimensionalise this system with time scale $\bar{t} \in \{\tau_{\mathrm{A}}, \tau_{\mathrm{P}}\}$. Then supposing that the system reaches steady state for t exceeding $10\bar{t}$ to $100\bar{t}$, this analysis suggests that the local cell state dynamics reach steady state after 10 to 100 days. (Note that the values of α_{A} and α_{P} may also affect the time to steady state, although we observe steady-state population dynamics by 15 days in the main simulation results.) This is significant, because it will allow us to calibrate the population dynamic parameters (α_{A}, α_{P}) without the inherent difficulty of estimating time derivatives from often noisy *in vitro* and immunohistochemistry data. This result is consistent with our earlier mathematical analysis in Macklin and Lowengrub (2007), which hypothesised "local equilibration" of the tumour microstructure, even during growth.

4 Patient-independent parameter estimation: expanded detail

We now give expanded detail on our estimates of the parameters that are common to all patients, based upon literature searches of theoretical and experimental biology, mathematical analysis, and prior modelling efforts. In many cases, the literature can give values that may span multiple orders of magnitude; other parameters may be difficult (or impossible) to measure directly, particularly *in vivo*. In such cases, we attempt to estimate the most likely order of magnitude of the parameter, and evaluate the *ensemble package* of model plus parameter estimates plus calibration. To the extent that the ensemble makes reasonable quantitative predictions that can be validated against the experimental and clinical literature, we can accept such estimates as sufficient. Discrepancies between the model predictions and observed biological and clinical behaviour may expose specific parameters that require better measurements. Thus, this work serves as a proof of concept to determine which parameters are most critical to accurate agent-based simulations. The full list of non-specific parameters and their physical meanings is given in Table 1.

4.1 Cell cycle timescales

We estimate that the cell cycle time τ_{P} is 18 hours by the modelling literature (e.g., Owen et al. (2004)). We estimate that $\tau_{\mathrm{G1}} \approx \frac{1}{2}\tau_{\mathrm{P}} = 9$ hours (e.g., see the S + G_2 + M time in Smith and Martin (1973)).

Parameter	Physical Meaning	Value	Section
τ_{P}	cell cycle time	18 hours	4.1
τ_{G1}	length of G_1	9 hours	4.1
τ_{A}	apoptosis time	8.6 hours	4.2
τ_{NL}	necrotic cell lysis time	6 hours	4.3
f_{NS}	necrotic cell volume increase	1.0	4.3
τ_{C}	necrotic debris calcification time	15 days	4.4
L	oxygen diffusion length scale	100 μm	4.5
λ	(viable) tumour cell oxygen uptake rate	0.1 min^{-1}	4.5
λ_{b}	oxygen uptake/decay rate for non-viable cells and background	0.01λ	4.5
σ_{H}	hypoxic oxygen threshold	0.2	4.5
R_{A}	maximum adhesion distance	$1.214\,R$	4.6
c_{ccr}	cell-cell repulsive force coefficient	10.0ν μm/min	4.6
c_{cbr}	cell-BM repulsive force coefficient	c_{ccr}	4.6
n_{cca}	cell-cell adhesion potential exponent	1	4.6
n_{ccr}	cell-cell repulsion potential exponent	1	4.6
n_{cba}	cell-BM adhesion potential exponent	n_{cca}	4.6
n_{cbr}	cell-BM repulsion potential exponent	n_{ccr}	4.6
M	maximum value of $\lvert\nabla\psi\rvert$	1	4.6

Table 1
Patient-independent parameters and values for DCIS.

4.2 Apoptosis timescale

The time course from the initial signal to commence apoptosis to final cell lysis has been difficult to quantify (Hu et al., 1997). Early reviews estimated the early cellular events in apoptosis comprise a fast process on the order of minutes, with digestion of apoptotic bodies occurring within hours of phagocytosis (Kerr et al., 1994). Hu et al. (1997) conducted a detailed *in vivo* observation of apoptosis in the rat hippocampus, observing cells breaking up in 12–24 hours and the complete elimination of apoptotic bodies within 72 hours. Experimental work by Scarlett et al. (2000) similarly observed most apoptotic processes on the order of hours. These provide a bound for $\tau_{\mathrm{A}} \leq 24$ h and suggest that apoptotic bodies are eliminated in under 48 hours after cell lysis. In total, the experimental observations in the literature lead us to estimate $\tau_{\mathrm{A}} \approx \mathcal{O}(10\mathrm{h})$.

We estimate τ_{A} for breast epithelial cells based upon the hypothesis that cancerous and noncancerous cells use the same basic mechanisms of proliferation and apoptosis, only with altered frequency (Hanahan and Weinberg, 2000).

Hence, we postulate that τ_A and τ_P are the same for DCIS cells and noncancerous breast epithelial cells. The total number of cells $N(t)$ in a fixed region of breast epithelium is given by

$$\dot{N} = \left(\frac{1}{\tau_P}\mathrm{PI} - \frac{1}{\tau_A}\mathrm{AI}\right) N, \tag{15}$$

where PI and AI are the proliferative and apoptotic indices (the fractions of proliferating and apoptosing cells), respectively (supplementary material). If we assume that noncancerous breast epithelial tissue is in homeostasis (when averaged through the duration of the menstrual cycle), then $\dot{N} = 0$, and

$$\tau_A\,\mathrm{PI} = \tau_P\,\mathrm{AI}. \tag{16}$$

In Lee et al. (2006b), the mean proliferative and apoptotic indices of noncancerous breast epithelial cells in several hundred pre-menopausal women were measured at 0.0252 ± 0.0067 and 0.0080 ± 0.0006, respectively. While the AI and PI can vary considerably in time due to hormone fluctuation during the menstrual cycle (Navarrete et al., 2005), when averaged over many women (who fall at different points in this cycle), the effects of the monthly variation should cancel. Assuming that $\tau_P = 18$ h, we estimate $\tau_A \approx 5.71$ h. This is consistent our estimates above.

Since DCIS occurs predominantly in postmenopausal women, any effect of monthly variation with the menstrual cycle is not pertinent for the majority of DCIS patients. Lee et al. (2006b) measured the PI and AI in several hundred postmenopausal women at 0.0138 ± 0.0069 and 0.0043 ± 0.0007, respectively. Using these data gives $\tau_A \approx 5.62$ h. The similarity of the estimates in pre- and post-menopausal women supports our working hypothesis that τ_A and τ_P are relatively fixed for the cell type, even in different hormonal environments.

We now account for detection shortcomings in the immunostaining. (See Duan et al. (2003) for a good overview of the current apoptosis marking methods in histologic tissue samples.) The AI measurements in Lee et al. (2006b) were obtained by TUNEL assay, which detects DNA fragmentation. According to the detailed work on Jurkat cell apoptosis in Scarlett et al. (2000), there was an approximately 3-hour lag between the inducement of apoptosis (observed as rapid changes in mitochondrial membrane voltage potential and the ratio of ATP to ADP) and the detection of DNA laddering and chromatin condensation. Cleaved Caspase-3 activity was negligible for the first 60 minutes and steadily climbed thereafter, peaking after 180 minutes and reaching approximately 10% of that peak in 50-60 minutes. On this basis, we would expect that TUNEL-assay-based AI figures fail to detect approximately the first 3 hours of apoptosis, and cleaved Caspase-3-based AI stains could underestimate the first one-to-two hours. Thus, we increase our estimate for τ_A to 8.6 hours. This also gives "correction factors" to account for undetected apoptotic cells

by TUNEL assay and cleaved Caspase-3 immunostaining:

$$\mathrm{AI}_{\mathrm{actual}} \approx \frac{8.6}{5.6}\mathrm{AI}_{\mathrm{TUNEL}}, \qquad \text{and} \tag{17}$$

$$\frac{8.6}{7.6}\mathrm{AI}_{\text{Caspase-3}} \leq \mathrm{AI}_{\mathrm{actual}} \leq \frac{8.6}{6.6}\mathrm{AI}_{\text{Caspase-3}}. \tag{18}$$

4.3 Necrosis parameters

Necrotic cells lack sufficient energy to maintain ion pumps that regulate intracellular H^+, K^+, Na^+ and Ca^+ concentrations. K^+ and Na^+ play key roles in modulating cell volume; pumps for these ions are active during apoptosis to promote orderly cell shrinking and prevent premature lysis (Majno and Joris, 1995; McCarthy and Cotter, 1997; Barros et al., 2001; Cantoni et al., 2005). On this basis, we estimate $\tau_{\mathrm{NL}} < \tau_{\mathrm{A}}$ (8.6 h). This is consistent with experimental reports of necrotic cell lysis times ranging from "immediate" (e.g., Cantoni et al. (2005)), 6-7 hours (e.g., in Majno and Joris (1995)), and "overnight" (e.g., Mattes (2007)). We use $\tau_{\mathrm{NL}} = 6$ hours for our initial estimate.

There has been a wide range of reported cell volume increase (f_{NS}) in necrotic cells prior to lysis. Jun et al. (2007) reported cell volume increase of approximately 30% within 60 minutes of the onset of necrosis in SN4741 neuron cells. Necrotic "blebs" on cultured liver cells were reported to increase their volume linearly in time for over 200 minutes in Barros et al. (2003), which supports our linear necrotic core volume increase, and suggests τ_{NL} is on the order of hours. Grönroos et al. (2005) observed a 1.5-fold increase in cell volume in necrotic renal tubular cells in approximately 12 hours. Wu et al. (2010) observed necrotic cells swelling between two- and five-fold ($1 \leq f_{\mathrm{NS}} \leq 4$) after 24 hours in rat adrenal medulla cells. We use $f_{\mathrm{NS}} = 1$ as our initial estimate; other values are briefly discussed in the supplementary material, but do not significantly affect the long-term rate of tumour growth.

4.4 Calcification timescale

Little-to-no literature data are available on the calcification process for necrotic breast epithelial cells. The best available experimental data are generally animal time course studies of arterial calcification; we use these to estimate the order of magnitude of τ_{C}. Time course studies on *post mortem* cardiac valves by Jian et al. (2003) observed significant tissue calcification between 7 days (10% increase in Ca incorporation) and 14 days (40% increase) after injection by TGF-β1. Lee et al. (2006a) examined a related process (elastin calcification) using a rat subdermal model; calcification occurred gradually over two

τ_C (days)	0.5	1	5	15	30
Fraction of core calcified after 30 days (%)	94.0	83.7	51.1	6.9	0.0

Table 2
Fraction of the necrotic core occupied by calcifed debris after 30 days of simulation.

to three weeks. Gadeau et al. (2001) measured calcium accumulation in rabbit aortas following oversized balloon angioplasty injury. Calcified deposits appeared as soon as 2-4 days after the injury, increased over the course of 8 days, and approached a steady state between 8 and 30 days. Hence, we estimate τ_C is on the order of days to a few weeks.

To sharpen our estimate, we conducted a parameter study on τ_C using a simplified form of the model in Macklin et al. (2009a). We varied τ_C from 12 hours to 30 days and calculated the percentage of the necrotic core occupied by calcified debris (calculated by area). The results are in Table 2. Calcification times under 15 days lead to necrotic cores that were nearly entirely calcified; this is not typically observed in H&E images of DCIS. On the other hand, the 30-day calcification time lead (as expected) to a complete absence of microcalcifications in the core at time 30 days. Edgerton et al. (2008) hypothesised that DCIS tumours may grow to steady state in as little as two-to-three months, and so we expect microcalcifications by this time. Hence, our sharpened estimate of τ_C is 15 days, consistent with the literature.

4.5 Oxygen transport parameters

By Owen et al. (2004), the oxygen diffusion length scale L is $\sim 100\ \mu$m, and the cellular oxygen uptake rate λ (in the viable rim) can be estimated at approximately 0.1 min^{-1} via $L = \sqrt{D/\lambda}$ and their published value of D. Other values of D (e.g., from Grote et al. (1977) and Evans et al. (1981)) give $0.1\ \text{min}^{-1} \leq \lambda \leq 10\ \text{min}^{-1}$. This does not majorly impact our results because we calibrate the proliferation and oxygenation sub-models in a self-consistent manner, and λ acts as an oxygen transport time scale, and all these values yield fast equilibration relative to the proliferation timescale.

To estimate the hypoxic threshold σ_H, we examine the mitosis function $k_m(\sigma)$ in Ward and King (1997). At the step function limit, $k_m(\sigma) \propto H(\sigma - \sigma_c)$, where H is the Heaviside function. The authors determined experimentally that $\sigma_c \approx 0.2$ when σ is nondimensionalised by the far-field substrate value in non-pathologic, well-vascularized tissue. Because the step function limit is similar to α_P, our σ_H is analogous to σ_c in Ward and King (1997), and as we have nondimensionalised oxygen similarly, we set $\sigma_H = 0.2$.

We observe in our histopathology images that the quiescent and proliferating

viable tumour cells have the same general size; this suggests that the quiescent tumour cells are relatively metabolically active compared to non-cancerous, long-term quiescent cells that generally are smaller with condensed nuclei (relates to lack of transcriptional activity), reduced mitochondrial populations (Freyer, 1998), and less cytosol. Hence, we estimate that proliferating and non-proliferating viable tumour cells uptake oxygen at comparable rates. This is consistent with evidence that a cell's mitochondrial and cytoplasmic volumes are proportional (James and Bohman, 1981), and that oxygen uptake (Hystad and Rofstad, 1994) correlates with mitochondrial volume. Similarly, experiments report a linear correlation between glucose uptake and cell volume (Miller, 1964). For simplicity, we assume all tumour cell uptake oxygen at the same rate, and $\lambda_{\mathrm{b}} = 0.01\lambda$. In Video S2, we show that setting the oxygen uptake rate significantly higher in the proliferating cells (while maintaining an identical mean oxygen uptake rate via $\lambda = \mathrm{PI}\lambda_{\mathrm{P}} + (1-\mathrm{PI})\lambda_{\mathrm{NP}}$) destabilises the perinecrotic border, and is not consistent with typical patient histopathology.

4.6 Cell mechanics

When patient-specific nuclear size measurements are unavailable, we consider nuclear grade, where the tumour cell diameter is compared to the size of a red blood cell (RBC: generally 6 to 8 μm (e.g., Dao et al. (2003))). Low-grade DCIS nuclei are 1.5 to 2.0 RBCs across ($4.5\,\mu\mathrm{m} \leq R_{\mathrm{N}} \leq 7\,\mu\mathrm{m}$), high-grade are 2.5 RBCs or more ($R_n \geq 7.5\,\mu\mathrm{m}$); intermediate grade lies between these (Tan et al., 2001).

We estimate the maximum adhesive interaction distance R_{A} using published measurements of breast cancer cell deformations. Byers et al. (1995) found the deformation of MCF-7 (an adhesive, moderately aggressive breast cancer cell line) and MCF-10A (a non-malignant but transformed cell line) breast epithelial cells to be bounded around 50% to 70% of the cell radius in shear flow conditions; this is an upper bound on R_{A}. Guck et al. (2005) measured breast epithelial cell deformability (defined as additional stretched length over relaxed length) after 60 seconds of stress. Deformability increased with malignant transformation: MCF10 deformed 10.5%, MCF7 deformed 21.4%, MCF7 deformed 30.4% after weakening the cytoskeleton, and MDA-MB-231 (an aggressive, often motile cancer cell line) deformed 33.7%. Because solid-type DCIS is adhesive but not invasive, we use the MCF7 estimate and set $R_{\mathrm{A}} = 1.214R$.

We also turn to the experimental literature to estimate the overall magnitude of the mechanical forces. Cell mechanics can operate over a large range of time scales (Bursac et al., 2005), ranging from ~ 0.1 seconds for immediate viscoelastic responses to 1 minute or more when exposed to prolonged stresses (Matthews et al., 2006). Matthews et al. (2006) applied magnetic forces to

distance from cell center r (μm)	$\frac{c_{\text{ccr}}}{\nu}\,\lvert\nabla\psi(r)\rvert$ (μm/min)	distance from cell center r (μm)	$\frac{c_{\text{ccr}}}{\nu}\,\lvert\nabla\psi(r)\rvert$ (μm/min)
$R-0.50=9.4530$	0.02524	$R-0.50=9.4530$	0.05047
$R-1.00=8.9530$	0.10095	$R-1.00=8.9530$	0.20189
$R-2.00=7.9530$	0.40379	$R-2.00=7.9530$	0.80757
$R-3.00=6.9530$	0.90852	$R-3.00=6.9530$	1.81704

Table 3
Cell relaxation rate given by $|\nabla\psi|$ for $R = 9.953$ μm, $n_{\text{ccr}} = 1$, and $c_{\text{ccr}}/\nu = 10.00$ μm/min (*left*) and 20.00 μm/min (*right*), for small and intermediate deformations. The value of M does not play a role when $r > R_{\text{N}}$ (typically 4 to 7 μm).

microbeads attached to cultured endothelial cells to measure their cytoskeletal response to mechanical stress. For longer-duration stresses, they observed bead displacement velocities on the order of 0.1 μm/min to 10 μm/min (after early transient dynamics). (See Figs. 6 and 7 in Matthews et al. (2006).) We find that $c_{\text{ccr}}/\nu = 10$ μm/min gives $|\mathbf{F}_{\text{ccr}}|/\nu$ within this range for typical cell-cell interaction distances $(R-3)$ μm $< r <$ $(R-0.5)$ μm; see Table 3. This is consistent with Macklin et al. (2009a), where setting $c_{\text{ccr}}/\nu = 8$ μm/min and $c_{\text{cbr}}/\nu = 5$ μm/min prevents unreasonable simulation behaviour (overlapping cell nuclei and cell penetration of the BM). In Section 7.2, we we show that our simulation results are resilient to error in c_{ccr}: the cell density and rate of tumour growth exhibit little change over a broad range of c_{ccr}. For simplicity, we set $M = 1$, $c_{\text{ccr}} = c_{\text{cbr}} = 10\nu$ μm/min, and $n_{\text{ccr}} = n_{\text{cbr}} = 1$ (to model anticipated nonlinear but smooth cell mechanical responses).

5 Numerical methods

We implement the model using object-oriented ANSI C++, where each agent is an instance of a `Cell` class. Each cell object is endowed with an instance of a `Cell_State` class, which contains the cell phenotypic parameters ($\overline{\alpha}_P$, α_A, τ_P, etc.), volumes (V_S, V_N, V), radii (R_N, R), maximum interaction distances (R_{A}, recorded as a multiple of R), position $\mathbf{x}$, and velocity $\mathbf{v}$. We order the cells with a doubly-linked list structure: each agent is given the memory addresses of the previous and next cells. This allows us to easily delete apoptosed cells and insert new daughter cells following proliferation events. Wrapping the phenotypic properties in a `Cell_State` class makes it easy to pass heritable properties from parent to daughter cells in a generalised manner.

We discretise microenvironmental field variables (e.g., oxygen σ) on an independent Cartesian mesh with uniform spacing $\Delta x = \Delta y = 0.1L$, where L is the oxygen diffusion length scale. We represent the BM shape with a level set

function, and we use an auxiliary data structure to reduce the computational cost of cell-cell interaction testing and evaluation. (See Section 5.1.)

We now describe the program flow of this numerical implementation. In the discussion below, $N(t)$ denotes the total number of cells at time t.

(1) **Initialisation Routines:**
   (a) **Parse simulation settings file:** Parse an XML file containing all information on the simulation domain, cell types and initial arrangement, phenotypic parameters, data output times, etc. Set global variables such as the current simulation time t, the current (dynamic) time step size (Δt, initially zero), etc.

   (b) **Initialise cells:** Create new cell objects and place them within the computational domain as indicated in the prior step. For each cell, set its phenotypic parameters, and randomly select its state $\mathcal{S}$ with probabilities specified in the settings (e.g., to match immunohistochemistry). Lastly, set its progression within its state randomly (with uniform distribution), and update its volume, etc. accordingly.

   (c) **Initialise BM morphology:** Create a level set function d on a mesh (with $\Delta x = \Delta y = 1\ \mu$m) to represent the basement membrane morphology as specified in the settings. Discretise the normal vector $\mathbf{n}$ on the same mesh, computing the gradient $\mathbf{n} = \nabla d$ either analytically or by the gradient discretisations in Macklin and Lowengrub (2006).

   (d) **Initialise microenvironmental variables:** Introduce a regular Cartesian mesh and discretise the microenvironmental field variables on that mesh. For oxygen, initialise $\sigma \equiv \sigma_B$ and solve to steady state.

(2) **Main program loop:** While $t < t_{\max}$:
   (a) **Update microenvironmental variables:** Each microenvironmental variable u must be updated from $u(\mathbf{x}, t - \Delta t)$ to $u(\mathbf{x}, t)$. Solve the various microenvironmental PDEs using standard finite difference schemes (Macklin and Lowengrub, 2005, 2008; Macklin et al., 2009b). Compute volume-weighted, upscaled uptake and other reaction rates as necessary. Use independent time step sizes for each variable according to standard CFL stability criteria, until each variable has been updated to time t.

   (b) **Update cell-cell interactions:** Update the data structure for cell-cell interaction testing and evaluation. See Section 5.1.

   (c) **Update the cells:** For each cell:
      (i) *Progress the current cell state:* Update the cell with the appro-

priate submodel for its previous state $\mathcal{S}(t - \Delta t)$ until reaching the current simulation time t. Any field variable values that are necessary for computing the cell phenotypic transition probabilities (e.g., oxygen) are interpolated at the cell's position $\mathbf{x}$.

(ii) *Choose the next cell state:* If the cell was not quiescent at the previous time step, then set $\mathcal{S}(t) = \mathcal{S}(t - \Delta t)$ unless it has been altered in (2c.i). If $\mathcal{S}(t - \Delta t) = \mathcal{Q}$, then choose $\mathcal{S}(t)$ by evaluating the (exponentially-distributed) random probabilities as described in Section 5.2. We note that the probabilities can likely be approximated by their linear Taylor expansions:

$$\Pr\left(\mathcal{S}(t + \Delta t) = \mathcal{S}_2 | \mathcal{S}(t) = \mathcal{S}_1\right) \approx \alpha_{12}(\mathcal{S}(t), \bullet, \circ)\Delta t. \tag{19}$$

We are testing this acceleration among other ongoing code optimisations.

(iii) *Set the cell velocity:* Set $\mathbf{v}$ according to Eq. 3. Use the optimisation in Section 5.1 to truncate the summation to a smaller set of interacting cells.

(d) **Set Δt:** Dynamically choose the simulation time step size via:

$$\Delta t = \frac{\epsilon}{\max\left\{|\mathbf{v}_i|\right\}_{i=1}^{N(t)}}. \tag{20}$$

Here, ϵ is the desired accuracy in the cell position; we use $\epsilon = 1\ \mu$m. Note that Δt is independent of the interaction and microenvironmental mesh sizes, as the agents themselves are lattice-free.

(e) **Update cell positions:** For each cell, update the position using:

$$\mathbf{x}(t) = \mathbf{x}(t - \Delta t) + \mathbf{v}\Delta t \tag{21}$$

While we use the forward Euler difference for simplicity, improved methods (e.g., Runge-Kutta (Gottlieb and Shu, 1997; Gottlieb et al., 2001)) are straightforward to implement.

(f) **Update the simulation time:** Increment t by Δt.

Each step in the main program loop requires at most iterating through the list of the cell agents. If interaction testing can be made similarly efficient, then the overall computational effort is linear in the number of cells. To attain this, we use an auxiliary cell-cell interaction testing data structure that can be constructed linearly in the number of cells, and allows a truncation of the summation in each cell's velocity in Eq. 3, thus rendering the overall algorithm linear in the number of cells. See Section 5.1.

Let $\{k\}_{k=1}^{N(t)} = \{1, 2, 3, \cdots, N(t)\}$ be a list of all simulated cells in the computational domain $\mathcal{D}$ at time t. We construct a data structure that lists all possible cell-cell interactions at any point in the computational domain $\mathcal{D}$. We first introduce a uniform Cartesian mesh $\mathbb{M} = \{\mathbf{x}_{i,j}\} = \{(x_i, y_j)\}$ (the interaction mesh) with spacing $\Delta x = \Delta y = 1\,\mu$m. At each $\mathbf{x}_{i,j} \in \mathbb{M}$, let $\{k_m^{i,j}\}_{m=1}^{N_{i,j}(t)}$ be the list of (potentially) interacting cells at $\mathbf{x}_{i,j}$ at time t.

Step 1: Compute the maximum cell-cell interaction distance by

$$R_{\text{cca,max}} = \max\left\{R_{\text{A}}^k\right\}_{k=1}^{N(t)}. \tag{22}$$

Step 2: For each $\mathbf{x}_{i,j} \in \mathbb{M}$, set $k_0^{i,j} = 0$ and $N_{i,j}(t) = 0$. Because no cell has index 0, this denotes the case of 0 possible interactions at $\mathbf{x}_{i,j}$.

Step 3: For each cell k and for each $\mathbf{x}_{i,j}$ satisfying:

$$|\mathbf{x}_k - \mathbf{x}_{i,j}| \leq R_{\text{cca,max}} + R_{\text{cca}}^k, \tag{23}$$

set:

$$k_{N_{i,j}(t)+1}^{i,j} = k \qquad \text{(append the cell to the list at } \mathbf{x}_{i,j}) \tag{24}$$

$$N_{i,j}(t) = N_{i,j}(t) + 1. \qquad \text{(increment the total at } \mathbf{x}_{i,j}) \tag{25}$$

At each $\mathbf{x}_{i,j}$, the result is a list of all cells that can interact with a cell centred at $\mathbf{x}_{i,j}$. In C++, we implement this scheme as a singly-linked list of cell memory addresses at each $\mathbf{x}_{i,j} \in \mathbb{M}$; a NULL pointer indicates either an empty list ($N_{i,j}(t) = 0$) or the end of the list (list member $N_{i,j}(t)$ points to NULL).

For fixed ℓ and $\mathbf{x} \in \mathcal{D}$, we use this list to evaluate expressions of the form

$$\text{for all cells } k \in \{k\}_{k=1}^{N(t)} \setminus \{\ell\} \text{ compute } f\left(\text{cell}_k, \text{cell}_\ell\right)(\mathbf{x}), \tag{26}$$

such as

$$\sum_{\substack{k=1 \\ k\neq\ell}}^{N(t)} f\left(\mathbf{x}_k, \mathbf{x}_\ell\right). \tag{27}$$

Let $\mathbf{x}_{i,j}$ denote the closest interaction mesh point to $\mathbf{x}_\ell$ (the position of cell ℓ). Then we evaluate Eq. 26 by truncating it to the members of the list at $\mathbf{x}_{i,j}$:

$$\text{for all cells } k \in \left\{k_m^{i,j}\right\}_{m=1}^{N_{i,j}(t)} \setminus \{\ell\} \text{ compute } f\left(\text{cell}_k, \text{cell}_\ell\right)(\mathbf{x}). \tag{28}$$

In the example above, we truncate the summation to

$$\sum_{\substack{m=1 \\ k_m^{i,j} \neq \ell}}^{N_{i,j}(t)} f\left(\mathbf{x}_{k_m^{i,j}}, \mathbf{x}_\ell\right). \tag{29}$$

Setting the interaction mesh spacing to 1 μm sufficiently resolves cells (generally 10 to 20 μm in diameter), which reduces the impact of the nearest-neighbour approximation above; in practice, a larger spacing may suffice.

Because our interaction potentials have compact support, there is a fixed upper bound M_1 for the number of operations required to update the interaction lists for each cell; the operation is linear in the number of cells. Similarly, each interaction mesh point $\mathbf{x}_{i,j}$ has a fixed maximum number of list elements M_2, and so evaluating Eq. 28 for all cells $1 \leq \ell \leq N(t)$ is linear in the number of cells. Contrast this with Eq. 26, which for each cell ℓ scales with $N(t)$; iterating this non-truncated form over all cells thus requires $N(t)^2$ computational effort.

5.2 Evaluating probabilities

Suppose we have (assumed independent) random variables $X_1, \ldots, X_n$ with cumulative probability distributions $F_i(t)$, $1 \leq i \leq n$. We test for the occurrence of one of the events X_i in the interval $[t, t + \Delta t]$ by:

(1) Choose $r \in [0, 1]$ with uniform random distribution. Numerically, we use the `ran2` pseudorandom generator procedure from Press et al. (1992); the Mersenne twister pseudorandom generator is also commonly used.
(2) Set $p_i = F_i(t + \Delta t) - F_i(t)$ for $1 \leq i \leq n$. Define $p_0 = 0$. Set $a = b = 0$.
(3) For $1 \leq i \leq n$:
  (a) Set $a = b$ and $b = a + p_i$. (i.e., $a = \sum_{j=0}^{i-1} p_j$ and $b = \sum_{j=0}^{i} p_j$.)
  (b) If $a \leq r \leq b$, then say that event X_i has occurred in $[t, t + \Delta t]$, and end the procedure. Otherwise proceed.
(4) If we exit the loop, none of the X_i events has occurred in $[t, t + \Delta t]$.

We note that in principle, this procedure can break down for large Δt, as $\sum_{j=1}^{n} p_j$ can exceed 1. In practice, we only evaluate probabilities on short time intervals, thus the p_i are small, and this is not an issue in simulations.

Suppose that Λ is a volume-averaged uptake (or reaction) term for a continuum equation, such as

$$\frac{\partial c}{\partial t} = \nabla \cdot (D \nabla c) - \Lambda c, \tag{30}$$

where the spatial length scale $L = \sqrt{D/\bar{\Lambda}}$ is at least one order of magnitude larger than cells (e.g., $L \approx 100\,\mu$m); $\bar{\Lambda}$ is a characteristic value of Λ. For $N(t)$ defined as above, and for $1 \leq k \leq N(t)$, suppose that $\lambda = \lambda_k$ inside cell k, and $\lambda = \lambda_0$ elsewhere. We construct a cell-scale uptake term λ as follows:

Step 1: Introduce a uniform Cartesian mesh $\mathbb{M} = \{\mathbf{x}_{i,j}\} = \{(x_i, y_j)\}$ with spacing $\Delta x = \Delta y = 1\,\mu$m to discretise the computational domain.
Step 2: Initialise $\lambda_{i,j} = 0$ for each $\mathbf{x}_{i,j} \in \mathbb{M}$.
Step 3: For each cell k, centered at position $\mathbf{x}_k$ and with (equivalent) radius R_k, and for each $\mathbf{x}_{i,j} \in \mathbb{M}$ satisfying $|\mathbf{x}_{i,j} - \mathbf{x}_k| \leq R_k$, set

$$\lambda_{i,j} = \lambda_{i,j} + \lambda_k \tag{31}$$

Step 4: For each $\mathbf{x}_{i,j} \in \mathbb{M}$ such that $\lambda_{i,j} = 0$, set $\lambda_{i,j} = \lambda_0$.

After completing this procedure, the uptake rate λ equals each cell's individual uptake rate where cells are present, and the background rate λ_0 elsewhere. Note that $\lambda > \max\{\lambda_k\}_{k=1}^{N(t)}$ wherever cells overlap. This ensures the correct total uptake throughout the domain, reflecting uncertainty in the cells' morphologies. If the morphologies were known, we would replace Step 3 with:

Step 3': For each cell k and for each $\mathbf{x}_{i,j} \in \mathbb{M}$ contained within cell k, set:

$$\lambda_{i,j} = \lambda_{i,j} + \lambda_k \tag{32}$$

Lastly, we upscale this to a mesh suitable for solving Eq. 30, thereby generating the continuum-scale uptake Λ:

Step 4: Introduce a Cartesian mesh $\{(x_i, y_j) : 1 \leq i \leq M, 1 \leq j \leq N\}$ with spacing $\Delta x = \Delta y = \frac{1}{10}L$ to discretise the computational domain. Often, $\Delta x = \Delta y = 10\,\mu$m. Suppose then that $1 \leq i \leq M$ and $1 \leq j \leq N$.
Step 5: For each (i, j) satisfying $2 \leq i \leq M - 1$ and $2 \leq j \leq N - 1$, let

$$\mathcal{S}_{i,j} = \mathbb{M} \cap \left(\left[x_i - \frac{1}{2}\Delta x, x_i + \frac{1}{2}\Delta x \right] \times \left[y_j - \frac{1}{2}\Delta y, y_j + \frac{1}{2}\Delta y \right] \right),$$

i.e., the points in $\mathbb{M}$ contained in a $\Delta x \times \Delta y$ square centred at (x_i, y_j).

Let $\#\mathcal{S}_{i,j}$ denote the number of points in $\mathcal{S}_{i,j}$. Then set

$$\Lambda_{i,j} = \frac{1}{\#\mathcal{S}_{i,j}} \sum_{\mathbf{x}_{k,\ell} \in \mathcal{S}_{i,j}} \lambda_{k,\ell} \tag{33}$$

If $\Delta x = \Delta y = 10\,\mu$m, then this is a sum of 100 values on the fine mesh.

Step 6: For boundary points, we re-weight according to the smaller number of points. For example, for each $2 \leq i \leq M-1$, let

$$\mathcal{S}_{i,1} = \mathbb{M} \cap \left(\left[x_i - \frac{1}{2}\Delta x, x_i + \frac{1}{2}\Delta x \right] \times \left[y_j, y_j + \frac{1}{2}\Delta y \right] \right),$$

and set

$$\Lambda_{i,1} = \frac{1}{\#\mathcal{S}_{i,1}} \sum_{\mathbf{x}_{k,\ell} \in \mathcal{S}_{i,1}} \lambda_{k,\ell} \tag{34}$$

If $\Delta x = \Delta y = 10\,\mu$m, then this is a sum of 50 values on the fine mesh. Handle the other edges and corners analogously.

6 `MultiCellXML`: An open multicell simulation data format

We have developed a human-readable, XML-based data format for agent-based, multicell simulations (`MultiCellXML`), which includes the random seed state, global variables, information on (and filenames of) microenvironmental field variables, and a list of each cell object and its current state. This structure allows us to easily parse the data (using standardised XML parsers, such as Expat (Clark, 2007), xmlParser (Berghen, 2009), and TinyXML (Thomason et al., 2010)) for use in data visualisation and post-processing. The list of cells in the XML file is similar to the object-oriented `Cell` data structure in the simulator, making the format well-suited to resuming simulations from saved states. Modifying simulation parameters during a simulation can be achieved with simple plaintext search/replace operations in the XML files. We note that the `MultiCellXML` format is under active development; readers should reference the project website [9] for updates, documentation, and software utilities. We put forth our data format as a potential draft for inclusion in the standard being developed by Sluka et al. (2011).

We begin with XML header information (`<?xml>`) for XML 1.0 standards compliance, followed by a "root" `<data_set>` tag. In the `<data_source>` section, we include information on the originating simulation software (`<simulator>`), the user (`<user>`), and any publication information that may assist the recipient of a data file in (1) locating the original source of the data, and (2) proper

[9] `http://multicellxml.sourceforge.net`

academic citation (<reference>). See Fig. 1. Future MultiCellXML versions may include reference and citation information for the simulation software.

Following the <data_source> section, the <globals> section includes information such as the current simulation time and the random seed state–this is important for resuming saved simulation states without affecting the pseudo-random number generator. Where possible, we include information on physical units as XML tag attributes. We note that because this was initially a format developed for internal use, we have not been entirely consistent in our conventions–improvements are planned in future drafts of the file specification. For dimensionless quantities, the scale should ideally be stated (e.g., as an additional XML attribute):

```
<local_oxygen units="dimensionless" scale="far-field">0.84</local_oxygen>
```

In future drafts, we may include a new <scales> section to facilitate this.

The file format continues with a list structure of all the cells (<cell_list>), with essentially all internal cell variables (i.e., member data of the Cell class) listed clearly. We give each <cell> both a numeric type (<cell_type_code>) to assist comparing and classifying cells in software, and a human-readable type (<cell_type_text>) to assist data recipients with interpreting the data. See Fig. 2. Note that we have included "type" attributes to indicate Boolean variables, rather than units. In future file version drafts, we may include both "type" and "units" attributes to all <cell> data fields. However, we can generally assume that the presence of units indicates a non-Boolean variable, and the presence of a Boolean type obviates "units."

Due to historical reasons stemming from code development, each <cell> is split into <cell_properties> and <cell_state> sections; future versions of the data standard will likely merge these into a single <cell_state> section, because many cell properties tend to change over time due to the cells' exposure to differing microenvironments.

After all data files have been listed, we include a <global_variables> section with a list of all saved field variables and file formation information. See Fig. 3. Note that we have included the full path of each data file; often all the files (including the XML file) are saved in the same directory, so postprocessing may need to strip part of the path by comparison to the <filename> filed in the <data_source> section. Due to the large size of 2-D and 3-D double-precision data arrays, we opted for a binary data format. For increased compatibility, we choose the MATLAB .MAT (Level 4) file format, which is relatively simple to implement directly from the published file format standard (Mathworks, 2010), and is simple to read and write with common open source software (e.g., Octave) as well as MATLAB. In the source code to follow, we include C++ code to read and write these MATLAB data.

```
<?xml version="1.0" encoding="UTF-8" ?>
<data_set MultiCellXML_version="1.0">
    <data_source>
        <filename>data/output00000117.xml</filename>
        <created>29 July 2010</created>
        <simulator>
            <program_name>DCIS_2D</program_name>
            <program_version>1.38</program_version>
            <compiled></compiled>
            <author>Paul Macklin</author>
            <contact>Paul.Macklin@MathCancer.org</contact>
            <URL>http://MathCancer.org</URL>
        </simulator>
        <user>
            <name>Paul Macklin</name>
            <contact>Paul.Macklin@usc.edu</contact>
        </user>
        <reference>
            <citation>Macklin et al. J. Theor. Biol. (2011) (in review)</citation>
            <URL>http://MathCancer.org/Publications.php#macklin11_jtb</URL>
            <note>User notes may go here.</note>
        </reference>
    </data_source>
    <globals>
        <time units="minutes">7020</time>
        <next_output_time units="minutes">7020</next_output_time>
        <frame_number>117</frame_number>
        <random_seed_state>769969952</random_seed_state>
        <Domain_width_in_microns>1000</Domain_width_in_microns>
        <Domain_height_in_microns>340</Domain_height_in_microns>
    </globals>
    ...
```

Fig. 1. **Start of a `MultiCellXML` file:** The first tag is for XML 1.0 standards compliance. The `<data_source>` section indicates the source of the data, including the originating program, information on the user, and requested reference for citation (if any). The `<globals>` section gives information on program globals, including (in particular) the current simulation time and the random seed state.

Lastly, note that a primary goal of our specification is to make the format as human-readable as possible, rendering the format (partially) "self-documenting". This will make it simpler to interpret archived data long after the originating software is out of use, thus eliminating the need for reverse engineering–hence our choice of human-readable, non-binary data. While this results in much larger files, we regard data compression as a separate software problem from the specification of content. Compression can readily be applied to the data files after creation with widespread software, such as gzip.

```
 ...
<cell_list>
    <cell>
        <cell_properties>
            <cell_type_code>0</cell_type_code>
            <cell_type_text>DCIS cell</cell_type_text>
            <radius units="microns">9.95299956207</radius>
            <nuclear_radius units="microns">5.295</nuclear_radius>
            <volume units="cubic microns">4130.00487398</volume>
            <mature_volume units="cubic microns">4130.00487398</mature_volume>
            <solid_volume units="cubic microns">413.000487398</solid_volume>
            <cell_adhesion_1_level units="dimensionless">1</cell_adhesion_1_level>
            <cell_adhesion_2_level units="dimensionless">0</cell_adhesion_2_level>
            <matrix_adhesion_level units="dimensionless">1</matrix_adhesion_level>
            <calcite_level units="dimensionless">0</calcite_level>
            <mean_cell_cycle_time units="minutes">1080</mean_cell_cycle_time>
            <mean_G1_time units="minutes">540</mean_G1_time>
            <mean_time_to_apoptosis units="minutes">47196.6</mean_time_to_apoptosis>
            <mean_time_to_mitosis units="minutes">115.27</mean_time_to_mitosis>
            <cell_adhesion_exponent units="dimensionless">1</cell_adhesion_exponent>
            <BM_adhesion_exponent units="dimensionless">1</BM_adhesion_exponent>
            <calcite_adhesion_exponent units="dimensionless">1</calcite_adhesion_exponent>
            <cell_repulsion_exponent units="dimensionless">1</cell_repulsion_exponent>
            <BM_repulsion_exponent units="dimensionless">1</BM_repulsion_exponent>
            <cell_adhesion_max_distance units="x radius">1.214</cell_adhesion_max_distance>
            <BM_adhesion_max_distance units="x radius">1.214</BM_adhesion_max_distance>
            <calcite_adhesion_max_distance units="x radius">1.214</calcite_adhesion_max_distance>
        </cell_properties>
        <cell_state>
            <is_cycling type="Boolean">true</is_cycling>
            <is_quiescent type="Boolean">false</is_quiescent>
            <is_apoptosing type="Boolean">false</is_apoptosing>
            <is_hypoxic type="Boolean">false</is_hypoxic>
            <is_necrosing type="Boolean">false</is_necrosing>
            <apoptosis_time units="minutes">360.85</apoptosis_time>
            <necrosis_time units="minutes">0</necrosis_time>
            <cell_cycle_time units="minutes">0</cell_cycle_time>
            <Position units="microns">(86.5665990925,53.5000597051,0)</Position>
            <Velocity units="microns/minute">(-0.108426856979,0.213070920989,0)</Velocity>
        </cell_state>
    </cell>
    <cell>
    ...
    </cell>
...
</cell_list>
...
```


Fig. 2. **Main content of a** `MultiCellXML` **file:** Within the `<cell_list>` section, we save each individual cell agent's data within a set of `<cell></cell>` tags, including `<cell_properties>` and the `<cell_state>`. In future revisions, these fields may be merged due to the fact that cell properties change in time. *Note 1:* These fields have been minimised from the actual published datasets to simplify the presentation. *Note 2:* the `<is_debris>` tag is from an earlier version of the model, but unused here.

```
    ...
    <global_variables>
        <variable>
            <name>oxygen</name>
            <format version="Level 4">MATLAB</format>
            <filename>data/oxygen_00000117.mat</filename>
        </variable>
        <variable>
            <name>Duct_Wall_Level_Set</name>
            <format version="Level 4">MATLAB</format>
            <filename>data/level_set.mat</filename>
        </variable>
    </global_variables>
</data_set>
```

Fig. 3. **End of a `MultiCellXML` file:** After the `cell_list` section, the `global_variables` section gives a list of all associated external field data (here saved in MATLAB format).

6.1 Benchmark datasets

To demonstrate our open data format and serve as benchmark datasets, we are releasing[10] the full datasets for simulation times 0, 15, 30, and 45 days from the "baseline" simulation in the main text. Included files:

(1) `output00.zip`: contains all data from 0 days:
   (a) `output000000000.xml`: `MultiCellXML` data
   (b) `oxygen_000000000.mat`: (dimensionless) oxygen data
   (c) `levelset.mat`: basement membrane morophology
(2) `output15.zip`: contains all data from 15 days:
   (a) `output000000360.xml`: `MultiCellXML` data
   (b) `oxygen_000000360.mat`: (dimensionless) oxygen data
   (c) `levelset.mat`: basement membrane morophology
(3) `output30.zip`: contains all data from 30 days:
   (a) `output000000720.xml`: `MultiCellXML` data
   (b) `oxygen_000000720.mat`: (dimensionless) oxygen data
   (c) `levelset.mat`: basement membrane morophology
(4) `output45.zip`: contains all data from 45 days:
   (a) `output0000001080.xml`: `MultiCellXML` data
   (b) `oxygen_000001080.mat`: (dimensionless) oxygen data
   (c) `levelset.mat`: basement membrane morophology

The most up-to-date version of these datasets will be maintained at the `MultiCellXML` project website.

[10] No license applies here, aside from standard scientific citation ethics. Please reference Macklin et al. (2011).

Because the cell data are saved in a standardised XML configuration, post-processing is a combination of XML parsing and visualisation (to interpret the data). In our implementation, we choose the relatively compact TinyXML library (Thomason et al., 2010) with customised interfaces to simplify the process; this allows us to distribute code as fully self-contained, without need for installation of external libraries. We use the open source EasyBMP library (Macklin, 2005–present) for image operations. Source code is provided at MathCancer.org [8]; this software has been tested in Windows (with the mingw implementation of the g++ compiler), Linux, and OSX 10.6 in 32-bit and 64-bit environments.

In our post-processing code, we do the following:

(1) Parse the `<cell_list>` XML data:
   (a) Create a singly-linked list of a simplified `cell` objects (read from the `<cell_list>` section), consisting primarily of cell location, radius, degree of calcification, and phenotypic state.
   (b) Plot the cells in a temporary BMP image (in the program memory space–this is not actually saved as a file) for use in further geometric processing.
   (c) Plot a virtual "buffer" around all the cells to help fill in holes in the viable rim–this is essential for later cell density calculations, as well as for identifying the entire viable rim.
(2) Fill remaining holes in the viable rim to ensure its correct identification.
(3) Crop the virtual images at the leading and trailing edges to eliminate the "edge effects" and best match the patient data images. Remove the corresponding `cell` objects from the linked list.
(4) Count the total, proliferating, and apoptotic cells within the viable rim. Use these to calculate the proliferative index (PI) and apoptotic index (AI).
(5) Count the number of coloured pixels of the viable rim in the temporary image, and use this to calculate the area of the viable rim. (1 pixel is 1 μm^2.)
(6) Use the known length of the cropped domain to calculate the mean viable rim thickness.
(7) Use the viable rim area and total cell count in the cropped areas to calculate the cell density.
(8) Calculate the position of the farthest tumour cell (uncropped). Do the same for calcified cells.
(9) Use the known total numbers of (uncropped) viable tumour cells to find the 95% position (i.e., x_{V95} such that 95% of the tumour is in the region $\{(x, y) : x \leq x_{\text{V95}}\}$). Do the same for calcified cells.

(10) Append these data to `PPdata.txt`.
(11) Once done looping over all specified files, write `legend.txt` to document the structure of `PPdata.txt`

To use this code, compile it according to your compiler instructions. (`g++` users may use the supplied makefile. Windows 32-bit binaries are included in our distributions. Please note that the compiler optimisations are oriented towards 32-bit Core2 Intel processors and above.) To apply the code to the supplied data for time 30 days, type:

```
> ./PostProcessing output00000720.xml
```

To apply the code to all the supplied data, type:

```
> ./PostProcessing output*.xml
```

This software is licensed under the GPL 3.0. It is packaged with TinyXML (zlib/libpng license – see Thomason et al. (2010)) and EasyBMP (Modified BSD license – see Macklin (2005–present)). We request that users cite this paper and the project website in their "methods" section when publishing results that make substantial use of the code or derivative works.

6.3 Sample visualisation

Visualisation is performed similarly, but requires much less processing. We plot each cell as a circle with correct colour, overlay a solid nucleus, and draw a dark border. We draw the BM based upon the zero contour of the level set function. We then overlay a scale bar, label the time, and save the image.

We regard image creation and image compression as separate software problems. We use the BMP format because it is simple, universally understood, can be implemented without need for complex external libraries, and does not introduce visual artefacts to the data (in contrast to formats with lossy compression, such as JPEG). The lack of visual artefacts is also helpful for pixel-based image processing operations by other software. Users can readily compress the images using standard tools (e.g., ImageMagick and GIMP), or combine the BMP frames into an (uncompressed) AVI animation using tools such as EasyBMPtoAVI (Macklin, 2006–present).

Source code is provided at MathCancer.org [8]; this software has been tested in Windows (with the mingw implementation of the g++ compiler), Linux, and OSX 10.6 in 32-bit and 64-bit environments. To use this code, compile it according to your compiler instructions. (`g++` users may use the supplied makefile. Windows 32-bit binaries are included in our distributions. Please

note that the compiler optimisations are oriented towards 32-bit Core2 Intel processors and above.) To apply the code to the supplied data for time 30 days, type:

```
> ./visualize_DCIS_2D output00000720.xml
```

This software is licensed under the GPL 3.0. It is packaged with TinyXML (zlib/libpng license – see Thomason et al. (2010)) and EasyBMP (Modified BSD license – see Macklin (2005–present)). We request that users cite this paper and the project website in their "methods" section when publishing results that make substantial use of the code or derivative works.

7 Additional numerical studies

We performed additional numerical studies, which support and/or further investigate the model, and may be of interest to the reader.

7.1 Analysis of calibration discrepancies

To understand the discrepancy in the mean PI between our simulation and the patient data for future improvement in our calibration protocols, we examined our proliferation model more closely. Our protocol was based upon an earlier version of the proliferation "sub-model", where the cells do not divide until spending τ_{P} time in the cell cycle (Macklin et al., 2009a); in our current model, cells divide after $\tau_{\mathrm{P}} - \tau_{\mathrm{G1}}$ and continue cycling and growing for τ_{G1} time. For a given set of α_{A} and $\overline{\alpha}_P$ parameters, this should increase the simulated proliferative index. To test this, we first modify our system of ODEs to include $P_{\mathrm{SG_2M}}$ (cycling cells in S, $\mathrm{G_2}$, and M phases), $P_{\mathrm{G_1}}$ (cycling cells in $\mathrm{G_1}$ phase), A (apoptotic cells), and Q (quiescent cells in $\mathrm{G_0}$):

$$\dot{P}_{\mathrm{SG_2M}} = \langle \alpha_{\mathrm{P}} \rangle Q - \frac{1}{\tau_{\mathrm{P}} - \tau_{\mathrm{G1}}} P_{\mathrm{SG_2M}} \tag{35}$$

$$\dot{P}_{\mathrm{G_1}} = \frac{2}{\tau_{\mathrm{P}} - \tau_{\mathrm{G1}}} P_{\mathrm{SG_2M}} - \frac{1}{\tau_{\mathrm{G1}}} P_{\mathrm{G_1}} \tag{36}$$

$$\dot{A} = \alpha_{\mathrm{A}} Q - \frac{1}{\tau_{\mathrm{A}}} A \tag{37}$$

$$\dot{Q} = \frac{1}{\tau_{\mathrm{G1}}} P_{\mathrm{G_1}} - \left(\langle \alpha_{\mathrm{P}} \rangle + \alpha_{\mathrm{A}} \right) Q \tag{38}$$

If $P = P_{\mathrm{SG_2M}} + P_{\mathrm{G_1}}$ and $N = P + A + Q$, then

$$\mathrm{PI} = \frac{P}{N}, \qquad \text{and} \qquad \mathrm{AI} = \frac{A}{N}. \tag{39}$$

If we solve the system for $0 \leq t \leq 720$ hours with $Q(0) = 1$ and $A(0), P(0) = 0$ with the parameter values in the main text, then we should be able to predict the simulation's mean AI and PI. By this calculation, the PI and AI approach 24.45% and 0.761%, respectively, for Eqs. 35-38. Both of these limits match our simulated mean PI (24.04%) and AI (0.738%) very well.

All measurements given as mean ± standard deviation

Quantity	Patient Data	Simulated: $\tau_{\mathrm{G1}} = 9$ hr	Simulated: $\tau_{\mathrm{G1}} = 1$ min
PI (%)	17.43 ± 9.25	24.04 ± 4.587	18.25 ± 4.25
AI (%)	0.831 ± 0.572	0.7378 ± 0.7146	1.204 ± 0.1102
Viable rim thickness (μm)	76.92 ± 12.51	80.73 ± 1.10	80.57 ± 1.68
Cell density (cells/μm^2)	0.003213 ± 5.95e-4	0.002950 ± 6.09e-5	0.002923 ± 7.67e-05

Table 4
Verification of the patient-specific calibration (expanded): Comparison of the patient (second column) and computed (third and fourth columns) mean and standard deviation for the proliferative index, apoptotic index, viable rim thickness, and cell densities. The fourth column demonstrates that future calibration protocols should incorporate the impact of τ_{G1} using a more sophisticated population dynamic analysis. All computed quantities are within the range of patient variation.

Conversely, setting $\tau_{\mathrm{G1}} = 1$ min minimises the impact of the $\mathrm{G_1}$ phase, and the simulated PI matches the calibration target very closely; see Table 4. Because we can fully account for the discrepancy between the patient and simulated data with our improved understanding of the model, we can safely conclude that the calibration is performing well, and should match patient data exceedingly well once taking into account the division of the cell cycle into $\mathrm{G_1}$ and non-$\mathrm{G_1}$ (S-$\mathrm{G_2}$-M) phases. We plan improvements to our calibration based upon these observations in ongoing work.

7.2 Robustness of the mechanics parameters

In Section 4.6, we estimated the cell-cell repulsion parameter c_{ccr}/ν to be on the order of 10 μm/min. To assess the sensitivity of the model to error in this estimate, we varied $c_{\mathrm{ccr}}/\nu \in \{1, 2, 5, 10, 20, 100\}$ μm/min and simulated 30 days of growth with all other parameters as in the baseline case in the main text. In particular, we kept $c_{\mathrm{cca}}/c_{\mathrm{ccr}}$ constant for all the simulations to

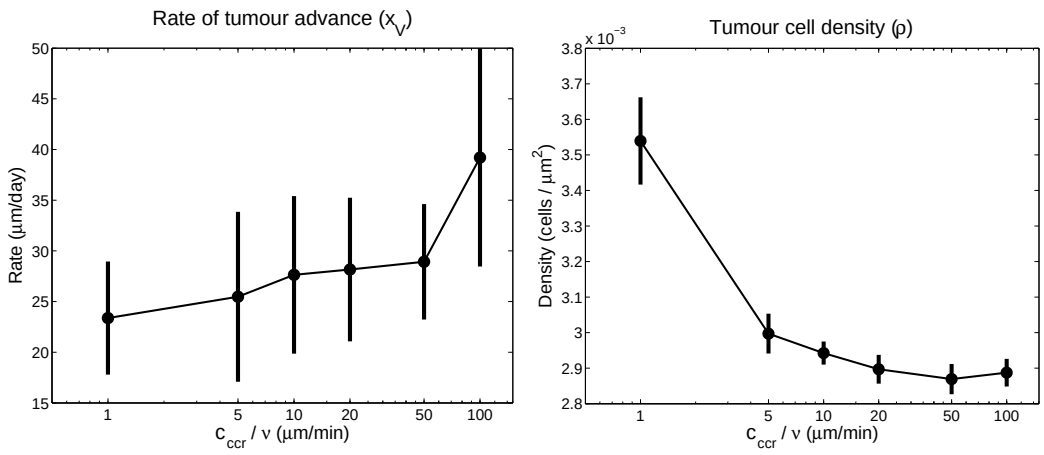

Fig. 4. **Robustness of the mechanics parameters:** We varied the cell-cell repulsive force c_{ccr}/ν while maintaining the relative balance of the forces; this is equivalent to varying the biomechanics time scale. For c_{ccr}/ν within an order of magnitude of our initial estimate, both the tumour front velocity (left plot) and viable rim cell thickness (right plot) varied little from our baseline simulation. Bars represent one standard deviation from the computed mean for each parameter value.

maintain the target cell density as in the calibration protocol, and we set $c_{\mathrm{cbr}} = c_{\mathrm{ccr}}$, and $c_{\mathrm{cba}} = 10c_{\mathrm{cca}}$. Changing c_{ccr} while maintaining these ratios of forces is equivalent to altering the biomechanics time scale.

For each combination of mechanics parameters, we calculated the smoothed tumour front velocity $x'_{\mathrm{V}}(t)$ at one-hour intervals from 10 to 25 days. (x_{V} exceeds 1 mm at 30 days for $c_{\mathrm{ccr}}/\nu = 100$ μm/min.) For any t, we calculated the smoothed $x'_{\mathrm{V}}(t)$ based upon the least-squares linear fit to x_{V} on $t \pm 24$ hours. In Fig. 4: left, we plot $\langle x'_{\mathrm{V}} \rangle$ versus c_{ccr}/ν; the bars denote one standard deviation above or below the mean. For $5 \leq c_{\mathrm{ccr}}/\nu \leq 50$ μm/min, the tumour front velocity was comparable, indicating that our simulations are robust so long as we can estimate the mechanics parameters within an order of magnitude. This is advantageous, as the individual cell mechanics parameters are among the most difficult to measure accurately.

To further evaluate the model's robustness, we calculated the mean cell density ρ throughout the viable rim at one-hour increments from 10 to 25 days for each of these simulations. In Fig. 4: right, we plot $\langle \rho \rangle$ versus c_{ccr}/ν; the bars denote one standard deviation above or below the mean. Similarly to $\langle x'_{\mathrm{V}} \rangle$, the mean cell density was comparable for $5 \leq c_{\mathrm{ccr}}/\nu \leq 100$ μm/min. This again indicates that our simulations are robust so long as we can estimate the mechanics parameters within an order of magnitude. Note that if $c_{\mathrm{ccr}}/\nu \leq 1$ μm/min, then the cell density increases rapidly due to overlapping cells, as we observed in Macklin et al. (2009a).

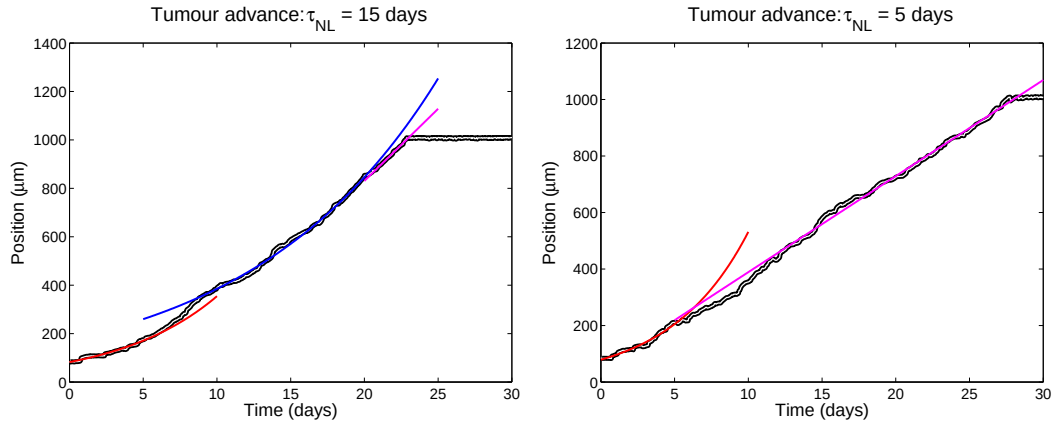

Fig. 5. The tumour advances exponentially (red and blue curves) until the viable rim is well-established and cell lysis has begun. Thereafter, the tumour advances through the duct at a constant rate (magenta curves).

7.3 On the balance of cell-cell and cell-BM adhesion

We studied the impact of the balance of cell-cell and cell-BM adhesive forces by varying $\frac{c_{\mathrm{cca}}}{c_{\mathrm{cba}}} \in \{0.01, 0.1, 1, 10\}$, while maintaining $\frac{c_{\mathrm{cca}}}{c_{\mathrm{ccr}}}$ constant. As this ratio is increased to 1 and above, the cells begin to pull off the BM except in regions of dense cell packing (results not shown). Such behviour, which may be reduced in 3D due to the duct curvature, is not consistent with typical patient histopathology. On the other hand, for $\frac{c_{\mathrm{cca}}}{c_{\mathrm{cba}}} = 0.01$, it was very difficult for daughter cells to push away from the BM after proliferation, leading to a non-physical "wetting" effect along the duct wall similar to a fluid capillary force (result not shown). We note that in the regime where $c_{\mathrm{cba}} \gg c_{\mathrm{cca}}$, the lack of tangential cell-BM force components can contribute significant errors to simulations (the BM becomes a frictionless surface).

7.4 Necrotic cell lysis is critical to linearity of tumour advance

In the main text, we found that necrotic cell lysis acts as a mechanical stress relief, and leads to a constant rate of tumour advance through the duct. To better understand this effect, we varied the necrotic cell lysis time scale $\tau_{\mathrm{NL}} \in \{2, 6 \text{ hours}, 1, 5, 15 \text{ days}\}$, with all other parameters as in the baseline simulation. To characterise the impact, we first examine the time evolution of the maximum tumour cell extent x_{V} for $\tau_{\mathrm{NL}} = 15$ days. See Fig. 5: left.

For the first 6 days, there is no necrosis, and the tumour grows exponentially; see the plot of x_{V} (black and white curve) versus the red fitted exponential curve on [0,6] in Fig. 5: left. (All exponential fits are linear least squares fits to $\log_{10} x_{\mathrm{V}}$.) At 6.08 days, the first cells necrose, and the viable region

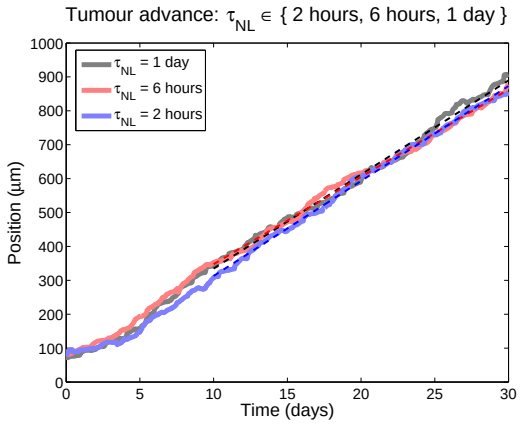

Fig. 6. For sufficiently small necrotic cell lysis times (τ_{NL}), the tumour's linear advance is roughly identical, with linear growth after the onset of necrotic cell lysis (around 5 to 6 days) that becomes clearer once the viable rim topology is well-established (around 10 days).

undergoes a topological change, splitting into upper and lower viable rims; in 3D, this would correspond to a hollow tube. Once this topological change is well-established (by approximately 10 days), growth continues exponentially at a lower rate; see the blue fitted exponential curve on [10,21] in Fig. 5: left. At 21.04 days (15 days after the first instance of necrosis), the necrotic cells begin to lyse, and the tumour's growth becomes linear as discussed earlier; see the magenta fitted line in Fig. 5: left on [21, 22.08]. At 22.08 days, the cells reach the edge of the computational domain at 1 mm.

The dynamic is the same for $\tau_{\mathrm{NL}} = 5$ days: growth is exponential at a high rate until the first instance of necrosis around 6 day; see the red fitted exponential curve on [0,6] in Fig.5: right. From 6 days to approximately 10 days, the tumour viable region is undergoing a topological change to a hollow tube; this can be observed by its transitional behaviour from 6 to approximately 10 days. The first necrotic cells begin to lyse at 11 days, and the tumour growth is linear until cells leave the simulation domain (1 mm) around 27 days; see the magenta fitted line in Fig. 5: right. Note that while growth is linear from 11 to 15 days, it appears to be at a faster rate, due to the dominance of the unlysed necrotic cells for these earlier times.

For lysis times under 1 day, growth is exponential for approximately the first 5 to 6 days, followed by a transitional period from approximately 6 to 10 days while the viable rim undergoes its topological change. Linear growth follows from 10 days until the end of the simulation at 30 days. This is the expected dynamic, given that necrotic cells begin lysing well before the end of the viable rim topological change. See Fig. 6. Note that all three tumours advance at approximately the same rate.

8 Videos

To better illustrate the key results, we include the following animations below. In each animation, the cells are labelled as follows:

- Dark blue circles: cell nuclei
- Green cells: proliferating cells ($\mathcal{S} = \mathcal{P}$; cells in non-$G_0$ phase)
- Pale blue cells: quiescent cells ($\mathcal{S} = \mathcal{Q}$; cells in G_0 phase)
- Red cells: apoptotic cells ($\mathcal{S} = \mathcal{A}$)
- grey cells: necrotic cells prior to lysis ($\mathcal{S} = \mathcal{N}, \tau < \tau_{\mathrm{NL}}$)
- Red circles: necrotic cellular debris after lysis ($\mathcal{S} = \mathcal{N}, \tau_{NL} < \tau < \tau_{\mathrm{C}}$); shade of red indicates the degree of calcification and cellular degradation
- Bright red circles: clinically-detectable calcified cellular debris $\mathcal{S} = \mathcal{N}, \tau > \tau_{\mathrm{C}}$)

The movies are in MP4 format. The open source VLC media player can play these movies on multiple platforms, including Windows, OSX, and Linux. Alternate formats are indicated below.

(1) **Video S1:** the "baseline" simulation from the main text, plotted in 1.5 mm of duct from 0 to 45 days.
Alternate format: `http://www.youtube.com/watch?v=b_GVnZWVhgk`.

(2) **Video S2:** unstable perinecrotic boundary (between the viable rim and the necrotic core) resulting from heterogeneous cellular oxygen uptake rates (proliferating cells uptake oxygen 100 times faster than quiescent cells), plotted from 0 to 30 days in 1 mm of duct.
Alternate format: `http://www.youtube.com/watch?v=Brgw8qI8k-k`.

References

A. R. A. Anderson. A hybrid mathematical model of solid tumour invasion: The importance of cell adhesion. *Math. Med. Biol.*, 22(2):163–86, 2005.

L. F. Barros, T. Hermosilla, and J. Castro. Necrotic volume increase and the early physiology of necrosis. *Comp. Biochem. Physiol. A. Mol. Integr. Physiol.*, 130(3):401–9, 2001.

L. F. Barros, T. Kanaseki, R. Sabriov, S. Morishima, J. Castro, C. X. Bittner, E. Maeno, Y. Ando-Akatsuka, and Y. Okada. Apoptotic and necrotic blebs in epithelial cells display similar neck diameters but different kinase dependency. *Cell Death Diff.*, 10(6):687–97, 2003.

F. V. Berghen. xmlParser project website, 2009. URL `http://www.applied-mathematics.net/tools/xmlParser.html`.

P. Bursac, G. Lenormand, B. Fabry, M. Oliver, D. A. Weitz, V. Viasnoff, J. P. Butler, and J. J. Fredberg. Cytoskeletal remodelling and slow dynamics in the living cell. *Nat. Mat.*, 4(7):557–61, 2005.

S. W. Byers, C. L. Sommers, B. Hoxter, A. M. Mercurio, and A. Tozeren. Role of E-cadherin in the response of tumor cell aggregates to lymphatic, venous and arterial flow: measurement of cell-cell adhesion strength. *J. Cell Sci.*, 108(5):2053–64, 1995.

H. M. Byrne and D. Drasdo. Individual-based and continuum models of growing cell populations: A comparison. *J. Math. Biol.*, 58(4–5):657–87, 2009.

O. Cantoni, A. Guidarelli, L. Palomba, and M. Fiorani. U937 cell necrosis mediated by peroxynitrite is not caused by depletion of ATP and is prevented by arachidonate via an atp-dependent mechanism. *Mol. Pharm.*, 67 (5):1399–1405, 2005.

J. Clark. Expat XML parser project website, 2007. URL `http://expat.sourceforge.net/`.

V. Cristini and J. Lowengrub. *Multiscale modeling of cancer*. Cambridge University Press, Cambridge, UK, 2010. ISBN 978-0521884426.

V. Cristini, J. S. Lowengrub, and Q. Nie. Nonlinear simulation of tumor growth. *J. Math. Biol.*, 46(3):191–224, 2003.

J. C. Dallon and H. G. Othmer. How cellular movement determines the collective force generated by the dictyostelium discoideum slug. *J. Theor. Biol.*, 231(2):203–22, 2004.

M. Dao, C. T. Lim, and S. Suresch. Mechanics of the human red blood cell deformed by optical tweezers. *J. Mech. Phys. Solids*, 51(11–12):2259–80, 2003.

T. S. Deisboeck, Z. Wang, P. Macklin, and V. Cristini. Multiscale cancer modeling. *Annu. Rev. Biomed. Eng.*, 13:127–55, 2011.

P. J. DiGregorio, J. A. Ubersax, and P. H. O'Farrell. Hypoxia and nitric oxide induce a rapid, reversible cell cycle arrest of the drosophila syncytial divisions. *J. Biol. Chem.*, 276(3):1930–7, 2001. doi: 10.1074/jbc.M003911200.

S. Dormann and A. Deutsch. Modeling of self-organized avascular tumor growth with a hybrid cellular automaton. *In Silico Biology*, 2(3):393–406, 2002.

D. Drasdo. Coarse graining in simulated cell populations. *Adv. Complex Sys.*, 8(2 & 3):319–63, 2005.

D. Drasdo and S. Höhme. Individual-based approaches to birth and death in avascular tumors. *Math. Comput. Modelling*, 37(11):1163–75, 2003.

D. Drasdo and S. Höhme. A single-scale-based model of tumor growth *in vitro*: monolayers and spheroids. *Phys. Biol.*, 2(3):133–47, 2005.

D. Drasdo, R. Kree, and J. S. McCaskill. Monte-carlo approach to tissue cell populations. *Phys. Rev. E*, 52(6):6635–57, 1995.

W. R. Duan, D. S. Garner, S. D. Williams, C. L. Funckes-Shippy, I. S. Spath, and E. A. G. Blomme. Comparison of immunohistochemistry for activated

caspase-3 and cleaved cytokeratin 18 with the TUNEL method for quantification of apoptosis in histological sections of PC-3 subcutaneous xenografts. *J. Pathol.*, 199(2):221–8, 2003.

M. Edgerton, Y.-L. Chuang, J. Kim, G. Tomaiuolo, P. Macklin, S. Sanga, W. Yang, A. Broom, K.-A. Do, and V. Cristini. Using mathematical models to understand the time dependence of the growth of ductal carcinoma in situ. *31st Annual San Antonio Breast Cancer Symposium*, Supplement to Volume 68(24):Abstract 1165, 2008.

A. W. El-Kareh and T. W. Secomb. Two-mechanism peak concentration model for cellular pharmacodynamics of doxorubicin. *Neoplasia*, 7(7):705–13, 2005.

J. Engelberg, G. Ropella, and C. A. Hunt. Essential operating principles for tumor spheroid growth. *BMC Systems Biology*, 2(1):110, 2008. doi: 10.1186/1752-0509-2-110.

N. T. S. Evans, P. F. D. Naylor, and T. H. Quinton. The diffusion coefficient of oxygen in respiring kidney and tumour tissue. *Respir. Physiol.*, 43(3): 179–88, 1981.

J. P. Freyer. Decreased mitochondrial function in quiescent cells isolated from multicellular tumor spheroids. *J. Cell. Physiol.*, 176(1):138–49, 1998.

H. B. Frieboes, J. S. Lowengrub, S. Wise, X. Zheng, P. Macklin, E. L. Bearer, and V. Cristini. Computer simulations of glioma growth and morphology. *NeuroImage*, 37(S1):S59–S70, 2007.

A.-P. Gadeau, H. Chaulet, D. Daret, M. Kockx, J.-M. Daniel-Lamaziè, and C. Desgranges. Time course of osteopontin, osteocalcin, and osteonectin accumulation and calcification after acute vessel wall injury. *J. Histochem. Cytochem.*, 49(1):79–86, 2001.

J. Galle, M. Loeffler, and D. Drasdo. Modeling the effect of deregulated proliferation and apoptosis on the growth dynamics of epithelial cell populations in vitro. *Biophys. J.*, 88(1):62–75, 2005.

J. Galle, M. Hoffmann, and G. Aust. From single cells to tissue architecture-a bottom-up approach to modelling the spatio-temporal organisation of complex multi-cellular systems. *J. Math. Biol.*, 58(1–2):261–83, 2009.

W. D. Gilliland, D. L. Vietti, N. M. Schweppe, F. Guo, T. J. Johnson, and R. S. Hawley. Hypoxia transiently sequesters Mps1 and Polo to collagenase-sensitive filaments in *Drosophila* prometaphase oocytes. *PLoS ONE*, 4(10): e7544, 10 2009. doi: 10.1371/journal.pone.0007544.

J. A. Glazier and F. Garner. Simulation of the differential adhesion driven rearrangement of biological cells. *Phys. Rev. E*, 47(3):2128–54, 1993.

S. Gottlieb and C.-W. Shu. Total variation diminishing runge-kutta schemes. *Math. Comp.*, 67(221):73–85, 1997.

S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM Review*, 43(1):89–112, 2001.

F. Graner and J. A. Glazier. Simulation of biological cell sorting using a two-dimensional extended potts model. *Phys. Rev. Lett.*, 69(13):2013–6, 1992.

S. L. Green, R. A. Freiberg, and A. J. Giaccia. p21Cip1 and p27Kip1

regulate cell cycle reentry after hypoxic stress but are not necessary for hypoxia-induced arrest. *Mol. Cell. Biol.*, 21(4):1196–1206, 2001. doi: 10.1128/MCB.21.4.1196-1206.2001.

M. Grönroos, M. Chen, T. Jahnukainen, A. Capitanio, R. I. Aizman, and G. Celsi. Methotrexate induces cell swelling and necrosis in renal tubular cells. *Ped. Blood Canc.*, 46(5):624–9, 2005.

J. Grote, R. Süsskind, and P. Vaupel. Oxygen diffusivity in tumor tissue (DS-Carcinosarcoma) under temperature conditions within the range of 20–40°C. *Pflügers Archiv Euro. J. Physiol.*, 372(1):37–42, 1977.

J. Guck, S. Schinkinger, B. Lincoln, F. Wottawah, S. Ebert, M. Romeyke, D. Lenz, H. M. Erickson, R. Ananthakrishnan, D. Mitchell, J. Käs, S. Ulvick, and C. Bilby. Optical deformability as an inherent cell marker for testing malignant transformation and metastatic competence. *Biophys. J.*, 88(5): 3689–98, 2005.

D. Hanahan and R. A. Weinberg. The hallmarks of cancer. *Cell*, 100(1):57–70, 2000.

Z. Hu, K. Yuri, H. Ozawa, H. Lu, and M. Kawata. The *In Vivo* time course for elimination of adrenalectomy-induced apoptotic profiles from the granule cell layer of the rat hippocampus. *J. Neurosci.*, 17(11):3981–9, 1997.

M. E. Hystad and E. K. Rofstad. Oxygen consumption rate and mitochondrial density in human melanoma monolayer cultures and multicellular spheroids. *Int. J. Canc.*, 57(4):532–37, 1994.

T. W. James and R. Bohman. Proliferation of mitochondria during the cell cycle of the human cell line (HL-60). *J. Cell Biol.*, 89(2):256–60, 1981.

B. Jian, N. Narula, Q.-Y. Li, E. R. Mohler III, and R. J. Levy. Progression of aortic valve stenosis: TGF-β1 is present in calcified aortic valve cusps and promotes aortic valve interstitial cell calcification via apoptosis. *Ann. Thoracic Surg.*, 75(2):457–65, 2003.

D.-J. Jun, J. Kim, S.-Y. Jung, R. Song, Y.-S. Park, S.-H. Ryu, J.-H. Kim, Y.-Y. Kong, J.-M. Chung, and K.-T. Kim. Extracellular ATP mediates necrotic cell swelling in SN4741 dopaminergic neurons through P2X7 receptors. *J. Biol. Chem.*, 282(52):37350–8, 2007.

J. F. R. Kerr, C. M. Winterford, and B. V. Harmon. Apoptosis. its significance in cancer and cancer therapy. *Cancer*, 73(8):2013–26, 1994.

J. S. Lee, D. M. Basalyga, A. Simionescu, J. C. Isenburg, D. T. Sinionescu, and N. R. Vyavahare. Elastin calcification in the rate subdermal model is accompanied by up-regulation of degradative and osteogenic cellular responses. *Am. J. Pathol.*, 168(2):490–8, 2006a.

S. Lee, S. K. Mohsin, S. Mao, S. G. Hilsenbeck, D. Medina, and D. C. Allred. Hormones, receptors, and growth in hyperplastic enlarged lobular units: early potential precursors of breast cancer. *Breast Canc. Res.*, 8(1):R6, 2006b.

J. S. Lowengrub, H. B. Frieboes, F. Jin, Y.-L. Chuang, X. Li, P. Macklin, S. M. Wise, and V. Cristini. Nonlinear modeling of cancer: Bridging the gap between cells and tumors. *Nonlinearity*, 23(1):R1–R91, 2010.

P. Macklin. EasyBMP Cross-Platform C++ BMP library project, 2005–present. URL `http://easybmp.sourceforge.net`.

P. Macklin. EasyBMPtoAVI movie creator project, 2006–present. URL `http://easybmptoavi.sourceforge.net`.

P. Macklin. Biological background. In *Multiscale modeling of cancer* Cristini and Lowengrub (2010), chapter 2, pages 8–24. ISBN 978-0521884426.

P. Macklin. Basic ductal carcinoma in situ (DCIS) pathobiology for modelers. 2012. URL `http://MathCancer.org/Resources.php#DCIS_biology_tutorial`.

P. Macklin and J. S. Lowengrub. Evolving interfaces via gradients of geometry-dependent interior poisson problems: application to tumor growth. *J. Comput. Phys.*, 203(1):191–220, 2005.

P. Macklin and J. S. Lowengrub. An improved geometry-aware curvature discretization for level set methods: application to tumor growth. *J. Comput. Phys.*, 215(2):392–401, 2006.

P. Macklin and J. S. Lowengrub. Nonlinear simulation of the effect of microenvironment on tumor growth. *J. Theor. Biol.*, 245(4):677–704, 2007.

P. Macklin and J. S. Lowengrub. A new ghost cell/level set method for moving boundary problems: Application to tumor growth. *J. Sci. Comp.*, 35(2–3): 266–99, 2008.

P. Macklin, J. Kim, G. Tomaiuolo, M. E. Edgerton, and V. Cristini. Agent-based modeling of ductal carcinoma in situ: Application to patient-specific breast cancer modeling. In T. Pham, editor, *Computational Biology: Issues and Applications in Oncology*, chapter 4, pages 77–112. Springer, 2009a. ISBN 978-1441908100.

P. Macklin, S. McDougall, A. R. A. Anderson, M. A. J. Chaplain, V. Cristini, and J. Lowengrub. Multiscale modeling and nonlinear simulation of vascular tumour growth. *J. Math. Biol.*, 58(4–5):765–98, 2009b.

P. Macklin, M. E. Edgerton, and V. Cristini. Agent-based cell modeling: application to breast cancer. In *Multiscale modeling of cancer* Cristini and Lowengrub (2010), chapter 10, pages 216–244. ISBN 978-0521884426.

P. Macklin, M. E. Edgerton, J. Lowengrub, and V. Cristini. Discrete cell modeling. In *Multiscale modeling of cancer* Cristini and Lowengrub (2010), chapter 6, pages 92–126. ISBN 978-0521884426.

P. Macklin, M. E. Edgerton, A. M. Thompson, and V. Cristini. Patient-calibrated agent-based modelling of ductal carcinoma in situ (DCIS): From microscopic measurements to macroscopic predictions of clinical progression. *J. Theor. Biol.*, 2011. (in final review).

G. Majno and I. Joris. Apoptosis, oncosis, and necrosis. an overview of cell death. *Am. J. Pathol.*, 146(1):3–15, 1995.

Mathworks. MATLAB 7 MAT-file format, 2010. URL `http://www.mathworks.com/help/pdf_doc/matlab/matfile_format.pdf`.

M. J. Mattes. Apoptosis assays within lymphoma cell lines: problems and pitfalls. *Brit. J. Canc.*, 96(6):928–36, 2007.

B. D. Matthews, D. R. Overby, R. Mannix, and D. E. Ingber. Cellular adap-

tation to mechanical stress: role of integrins, Rho, cytoskeletal tension and mechanosensitive ion channels. *J. Cell. Sci.*, 119(3):508–18, 2006.

J. V. McCarthy and T. G. Cotter. Cell shrinkage and apoptosis: a role for potassium and sodium ion efflux. *Cell Death Diff.*, 4(8):756–70, 1997.

D. M. Miller. Sugar uptake as a function of cell volume in human erythrocytes. *J. Physiol.*, 170(1):219–25, 1964.

M. A. H. Navarrete, C. M. maier, R. Falzoni, L. G. d. A. Quadros, E. C. Baracat, and A. C. P. Nazário. Assessment of the proliferative, apoptotic, and cellular renovation indices of the human mammary epithelium during the follicular and luteal phases of the menstrual cycle. *Breast Cancer Res.*, 7(3):R306–13, 2005.

K.-A. Norton, M. Wininger, G. Bhanot, S. Ganesan, N. Barnard, and T. Shinbrot. A 2D mechanistic model of breast ductal carcinoma in situ (DCIS) morphology and progression. *J. Theor. Biol.*, 263(4):393–406, 2010.

M. R. Owen, H. M. Byrne, and C. E. Lewis. Mathematical modelling of the use of macrophages as vehicles for drug-delivery to hypoxic tumour sites. *J. Theor. Biol.*, 226(4):377–391, 2004.

E. Palsson and H. G. Othmer. A model for individual and collective cell movement in Dictyostelium discoideum. *Proc. Nat. Acad. Sci. USA*, 97 (19):10338–453, 2000.

T. Pham, editor. *Computational Biology: Issues and Applications in Oncology.* Springer, New York, NY USA, 2009. ISBN 978-1441908100.

W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press, 2nd edition, 1992. ISBN 0-521-43108-5.

L. Preziosi and A. Tosin. Multiphase modeling of tumor growth and extracellular matrix interaction: Mathematical tools and applications. *J. Math. Biol.*, 58(4-5):625–56, 2009.

I. Ramis-Conde, M. A. J. Chaplain, and A. R. A. Anderson. Mathematical modelling of cancer cell invasion of tissue. *Math. Comp. Model.*, 47(5–6): 533–545, 2008a.

I. Ramis-Conde, D. Drasdo, A. R. A. Anderson, and M. A. J. Chaplain. Modeling the influence of the e-cadherin-beta-catenin pathway in cancer cell invasion: A multiscale approach. *Biophys. J.*, 95(1):155–165, 2008b.

K. A. Rejniak. An immersed boundary framework for modeling the growth of individual cells: An application to the early tumour development. *J.Theor. Biol.*, 247(1):186–204, 2007.

K. A. Rejniak and A. R. A. Anderson. A computational study of the development of epithelial acini: I. sufficient conditions for the formation of a hollow structure. *Bull. Math. Biol.*, 70(3):677–712, 2008a.

K. A. Rejniak and A. R. A. Anderson. A computational study of the development of epithelial acini: II. necessary conditions for structure and lumen stability. *Bull. Math. Biol.*, 70(5):1450–79, 2008b.

K. A. Rejniak and R. H. Dillon. A single cell-based model of the ductal tumor microarchitecture. *Comp. Math. Meth. Med.*, 8(1):51–69, 2007.

J. L. Scarlett, P. W. Sheard, G. Hughes, E. C. Ledgerwood, H.-K. Ku, and M. P. Murphy. Changes in mitochondrial membrane potential during staurosporine-induced apoptosis in Jurkat cells. *FEBS Letters*, 475(3): 267–72, 2000.

G. Schaller and M. Meyer-Hermann. Multicellular tumor spheroid in an off-lattice Voronoi-Delaunay cell model. *Phys. Rev. E*, 71(5):051910, 2005.

A. S. Silva and R. A. Gatenby. A theoretical quantitative model for evolution of cancer chemotherapy resistance. *Biol. Direct*, 5(1):25ff, 2010.

J. P. Sluka, J. A. Glazier, and M. Swat. Frameworks for shareable multiscale modeling: Goals, philosophy, tools, languages and needs. *Phil. Trans. Royal Soc. A*, 2011. URL `http://bit.ly/nih_sluka`. (in preparation).

K. Smallbone, R. A. Gatenby, R. J. Gillies, P. K. Maini, and D. J. Gavaghan. Metabolic changes during carcinogenesis: Potential impact on invasiveness. *J. Theor. Biol.*, 244(2):703–713, 2007.

J. A. Smith and L. Martin. Do cells cycle? *Proc. Natl. Acad. Sci. USA*, 70: 1263–67, 1973. URL `http://www.pnas.org/content/70/4/1263`.

P. H. Tan, B. B. Goh, G. C. Chiang, and B. H. Bay. Correlation of nuclear morphometry with pathologic parameters in ductal carcinoma *in situ* of the breast. *Mod. Pathol.*, 14(10):937–41, 2001.

T. Tatsumi, J. Shiraishi, N. Keira, K. Akashi, A. Mano, S. Yamanaka, S. Matoba, S. Fushiki, H. Fliss, and M. Nakagawa. Intracellular ATP is required for mitochondrial apoptotic pathways in isolated rat cardiac myocytes. *Cardiovascular Res.*, 59(2):428–40, 2003.

L. Thomason et al. TinyXML project website, 2010. URL `http://www.sourceforge.net/projects/tinyxml`.

J. P. Ward and J. R. King. Mathematical modelling of avascular tumour growth. *IMA J. Math. Appl. Medicine Biol.*, 14(1):36–69, 1997.

L.-Y. Wu, Z.-M. Ma, X.-L. Fan, T. Zhao, Z.-H. Liu, X. Huang, M.-M. Li, L. Xiong, K. Zhang, L.-L. Zhu, and M. Fan. The anti-necrosis role of hypoxic preconditioning after acute anoxia is mediated by aldose reductase and sorbitol pathway in PC12 cells. *Cell Stress and Chaperones*, 15(4): 387–94, 2010.