Patient-calibrated agent-based modelling of ductal carcinoma in situ (DCIS): From microscopic measurements to macroscopic predictions of clinical progression

Paul Macklin*, Mary E. Edgerton, Alastair M. Thompson, and Vittorio Cristini

* Corresponding author: Paul.Macklin@usc.edu

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):

- DCIS biology
- Sampling of significant agent-based modelling beyond DCIS
- Agent model generalisations
- Additional mathematical notes and analyses (including volume averaging)
- Examination of volume-averaged analysis in DCIS ducts (testing Michaelis-Menten predictions)
- Expanded detail for parameter estimations based on the literature
- Numerical methods and `MultiCellXML`
- Additional numerical results
- Videos

Extensive resources, including this document, C++ source code, benchmark datasets, and animations are mirrored and maintained at MathCancer.org. [8]

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_2011/`

1 Ductal carcinoma in situ (DCIS) of the breast

Patient-specific DCIS simulation both motivates our model development and serves as a test bed for the resulting framework. We now discuss the specific biology of normal breast tissue and how that biology is subverted in DCIS. For further pertinent biological background (including support for some modelling assumptions), see Macklin (2010) and the references therein. The biology discussed below is generally applicable to most epithelial malignancies.

1.1 Biology of breast duct epithelium

The breast is organised as a system of 12-15 independent, largely parallel duct systems: clusters of milk-producing lobules that feed into a branched duct system that terminates at the nipple (Wellings et al., 1975; Moffat and Going, 1996; Ohtake et al., 2001; Going and Mohun, 2006). The duct systems are separated by supporting ligaments and fatty tissue and drained by the lymphatic system (Tannis et al., 2001). Each duct is a tubular arrangement of epithelial cells that enclose a fluid-filled lumen. The epithelium, in turn, is surrounded by myoepithelial cells (epithelial cells with contractile properties to transport milk) and a basement membrane. Surrounding the duct is the stroma, which is comprised primarily of a supporting scaffolding of fibres (the extracellular matrix, or ECM) and mesenchymal cells that maintain the ECM. The stroma is interlaced by blood vessels that supply oxygen and other vital substrates to the tissue. See Fig. 1 (top left). Note that the breast epithelium has no direct access to oxygen and nutrients; these must diffuse into the duct through the BM.

The epithelial cells are *polarised*: integrins on a well-defined basal side adhere to the basement membrane, E-cadherin molecules on the lateral sides adhere to neighbouring cells, and the apical side has relatively few adhesion molecules. See Fig. 1 (top right). The epithelial cell arrangement in the duct depends critically upon this polarisation and the resulting nonuniform distribution of adhesive forces (Jiang and Chuong, 1992; Hansen and Bissell, 2000; Wei et al., 2007; Butler et al., 2008).

While the epithelial cell population oscillates with the menstrual cycle (Khan et al., 1998, 1999), on average proliferation and apoptosis balance to maintain homeostasis. Microenvironmental changes can trigger signalling responses that lead to proliferation or apoptosis, which ordinarily helps to safeguard the normal tissue architecture. For example, a decrease of E-cadherin signalling (following apoptosis in a neighbouring cell) can increase β-catenin signalling, which eventually increases proliferation to replace the missing cell (Conacci-
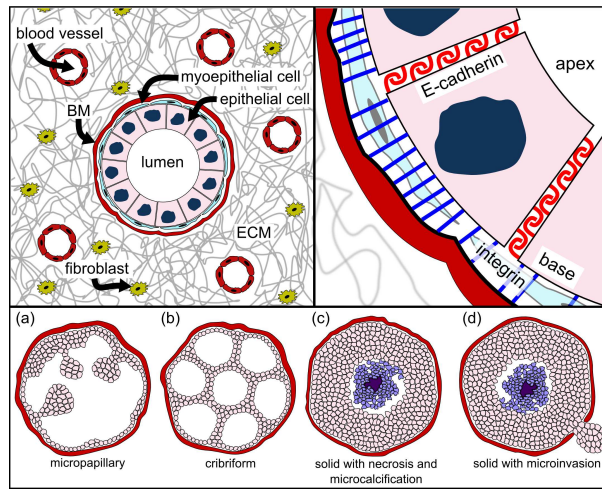
Fig. 1. *Top Left:* Typical breast duct microanatomy. *Top Right:* Breast duct epithelial cell polarisation. *Bottom:* Major DCIS types and invasive ductal carcinoma. An advance copy of this figure appeared in Macklin et al. (2009a, 2010a).

Sorrell et al., 2002; Hansen and Bissell, 2000; Wei et al., 2007). Adhesion to the BM triggers integrin signalling and downstream production of survival proteins that inhibit apoptosis (Ilić et al., 1998; Giancotti and Ruoslahti, 1999; Stupack and Cheresh, 2002). Loss of attachment to the BM therefore allows one type of apoptosis (*anoikis*) to occur, thus preventing overgrowth of cells into the lumen (Danes et al., 2008). Hormones such as estrogen, progesterone, prolactin, and epidermal growth factor can affect epithelial cell proliferation and apoptosis prior to lactation (Anderson, 2004), during breast involution (Baxter et al., 2007), and in cancer (Simpson et al., 2005).

1.2 Biology of DCIS

Overexpressed oncogenes and underexpressed tumour suppressor genes can disrupt the balance of epithelial cell proliferation and apoptosis, leading to overproliferation. This can occur typically either by the accumulation of DNA mutations (genetic damage) or DNA amplification (Simpson et al., 2005), or epigenetic anomalies (Ai et al., 2006). The transformation from regular breast epithelium to carcinoma is thought to occur in stages. For simplicity, we set aside the relatively benign precursor transformations (e.g., atypical ductal hyperplasia) which have a low risk for subsequent invasive breast cancer (Page, 1992) and focus on DCIS.

In the most well-differentiated classes of DCIS, the epithelial cells maintain their polarity and anisotropic adhesion receptor distributions, resulting in partial recapitulation of the non-pathological duct structure within the lumen. These demonstrate either finger-like growths into the lumen (micropapillary:

3

see Fig. 1 (bottom:a)), or arrangements of duct-like structures (cribriform: see Fig. 1 (bottom:b)) (Silverstein, 2000). The cells in solid type DCIS lack polarity and do not develop these microstructures. Instead, the cells proliferate until filling the entire lumen (Fig. 1 (bottom:c)) (Danes et al., 2008). The proliferating cells uptake oxygen and nutrients as they diffuse into the duct, causing substrate gradients to form. If the central oxygen level is sufficiently depleted, a necrotic core of debris forms (comedo-type solid DCIS: see Fig. 1 (bottom:c)) (Silverstein, 2000). These necrotic cells are typically not phagocytosed; instead, they swell and burst (Barros et al., 2001), and their solid (i.e., non-water) components are slowly calcified (Stomper and Margolin, 1994). It is these calcifications that are generally detected by mammograms when diagnosing DCIS (Ciatto et al., 1994). The BM blocks DCIS from invading the stroma, thereby impeding spread through the stroma, invasion into lymphovascular channels and hence metastasis. Further mutations can transform DCIS into invasive ductal carcinoma, whose cells move along the duct, secrete matrix metalloproteinases (MMPs) to degrade the BM and subsequently invade the stroma (Fig. 1 (bottom:d)). See Silver and Tavassoli (1998) and Adamovich and Simmons (2003).

While it is tempting to regard DCIS as a linear progression from regular epithelium to cribriform or micropapillary ("partially transformed") to solid type ("fully transformed"), the morphological and molecular pathway is currently an open question (Erbas et al., 2006; Rennstam and Hedenfalk, 2006). The excellent modelling and analysis by Sontag and Axelrod (2005) strongly refutes a linear progression model. The dominant type of DCIS in any particular case may depend upon the underlying molecular changes. For example, cribriform DCIS could arise from hyperproliferative cells where genes regulating polarisation are functionally intact.

2 A sampling of significant agent-based modelling beyond DCIS

It is beyond the scope of this overview to exhaustively review all discrete biomathematics modelling; instead, 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 of discrete modelling, please see Lowengrub et al. (2010); Macklin et al. (2010b); Deisboeck et al. (2011) and the references therein, such as Anderson and Quaranta (2008); Deisboeck et al. (2009) and Zhang et al. (2009).

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 co-workers (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 biophysical and kinetic 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 dif-

5

ferentiation and tumour monolayer progression. Ramis-Conde et al. (2008a,b) used their model to investigate 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 (e.g., as in Fig. 2). 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 over-proliferation 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 a particularly rigorous calibration of their biomechanical models, Galle et al. (2005, 2009) produced quantitatively accurate predictions of *in vitro* monolayer growth in several cell lines. Additionally, they 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.

3 Agent model generalisations

Heterophilic cell-cell adhesion ($\mathbf{F_{cca}}$): In heterophilic cell-cell adhesion (e.g., Springer (1990); Terol et al. (2003); Lucio et al. (1998)), adhesion receptors $\mathcal{I}_A$ bond with dissimilar ligands $\mathcal{I}_B$, and vice versa. Hence,

$$f_{i,j} = \mathcal{I}_{A,i}\mathcal{I}_{B,j} + \mathcal{I}_{B,i}\mathcal{I}_{A,j}, \tag{1}$$

where $\mathcal{I}_{A,i}$ and $\mathcal{I}_{B,i}$ are cell i's (nondimensionalised) $\mathcal{I}_A$ and $\mathcal{I}_B$ expressions.

Cell-ECM adhesion ($\mathbf{F_{cma}}$): Integrins $\mathcal{I}_E$ on the cell surface form heterophilic bonds with suitable ligands $\mathcal{L}_E$ in the ECM. We assume that $\mathcal{L}_E$ is distributed proportionally to the (nondimensional) ECM density E. If $\mathcal{I}_E$ 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 $\mathbf{F}_{cma}$ in all directions, resulting in zero net cell-ECM force. For cells in motion, $\mathbf{F}_{cma}$ resists that motion similarly to drag due to the energy required to overcome $\mathcal{I} - \mathcal{L}$ bonds:

$$\mathbf{F}_{cma} = -c_{cma}\mathcal{I}_{E,i}E\mathbf{v}_i. \tag{2}$$

Here, c_{cma} is a constant. If E or $\mathcal{L}_E$ varies with a higher spatial frequency, or if $\mathcal{I}_E$ is not uniformly distributed, then the finite half-life of $\mathcal{I}_E - \mathcal{L}_E$ 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 $\mathbf{F}_{loc}$.

For cells in a lumen where E is zero, $\mathbf{F}_{cma} = \mathbf{0}$. 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, $\mathbf{F}_{cma}$ is important for understanding Darcy's law formulations of tissue mechanics. We note that Preziosi and Tosin (2009) discussed a generalisation of Eq. 2 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 the hypoxic state at any time that $\sigma < \sigma_H$. Hypoxic cells have an exposure time-dependent probability of becoming necrotic:

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

$$\approx 1 - \exp\left(-\beta_{\mathrm{H}}\left(\sigma\right)(t)\Delta t\right). \tag{3}$$

We currently model $\beta_{\mathrm{H}}(\sigma)(t)$ as constant, although it could readily be made dependent upon σ to more explicitly model energy depletion, such as in Smallbone et al. (2007a); 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 (e.g., G_1 and metaphase during M), 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. 3, 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)).

4   Additional mathematical notes and analyses

4.1   Relationship of the inertialess assumption to Darcy's law

Recall the inertialess formulation of the agent velocity:

$$\mathbf{v}_i = \frac{1}{\nu + c_{\mathrm{cma}}\mathcal{I}_{\mathrm{E},i}E}\left(\sum_{\substack{j=1 \\ j \neq i}}^{N(t)} \left(\mathbf{F}_{\mathrm{cca}}^{ij} + \mathbf{F}_{\mathrm{ccr}}^{ij}\right) + \mathbf{F}_{\mathrm{cba}}^i + \mathbf{F}_{\mathrm{cbr}}^i + \mathbf{F}_{\mathrm{loc}}^i\right). \tag{4}$$

It is interesting to compare Eq. 4 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 mechan-

ical 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

$$\mathbf{u} = -\mu \nabla P. \tag{5}$$

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

$$\mu = \frac{1}{\alpha + \beta E + \frac{1}{\epsilon}S}, \tag{6}$$

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. 6 is identical to the coefficient in Eq. 4. While Eq. 6 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.

4.2 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

$$
\begin{aligned}
\Pr\left(\mathcal{S}(t + \Delta t) = \mathcal{X} | \mathcal{S}(t) = \mathcal{Q}\right) &= 1 - \exp\left(-\int_t^{t+\Delta t} \alpha(s)\, ds\right) \\
&\approx 1 - e^{-\alpha(t)\,\Delta t} \\
&= \alpha(t)\Delta t + \mathcal{O}\left(\Delta t^2\right).
\end{aligned}
\tag{7}
$$

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

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. 7 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

$$
\begin{aligned}
\Pr(\mathcal{S}(t + \Delta t) = \mathcal{X} | \mathcal{S}(t) = \mathcal{Q}) &= \Pr(X_{t+\Delta t} - X_t = 1) \\
&= \exp\left(-\int_t^{t+\Delta t} \alpha(s)\,ds\right) \int_t^{t+\Delta t} \alpha(s)\,ds \\
&\approx e^{-\alpha(t)\,\Delta t}\alpha(t)\Delta t.
\end{aligned}
\tag{8}
$$

However, by Property 1, Eq. 7 can be regarded as a second-order approximation of Eq. 8. 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. 7 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.

4.3   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 $\langle\alpha_P\rangle(t) = \frac{1}{|\Omega|}\int_\Omega \alpha_P\, dV$ 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

$$P(t + \Delta t) = P(t) + \Pr\left(\mathcal{S}(t + \Delta t) = \mathcal{P}\,|\,\mathcal{S}(t) = \mathcal{Q}\right)Q(t) - \frac{1}{\tau_P}P(t)\Delta t$$

$$\approx P(t) + \left(1 - e^{-\langle\alpha_P\rangle\Delta t}\right)Q(t) - \frac{1}{\tau_P}P(t)\Delta t, \tag{9}$$

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

$$\dot{P} = \langle\alpha_P\rangle Q - \frac{1}{\tau_P}P. \tag{10}$$

Similarly,

$$\dot{A} = \alpha_A Q - \frac{1}{\tau_A}A \tag{11}$$

$$\dot{Q} = 2\frac{1}{\tau_P}P - \left(\langle\alpha_P\rangle + \alpha_A\right)Q. \tag{12}$$

Summing these, we obtain

$$\dot{N} = \frac{1}{\tau_P}P - \frac{1}{\tau_A}A. \tag{13}$$

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. 13 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:

$$\dot{\mathrm{PI}} = \langle\alpha_P\rangle\left(1 - \mathrm{AI} - \mathrm{PI}\right) - \frac{1}{\tau_P}\left(\mathrm{PI} + \mathrm{PI}^2\right) + \frac{1}{\tau_A}\mathrm{AI}\cdot\mathrm{PI} \tag{14}$$

$$\dot{\mathrm{AI}} = \alpha_A\left(1 - \mathrm{AI} - \mathrm{PI}\right) - \frac{1}{\tau_A}\left(\mathrm{AI} - \mathrm{AI}^2\right) - \frac{1}{\tau_P}\mathrm{AI}\cdot\mathrm{PI}. \tag{15}$$

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

Lastly, note that Eqs. 14-15 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 6. Suppose then that we nondimensionalise this system with time scale $\bar{t} \in \{\tau_A, \tau_P\}$. 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.

5 Application of the volume-averaged analysis: Relationship between proliferation and oxygenation in breast ducts; validation against patient Ki-67 immunohistochemistry

We apply volume-averaged analysis the volume-averaged analysis from Section 4.3 to the viable rim in DCIS to generate biological hypotheses that we test against immunohistochemistry data. For fixed AI, PI, τ_A, and τ_P, we can use Eqs. 14-15 to determine $\langle \alpha_P \rangle$ and α_A, and ultimately, $\overline{\alpha}_P$. Indeed, this is the basis of our calibration procedure.

In Macklin et al. (2009a), we instead treated α_A and $\overline{\alpha}_P$ and constant and solved the nonlinear ODE system for PI and AI to steady state as a function of $0 \leq \sigma \leq 1$. This analysis led us to predict Michaelis-Menten population kinetics as an emergent model phenomenon: for sufficient oxygen availability, proliferation saturates, indicating that oxygenation is no longer the primary growth-limiting factor.

We now test this hypothesis based upon a careful analysis of Ki-67 immunohistochemistry in two exemplar ducts (F3 and F19) for a DCIS patient (anonymised case 100019) (Edgerton et al., 2011). See Fig. 2. For each of these ducts, we calculate the distance of all nuclei and Ki-67 positive nuclei to the duct wall, the mean distance from the duct centroid to the duct wall (i.e., the radius R_{duct}), and the mean duct viable rim thickness T. See the full description of the image processing in Section 5.1, along with C++ code and the annotated images.
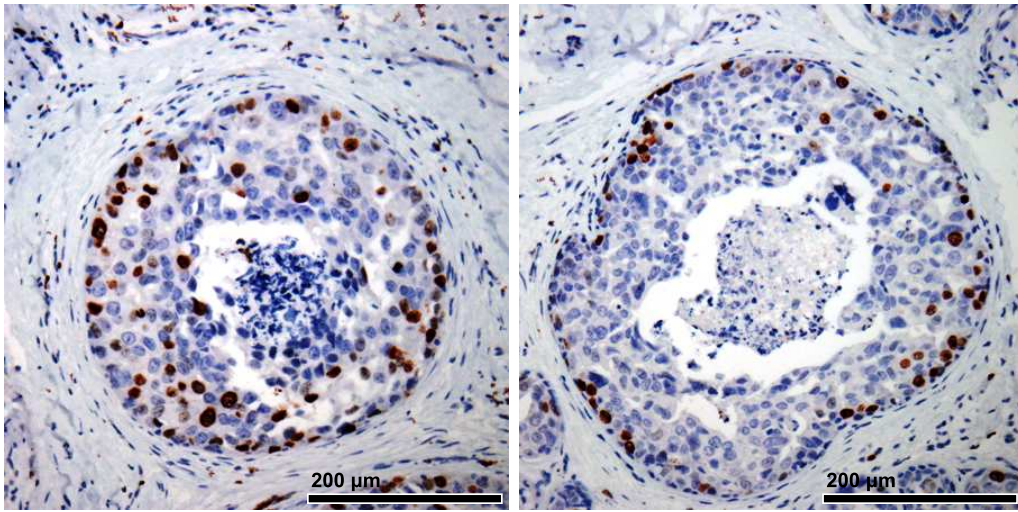
12

Fig. 2. Ki-67 immunohistochemistry for ducts F3 (left) and F19 (right) for anonymised case 100019. Ki-67 positive nuclei stain dark red; Ki-67 negative nuclei are counterstained light blue. A colour version of this image is available online. An advance copy of this figure appeared in Macklin et al. (2010a).

Next, we create a histogram of Ki-67-positive nucleus distances to the duct wall (Fig. 3, first row), all nucleus distances to the duct wall *using the same histogram "bins"* (Fig. 3, second row), and divide these to obtain the proliferative index (PI) versus distance from the duct wall (Fig. 3, third row).

Next, we estimate the 3-D steady-state oxygen profile through the ducts (assumed radially symmetric with no variation in the longitudinal direction):

$$0 = L^2 \left(\sigma'' + \frac{1}{r}\sigma' \right) - \sigma, \quad 0 < r < R_{\text{duct}} \tag{16}$$

with boundary conditions

$$\sigma(R_{\text{duct}} - T) = \sigma_{\text{H}}, \qquad \sigma'(0) = 0, \tag{17}$$

The solution is

$$\sigma(r) = \frac{\sigma_{\text{H}}}{I_0 \left(\frac{R_{\text{duct}} - T}{L} \right)} I_0 \left(\frac{r}{L} \right), \tag{18}$$

where I_n is the nth-order modified Bessel function of the first kind, σ is nondimensionalised by the normoxic oxygen level in non-pathological tissue, $L = 100\,\mu\text{m}$, and $\sigma_{\text{H}} = 0.2$. The mean value of the oxygen solution in the viable rim ($R_{\text{duct}} - T < r < R_{\text{duct}}$) is given explicitly by

$$\langle \sigma \rangle = \left(\frac{2L\sigma_{\text{H}}}{2R_{\text{duct}}T - T^2} \right) \left(\frac{R_{\text{duct}} I_1 \left(\frac{R_{\text{duct}}}{L} \right) - (R_{\text{duct}} - T) I_1 \left(\frac{R_{\text{duct}} - T}{L} \right)}{I_0 \left(\frac{R_{\text{duct}} - T}{L} \right)} \right). \tag{19}$$

For the duct in F3,

13
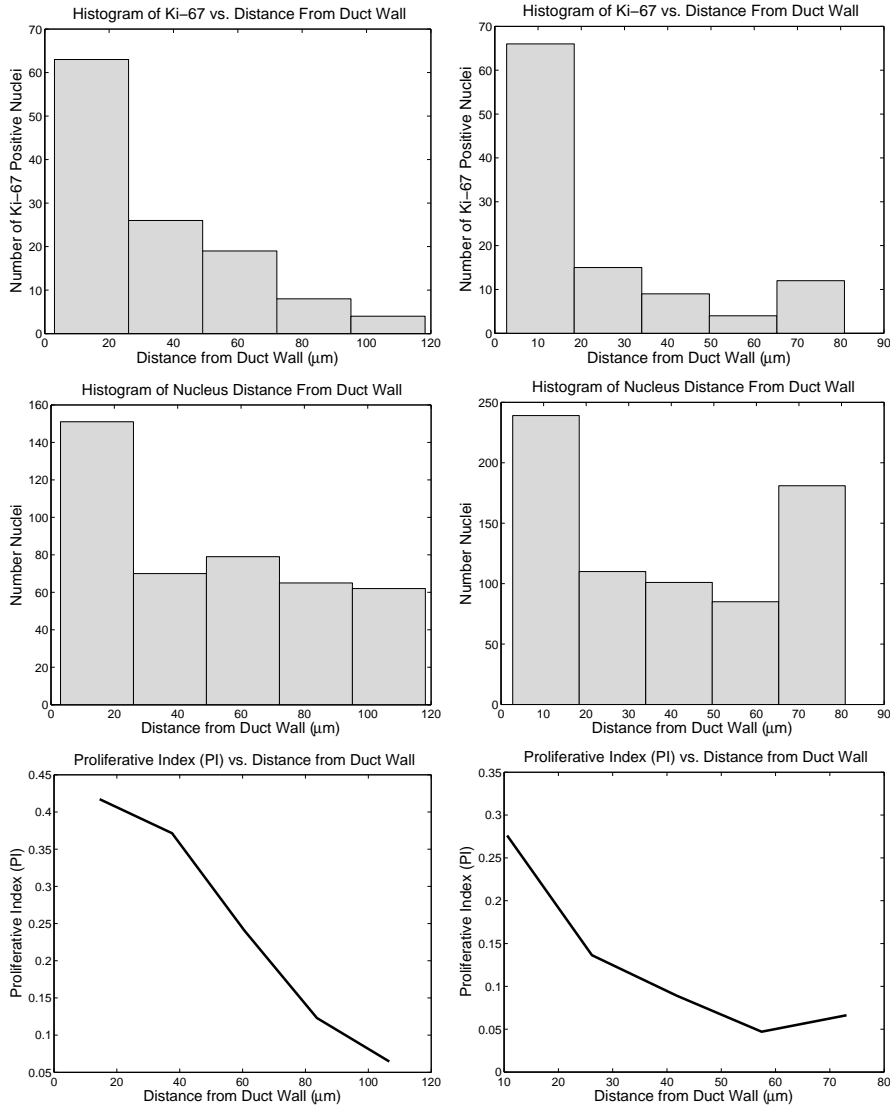
Fig. 3. Histograms of Ki-67 positive nuclei vs. distance from duct wall (top row), all nuclei vs. distance from duct wall (middle row), and proliferative index vs. distance from the duct wall (bottom row). *Left column:* Duct F3. *Right column:* Duct F19. An advance copy of this figure appeared in Macklin et al. (2010a).

$$R_{\mathrm{duct}} \approx 188.4634 \, \mu m, \qquad T \approx 119.0256 \, \mu m, \qquad \text{and} \qquad \langle \sigma \rangle \approx 0.282145,$$

and for the duct in F19,

$$R_{\mathrm{duct}} \approx 217.5548 \, \mu m, \qquad T \approx 97.9602 \, \mu m, \qquad \text{and} \qquad \langle \sigma \rangle \approx 0.280459.$$

By correlating the oxygen solutions with the PI profiles, we estimate the relationship between the measured PI and σ in the ducts. We plot these curves for F3 (dashed curve) and F19 (dotted curve) against the predicted curve (solid curve) from Macklin et al. (2009a) in Fig. 4. The theoretical predictions and
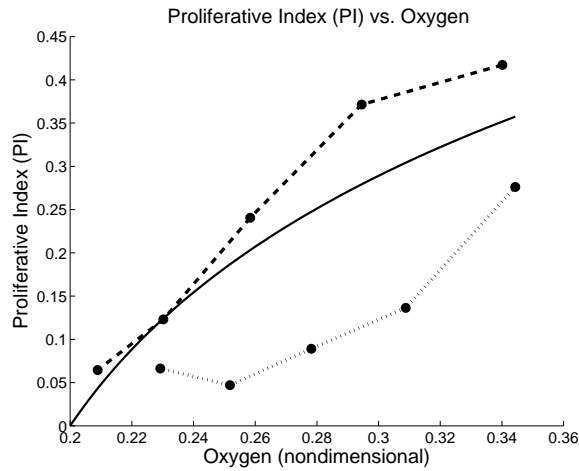
Fig. 4. Comparison of the predicted PI curve (solid curve) with data from duct F3 (dashed curve) and duct F19 (dotted curve) for case 100019. An advance copy of this figure appeared in Macklin et al. (2010a).

measurements agree qualitatively but not quantitatively. We conclude that while proliferation correlates with oxygen levels throughout the tumour, oxygenation alone cannot fully determine PI. Hence, there must be additional heterogeneities in other microenvironmental factors (e.g., EGF), gene expression, or protein signalling across the tumour.
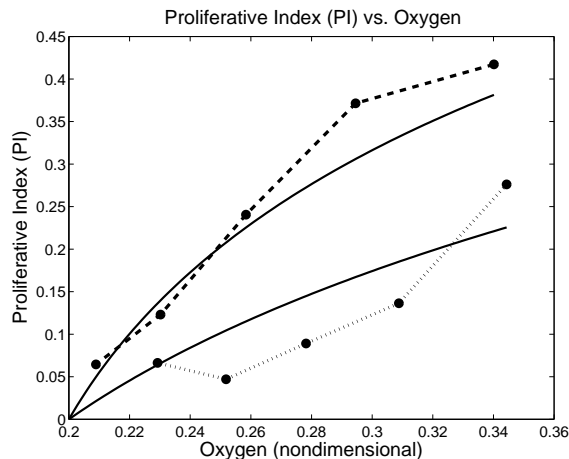
Fig. 5. Comparison of the hypothesised (solid) and measured (dashed and dotted) PI vs. σ curves for duct F3 (dashed) and duct F19 (dotted). An advance copy of this figure appeared in Macklin et al. (2010a).

The next natural question is whether we can account for these heterogeneities with our current functional form by applying the same analysis to the individual ducts. We use AI $= 0.008838$ in each duct, and PI, R_{duct}, and T as measured separately for each duct above. For the duct in F3,

15

$$\mathrm{PI} = 0.281030, \qquad \alpha_{\mathrm{A}} \approx 0.00162405 \, \mathrm{h}^{-1},$$
$$\langle \alpha_{\mathrm{P}} \rangle \approx 0.0277579 \, \mathrm{h}^{-1}, \qquad \text{and} \qquad \overline{\alpha}_P(\mathcal{S}, \bullet) \approx 0.270331 \, \mathrm{h}^{-1};$$

and for the duct in F19,

$$\mathrm{PI} = 0.148045, \qquad \alpha_{\mathrm{A}} \approx 0.00129067 \, \mathrm{h}^{-1},$$
$$\langle \alpha_{\mathrm{P}} \rangle \approx 0.0110190 \, \mathrm{h}^{-1}, \qquad \text{and} \qquad \overline{\alpha_{\mathrm{P}}}(\mathcal{S}, \bullet) \approx 0.109562 \, \mathrm{h}^{-1}.$$

Using this, we generate PI-vs-σ curves for the individual ducts based upon Eq. 14 and compare them to the measured data in Fig. 5. There is generally much improved quantitative agreement between the predicted (solid) and measured (dashed and dotted) curves. The difference in the predicted curves for the two ducts is due to the substantial difference in $\overline{\alpha}_P$: $\overline{\alpha}_P$ is much greater for F3, which has the overall higher PI curve.

We next examine the data in the ducts (Fig. 2) within the context of our modelling framework and the predicted PI-vs-σ curves to generate additional biological hypotheses. Notice that the cell density is lower in F3 (Fig. 2 left: larger nuclei with greater spacing between cells) than in F19 (Fig. 2 right: smaller nuclei with less spacing between cells). These lead us to hypothesise that $\overline{\alpha}_P$ decreases with increasing cell density. E-cadherin/β-catenin signalling may be the physiological explanation of the phenomenon: when E-cadherin is bound to E-cadherin on a neighbouring cell, β-catenin binds to the phosphorylated receptors, blocking its downstream pro-proliferative activity. (See Section 1.) For higher cell densities, more cell surfaces are in contact with each other, providing greater opportunities for E-cadherin binding; we consequently hypothesise that cell density correlates with cell cycle blockade by the E-cadherin/β-catenin pathway, resulting in the apparent relationship between cell density and $\overline{\alpha}_P$. Further evidence can be seen in duct F19 (Fig. 2, right): the majority of the proliferation activity is in a single layer of cells along the duct wall. Because these cells are adhered to the basement membrane, they present less surface for E-cadherin binding activity (relative to the interior cells), resulting in reduced E-cadherin blockade of proliferation.

These hypotheses can be tested by correlating $\overline{\alpha}_P$ with cell density in a larger number of ducts, performing IHC for β-catenin activity, and correlating β-catenin-mediated transcription (indicated by presence of β-catenin in the nuclei) with cell density and distance from the duct wall. One could use these data to hypothesise, calibrate, and test new functional forms for α_{P}, such as:

$$\alpha_{\mathrm{P}}(\mathcal{S}, \sigma, \bullet, \circ) = \overline{\alpha}_P(\bullet, \circ) \left(1 - \mathcal{E} \langle \mathcal{E} \rangle \frac{\rho}{\rho_{\max}} \right) \left(\frac{\sigma - \sigma_{\mathrm{H}}}{1 - \sigma_{\mathrm{H}}} \right), \qquad (20)$$

where ρ is the local cell density, PI ≈ 0 when $\rho = \rho_{\max}$, $\mathcal{E}$ is the cell's (nondimensional) E-cadherin expression, and $\langle \mathcal{E} \rangle$ is the tumour's mean E-cadherin

16

expression. In such a formulation, $\overline{\alpha}_P(\bullet, \circ)$ determines the cell's $\mathcal{Q} \to \mathcal{P}$ transition rate in normoxic conditions with minimal E-cadherin signalling.

5.1 Image processing algorithm

To process each image, we hand-annotated each nucleus with a single red pixel (RGB (255,0,0) for Ki-67 positive nuclei) or single green pixel (RGB (0,255,0) for Ki-67 negative nuclei). We hand-drew a magenta curve (RGB (255,0,255)) around the duct outer boundary, and a blue curve (RGB (0,0,255)) around the peri-necrotic boundary.

We then implemented in C++ a simple image processing routine to analyse the manually-annotated plot. The code first creates lists of the coordinates $(x, y) \in \mathbb{N}^2$ of the annotated pixels. ($0 \leq x < M$ and $0 \leq y < N$, where the image is M pixels wide and N pixels tall.) For any set $\mathcal{X}$, denote $\#\mathcal{X}$ to be the number of points in $\mathcal{X}$.

The algorithm creates lists $\mathcal{R}$, $\mathcal{G}$, $\mathcal{B}$, and $\mathcal{M}$ of the red, green, blue, and magenta pixel coordinates from the manual annotation. The remainder of the algorithm works on evaluating geometric distances amongst members of these sets. Note that for any $\mathbf{x} \in \mathcal{X}$, the distance from $\mathbf{x}$ to the set $\mathcal{Y}$ is given by $\min_{\mathbf{y} \in \mathcal{Y}} |\mathbf{x} - \mathbf{y}|$, where we use the standard Euclidean distance.

With these notions in mind:

(1) We compute
$$\mathbf{x}_{\text{centroid}} = \frac{1}{\#\mathcal{M}} \left(\sum_{(x,y) \in \mathcal{M}} x, \sum_{(x,y) \in \mathcal{M}} y \right),$$
the centroid of the set $\mathcal{M}$. This is the duct centre.

(2) We compute
$$\frac{1}{\#\mathcal{M}} \sum_{\mathbf{x} \in \mathcal{M}} |\mathbf{x} - \mathbf{x}_{\text{centroid}}|.$$
This is the mean duct radius R (in pixels).

(3) We compute
$$\frac{1}{\#\mathcal{B}} \sum_{\mathbf{x} \in \mathcal{B}} \min_{\mathbf{y} \in \mathcal{M}} |\mathbf{x} - \mathbf{y}|.$$
This is the mean thickness of the viable rim T.

(4) For each $\mathbf{r} \in \mathcal{R}$, we compute
$$d(\mathbf{r}) = \min_{\mathbf{y} \in \mathcal{M}} |\mathbf{r} - \mathbf{y}|,$$
the distance from the Ki-67 positive nucleus and the duct wall.

(5) For each $\mathbf{r} \in \mathcal{G}$, we compute

$$d(\mathbf{g}) = \min_{\mathbf{y} \in \mathcal{M}} |\mathbf{g} - \mathbf{y}|,$$

the distance from the Ki-67 negative nucleus and the duct wall.

We then store the distances in 4 and 5 to simple text files, which MATLAB can process using standard histogram procedures. Because these distances are given in pixel units, we must convert to physical units according the image scaling. For these images, 200 pixels correspond to 930 microns.

Complete numerical code (C++ for the image processing, and MATLAB for creating the histograms and figures), unannotated images, and annotated images are provided at MathCancer.org.[8].

6   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. The full list of non-specific parameters and their physical meanings is given in Table 1.

6.1   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)).

6.2   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

Parameter	Physical Meaning	Value	Section
τ_P	cell cycle time	18 hours	6.1
τ_{G1}	length of G_1	9 hours	6.1
τ_A	apoptosis time	8.6 hours	6.2
τ_{NL}	necrotic cell lysis time	6 hours	6.3
f_{NS}	necrotic cell volume increase	1.0	6.3
τ_C	necrotic debris calcification time	15 days	6.4
L	oxygen diffusion length scale	100 μm	6.5
$\langle \lambda \rangle$ $(= \lambda_p = \lambda_{np})$	mean tumour cell oxygen uptake rate	0.1 min^{-1}	6.5
λ_b	oxygen uptake/decay rate for non-viable cells and background	$0.01\langle \lambda \rangle$	6.5
σ_H	hypoxic oxygen threshold	0.2	6.5
R_A	maximum adhesion distance	$1.214\,R$	6.6
c_{ccr}	cell-cell repulsive force coefficient	10.0ν μm/min	6.6
c_{cbr}	cell-BM repulsive force coefficient	c_{ccr}	6.6
n_{cca}	cell-cell adhesion potential exponent	1	6.6
n_{ccr}	cell-cell repulsion potential exponent	1	6.6
n_{cba}	cell-BM adhesion potential exponent	n_{cca}	6.6
n_{cbr}	cell-BM repulsion potential exponent	n_{ccr}	6.6
M	maximum value of $|\nabla \psi|$	1	6.6

Table 1
Patient-independent parameters and values for DCIS.

apoptotic bodies are eliminated in under 48 hours after cell lysis. In total, the experimental observations in the literature lead us to estimate $\tau_A \approx \mathcal{O}(10\text{h})$.

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

$$\dot{N} = \left(\frac{1}{\tau_P}\text{PI} - \frac{1}{\tau_A}\text{AI} \right) N, \tag{21}$$

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

$$\tau_A \text{ PI} = \tau_P \text{ AI}. \tag{22}$$

In Lee et al. (2006b), the mean proliferative and apoptotic indices of non-cancerous 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 with the experimental literature; see Macklin et al. (2010a).

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:

$$AI_{actual} \approx \frac{8.6}{5.6} AI_{TUNEL}, \qquad \text{and} \qquad (23)$$

$$\frac{8.6}{7.6} AI_{Caspase-3} \leq AI_{actual} \leq \frac{8.6}{6.6} AI_{Caspase-3}. \qquad (24)$$

6.3 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

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.

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 $\tau_{NL} < \tau_A$ (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 $\tau_{NL} = 6$ 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 ($1 \leq f_{NS} \leq 4$) after 24 hours in rat adrenal medulla cells. We use $f_{NS} = 1$ 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.

6.4 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 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 sim-

plified 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.

6.5   Oxygen transport parameters

By Owen et al. (2004), the oxygen diffusion length scale L is $\sim 100\,\mu$m, and the mean cellular oxygen uptake rate $\langle\lambda\rangle$ (in the viable rim) can be estimated at approximately 0.1 min^{-1} via $L = \sqrt{D/\langle\lambda\rangle}$ 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 min$^{-1} \leq \langle\lambda\rangle \leq 10$ min^{-1}. This does not majorly impact our results because (1) we calibrate the proliferation and oxygenation sub-models in a self-consistent manner, and (2) $\langle\lambda\rangle$ 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 $\lambda_p \sim \lambda_{np}$. 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 set $\lambda_p = \lambda_{np}$ and $\lambda_b = 0.01\langle\lambda\rangle$. In the supplementary material, we show that setting $\lambda_p/\lambda_{np} > 1$ destabilises the perinecrotic border, and is not consistent with typical patient histopathology.

22

distance from cell center r (μm)	$\frac{c_{\text{ccr}}}{\nu}\|\nabla\psi(r)\|$ (μm/min)	distance from cell center r (μm)	$\frac{c_{\text{ccr}}}{\nu}\|\nabla\psi(r)\|$ (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).

6.6   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 μm $\leq R_{\text{N}} \leq 7$ μm), high-grade are 2.5 RBCs or more ($R_n \geq 7.5$ μm); 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 $R_{\text{A}} = 1.214R$.

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 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 $(R-3)\,\mu\mathrm{m}\ <r<(R-0.5)\,\mu\mathrm{m}$; see Table 3. This is consistent with Macklin et al. (2009a), where setting $c_{\mathrm{ccr}}/\nu = 8\,\mu\mathrm{m/min}$ and $c_{\mathrm{cbr}}/\nu = 5\,\mu\mathrm{m/min}$ prevents unreasonable simulation behaviour (overlapping cell nuclei and cell penetration of the BM). In the supplementary material, 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 $M = 1$, $c_{\mathrm{ccr}} = c_{\mathrm{cbr}} = 10\nu\,\mu\mathrm{m/min}$, and $n_{\mathrm{ccr}} = n_{\mathrm{cbr}} = 1$ (to model anticipated nonlinear but smooth cell mechanical responses).

7   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, R_A, recorded as multiples 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 morphology with an auxiliary level set function, and we use an auxiliary data structure to reduce the computational cost of cell-cell interaction testing and evaluation. (See Section 7.1.)

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 immunohisto-

24

chemistry). 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 $\Delta x = \Delta y = 1\ \mu\mathrm{m}$) 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 7.1.

   (c) **Update the cells:** For each cell:
      (i) *Progress the current cell state:* Update the cell with the appropriate 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 7.2. We note that the probabilities can likely be approximated by their linear Taylor expansions:

      $$\Pr\left(\mathcal{S}(t + \Delta t) = \mathcal{S}_2 | \mathcal{S}(t) = \mathcal{S}_1\right) \approx \alpha_{12}(\mathcal{S}(t), \bullet, \circ)\Delta t. \qquad (25)$$

      we are testing this acceleration among other ongoing code optimisations.
      (iii) *Set the cell velocity:* Set $\mathbf{v}$ according to Eq. 4. Use the optimi-

25

sation in Section 7.1 to truncate the summation to a smaller set of interacting cells.

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

$$\Delta t = \frac{\epsilon}{\max\left\{|\mathbf{v}_i|\right\}_{i=1}^{N(t)}}. \tag{26}$$

Here, ϵ is the desired accuracy in the cell position; we use $\epsilon = 1\ \mu\text{m}$. 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:

$$\mathbf{x}(t) = \mathbf{x}(t - \Delta t) + \mathbf{v}\Delta t \tag{27}$$

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. 4, thus rendering the overall algorithm linear in the number of cells. See Section 7.1.

7.1 Accelerated cell-cell interaction testing

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 $\Delta x = \Delta y = 1\ \mu\text{m}$. 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

$$R_{\text{cca,max}} = \max\left\{R_A^k\right\}_{k=1}^{N(t)}. \tag{28}$$

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:

$$|\mathbf{x}_k - \mathbf{x}_{i,j}| \leq R_{\text{cca,max}} + R_{\text{cca}}^k, \tag{29}$$

set:

$$k_{N_{i,j}(t)+1}^{i,j} = k \qquad \text{(append the cell to the list at } \mathbf{x}_{i,j}) \tag{30}$$
$$N_{i,j}(t) = N_{i,j}(t) + 1. \qquad \text{(increment the total at } \mathbf{x}_{i,j}) \tag{31}$$

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

$$\text{for all cells } k \in \{k\}_{k=1}^{N(t)} \setminus \{\ell\} \text{ compute } f\left(\text{cell}_k, \text{cell}_\ell\right)(\mathbf{x}), \tag{32}$$

such as

$$\sum_{\substack{k=1 \\ k \neq \ell}}^{N(t)} f\left(\mathbf{x}_k, \mathbf{x}_\ell\right). \tag{33}$$

Let $\mathbf{x}_{i,j}$ denote the closest interaction mesh point to $\mathbf{x}_\ell$ (the position of cell ℓ). Then we evaluate Eq. 32 by truncating it to the members of the list at $\mathbf{x}_{i,j}$:

$$\text{for all cells } k \in \left\{k_m^{i,j}\right\}_{m=1}^{N_{i,j}(t)} \setminus \{\ell\} \text{ compute } f\left(\text{cell}_k, \text{cell}_\ell\right)(\mathbf{x}). \tag{34}$$

In the example above, we truncate the summation to

$$\sum_{\substack{m=1 \\ k_m^{i,j} \neq \ell}}^{N_{i,j}(t)} f\left(\mathbf{x}_{k_m^{i,j}}, \mathbf{x}_\ell\right). \tag{35}$$

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. 34 for all cells $1 \leq \ell \leq N(t)$ is linear in the number of

cells. Contrast this with Eq. 32, which for each cell ℓ scales with $N(t)$; iterating this non-truncated form over all cells thus requires $N(t)^2$ computational effort.

7.2 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.

8 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 very 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 readily 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 the very latest standards, documentation, and software utilities. We put forth our data format as a

---

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

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 academic citation (`<reference>`). See Fig. 6. 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. 7. 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. 8. 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

```xml
<?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. 6. **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.

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.

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

```
    ...
<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. 7. **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. 8. **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).

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.

8.1 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

---

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

    (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.

8.2 Sample post-processing

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 the `MultiCellXML` project website; 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 $\{(x, y) : x \leq x_{V95}\}$). 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.

8.3 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 basement membrane based upon the zero contour of the level set function. Afterwards, we 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 the `MultiCellXML` project website; 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.

9   Additional numerical studies

We performed additional numerical studies, which we cut from the main manuscript but nonetheless support and/or further investigate the model, and may be of interest to the reader.

9.1   *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, G_2, and M phases), $P_{\mathrm{G_1}}$ (cycling cells in G_1 phase), A (apoptotic cells), and Q (quiescent cells in G_0):

All measurements given as mean ± standard deviation

Quantity	Patient Data	Simulated: $\tau_{G1} = 9$ hr	Simulated: $\tau_{G1} = 1$ 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.

$$\dot{P}_{SG_2M} = \langle \alpha_P \rangle Q - \frac{1}{\tau_P - \tau_{G1}} P_{SG_2M} \tag{36}$$

$$\dot{P}_{G_1} = \frac{2}{\tau_P - \tau_{G1}} P_{SG_2M} - \frac{1}{\tau_{G1}} P_{G_1} \tag{37}$$

$$\dot{A} = \alpha_A Q - \frac{1}{\tau_A} A \tag{38}$$

$$\dot{Q} = \frac{1}{\tau_{G1}} P_{G_1} - (\langle \alpha_P \rangle + \alpha_A) Q \tag{39}$$

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

$$\mathrm{PI} = \frac{P}{N}, \qquad \text{and} \qquad \mathrm{AI} = \frac{A}{N}. \tag{40}$$

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 simulated PI and AI should approach 24.45% and 0.761%, respectively; both of these limits match our simulated mean PI (24.04%) and AI (0.738%) very well.

Conversely, setting $\tau_{G1} = 1$ min minimises the impact of the 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 volume growth (G_1) and non-growth (S-G_2-M) phases. This provides new potential for matching bimodal distributions of cell sizes and cell cycle states

often observed *in vitro* (e.g., Tzur et al. (2009)). We plan improvements to our calibration based upon these observations in ongoing work.

9.2 Testing biological hypotheses: heterogeneity in PI is not solely due to oxygenation gradients; apoptosis is a "background" process

Because our calibration protocol only prescribed the mean behaviour, we can use the simulated variation to test the model's underlying biological hypotheses. First, we note that the simulated variation in PI is significantly smaller than the actual variation; this suggests that our current model (which varies proliferation solely with the local level of oxygenation) does not fully characterise the signalling mechanisms responsible for heterogeneity in DCIS proliferation. This is therefore additional confirmation of our hypothesis that while oxygenation dominates the variation in proliferation, other signalling mechanisms (e.g., contact inhibition by E-cadherin/β-catenin pathways) play a significant role in controlling DCIS proliferation as well.

Continuing, we assumed that apoptosis occurs at a low "background" rate that is independent of oxygenation and any other signalling, and hence occurs with probability that is independent of position within the viable duct. We see that our standard deviation in AI is quite similar to the actual standard deviation in the patient AI, which supports this biological hypothesis. In fact, the patient's mean and standard deviation are of comparable magnitude, which is consistent with an exponentially-distributed random variable, just as in our model.

9.3 Robustness of the mechanics parameters

In Section 6.6, 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 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. 9: 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
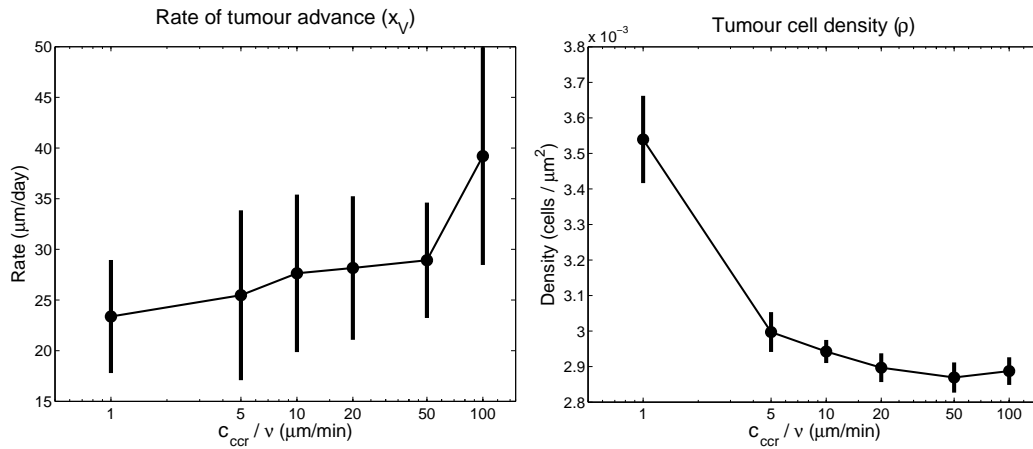
Fig. 9. **Robustness of the mechanics parameters:** We varied the cell-cell repulsive force c_{ccr}/ν while maintaining the relative balance of the forces; this is equivalent to varying the biomechanics time scale. For c_{ccr}/ν within an order of magnitude of our initial estimate, both the tumour front 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.

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. 9: 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. This is consistent with our earlier results that the cell-cell repulsion parameter must be on the order of 10 to prevent overlapping cells.

9.4 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. See Fig. 10: left for a typical example (at 15 days) with $\frac{c_{\mathrm{cca}}}{c_{\mathrm{cba}}} = 1$. In 3D, the curvature of the duct may reduce this effect.

On the other hand, for $\frac{c_{\mathrm{cca}}}{c_{\mathrm{cba}}} = 0.01$, it was very difficult for daughter cells to
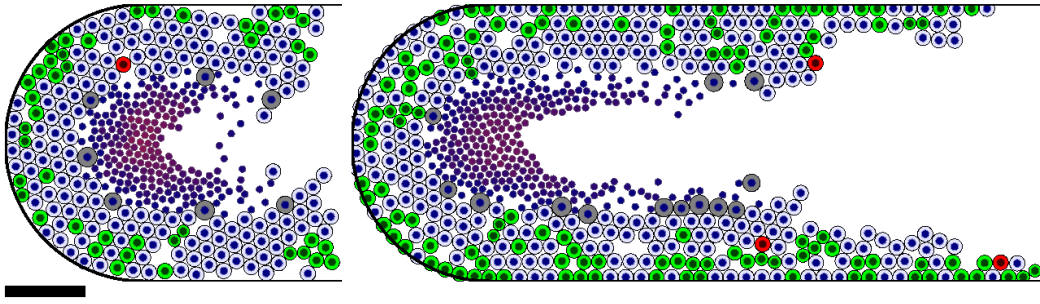
Fig. 10. **Impact of the ratio of cell-cell and cell-BM adhesive forces** $c_{\mathrm{cca}}/c_{\mathrm{cba}}$: <u>Left</u>: For $c_{\mathrm{cca}}/c_{\mathrm{cba}} = 1$ and greater, cells easily pull away from the duct wall except in regions of densely packed duct. <u>Right</u>: For $c_{\mathrm{cca}}/c_{\mathrm{cba}} = 0.01$, cells get trapped near the BM, leading to a "wetting" effect that accelerates the tumour's advance through the duct. High oxygenation near the duct wall acts as nonlinear feedback to accelerate the process. Cells are coloured as in the main text. <u>Bar</u>: 100 μm. A colour version of this figure is available in the online edition.

push away from the BM after proliferation, leading to a "wetting" effect along the duct wall similar to a fluid capillary force. See Fig. 10: right (plotted at 15 days). This increased the percentage of cells near the high-oxygen regions of the duct, which acted as a feedback that accelerates the tumour's advance through the duct at unrealistic rates. See Video S2. Because the cell-BM adhesive force is modelled as normal to the BM (thereby neglecting any tangential component), the BM is effectively frictionless, further exacerbating this effect.

9.5   *Amount of necrotic cell swelling primarily influences the gap*

We varied the level of necrotic cell swelling $f_{\mathrm{NS}} \in \{0\%, 30\%, 100\%\}$ and found virtually no impact on the least-squares fit of the rate of tumour advance from 15 to 30 days. (24.78, 24.38, and 24.25 μm/day, respectively.) Instead, the primary impact was to increase the size of the physical gap, with increases scaling roughly as $(1 + f_{\mathrm{NS}})^{\frac{1}{3}}$ (result not shown).

9.6   *Heterogeneous oxygen uptake causes perinecrotic boundary instability*

In Section 6.5, we estimated that proliferating and non-proliferating tumour cells uptake oxygen at approximately the same rate, with $\lambda_{\mathrm{p}} = \lambda_{\mathrm{np}}$. We close with an investigation of $\frac{\lambda_{\mathrm{p}}}{\lambda_{\mathrm{np}}} \in \{1, 10, 100\}$. In each simulation, we maintained $\langle \lambda \rangle = \mathrm{PI}\lambda_{\mathrm{p}} + (1 - \mathrm{PI})\lambda_{\mathrm{np}}$ constant. The results at time 30 days are plotted in Fig. 11. As $\frac{\lambda_{\mathrm{p}}}{\lambda_{\mathrm{np}}}$ is increased, the stability of the perinecrotic boundary is reduced, with greater mixing of necrotic cellular debris and viable tumour cells. This occurs because high oxygen uptake by isolated proliferating cells
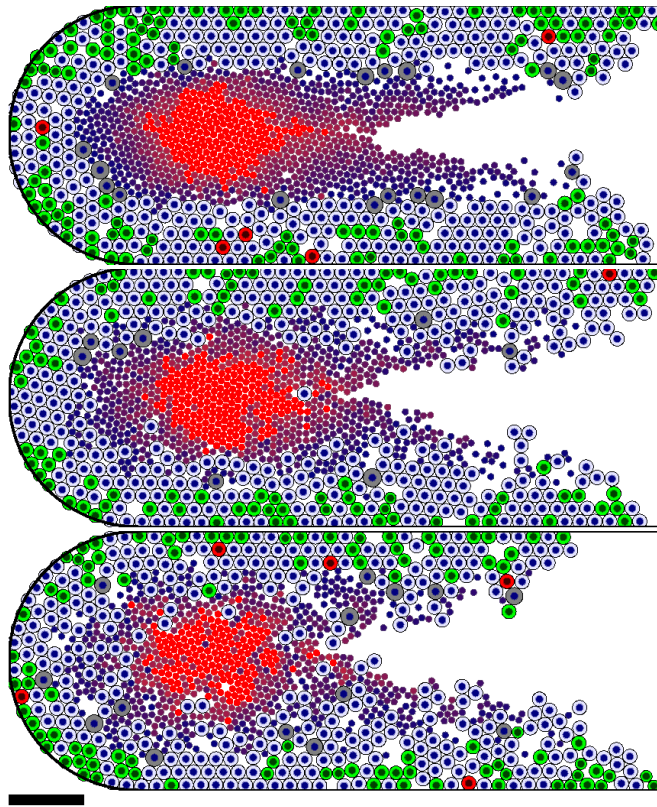
39

Fig. 11. **Impact of heterogeneous oxygen uptake rates:** As the ratio $\lambda_{\mathrm{p}}/\lambda_{\mathrm{np}}$ of the oxygen uptake rates by the proliferating (λ_{p}) and nonproliferating cells (λ_{np}) is increased from 1 (**top**) to 10 (**middle**) and 100 (**bottom**), the stability of the perinecrotic boundary is reduced, with greater mixing of necrotic cellular debris and viable cells. These morphologies are not typical in DCIS, supporting the estimate that $\lambda_{\mathrm{p}} \approx \lambda_{\mathrm{np}}$. Cells are coloured as in the main text. <u>Bar</u>: 100 μm. A colour version of this figure is available in the online edition.

creates small pockets of hypoxia between these cells and the necrotic core. See Video S3. The result is a ragged perinecrotic boundary not typically observed in DCIS, further supporting the estimate that $\lambda_{\mathrm{p}} \approx \lambda_{\mathrm{np}}$.

However, it is interesting to observe that such an instability can result from microscopic variations in cell metabolism caused by cell-induced, fine-scale alterations in the tumour microenvironment; such instabilities are often attributed to variations in cellular adhesion. While the result here is likely non-physical for oxygen transport, similar behaviour could occur in glucose transport, where glucose uptake is much greater for hypoxic cells than non-hypoxic cells (Smallbone et al., 2007b,a; Gatenby et al., 2007).
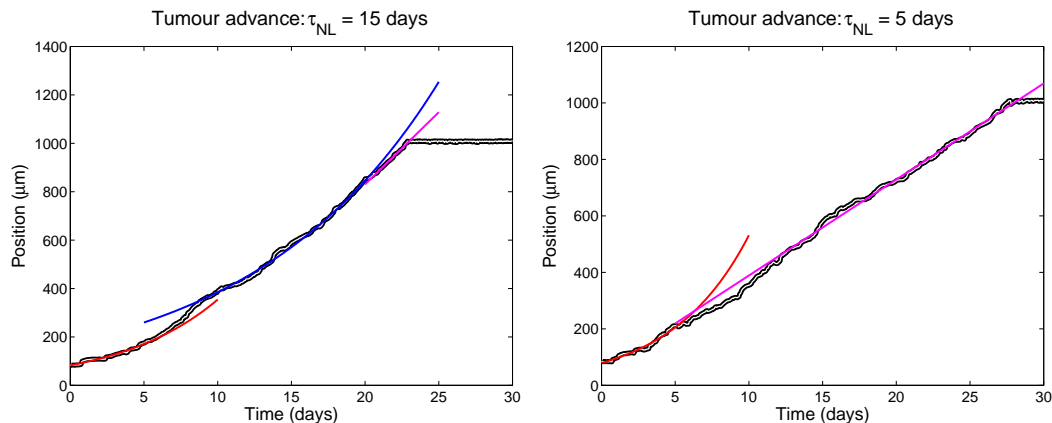
Fig. 12. The tumour advances exponentially (red and blue curves) until the viable rim is well-established and cell lysis has begun. Thereafter, the tumour advances through the duct at a constant rate (magenta curves).

9.7 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. 12: 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. 12: 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 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. 12: 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. 12: 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.12: 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
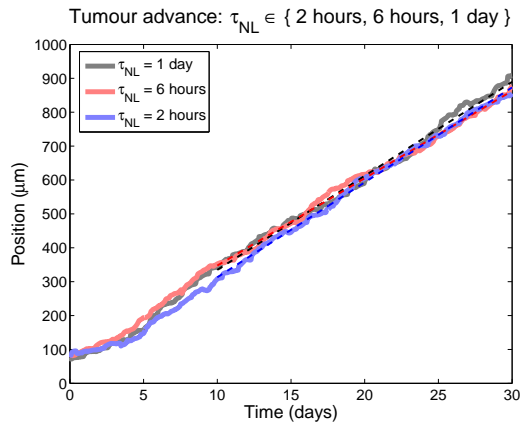
41

Fig. 13. 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).

the magenta fitted line in Fig. 12: 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. 13. Note that all three tumours advance at approximately the same rate.

10   Simulation Animations

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 ($\mathcal{S} = \mathcal{N}$)
- Red circles: necrotic cellular debris after lysis ($\mathcal{S} = \mathcal{N}$); shade of red indicates the degree of calcification
- Bright red circles: clinically-detectable calcified cellular debris ($\mathcal{S} = \mathcal{C}$)

42

The movies are in AVI format and compressed with the Xvid codec. 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:** "wetting" behaviour when cell-BM adhesion is strong relative to cell-cell adhesion ($c_{\mathrm{cba}} = 100c_{\mathrm{cca}}$; see Section 9.4), plotted from 0 to 30 days in 1 mm of duct.
Alternate format: `http://www.youtube.com/watch?v=9q9LGzX9fok`.

(3) **Video S3:** unstable perinecrotic boundary (between the viable rim and the necrotic core) resulting from heterogeneous cellular oxygen uptake rates ($\lambda_{\mathrm{p}} = 100\lambda_{\mathrm{np}}$; see Section 9.6), plotted from 0 to 30 days in 1 mm of duct.
Alternate format: `http://www.youtube.com/watch?v=Brgw8qI8k-k`.

References

T. L. Adamovich and R. M. Simmons. Ductal carcinoma in situ with microinvasion. *Am. J. Surg.*, 186(2):112–6, 2003.

L. Ai, W.-J. Kim, T.-Y. Kim, C. R. Fields, N. A. Massoll, K. D. Robertson, and K. D. Brown. Epigenetic silencing of the tumor suppressor cystatin m occurs during breast cancer progression. *Canc. Res.*, 66(16):7899–909, 2006.

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.

A. R. A. Anderson and V. Quaranta. Integrative mathematical oncology. *Nat. Rev. Canc.*, 8(3):227–44, 2008.

E. Anderson. Cellular homeostasis and the breast. *Maturitas*, 48(S1):13–7, 2004.

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. O. Baxter, K. Neoh, and M. C. Tevendale. The beginning of the end: Death

signaling in early involution. *J. Mamm. Gland Biol. Neoplas.*, 12(1):3–13, 2007.

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.

L. M. Butler, S. Khan, G. E. Rainger, and G. B. Nash. Effects of endothelial basement membrane on neutrophil adhesion and migration. *Cell. Immun.*, 251(1):56–61, 2008.

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.

S. Ciatto, S. Bianchi, and V. Vezzosi. Mammographic appearance of calcifications as a predictor of intraductal carcinoma histologic subtype. *Eur. Radiology*, 4(1):23–6, 1994.

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

M. Conacci-Sorrell, J. Zhurinsky, and A. Ben-Zeév. The cadherin-catenin adhesion system in signaling and cancer. *J. Clin. Invest.*, 109(8):987–91, 2002.

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.

C. G. Danes, S. L. Wyszomierski, J. Lu, C. L. Neal, W. Yang, and D. Yu. 14-3-3ζ down-regulates p53 in mammary epithelial cells and confers luminal filling. *Canc. Res.*, 68(6):1760–7, 2008.

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, L. Zhang, J. Yoon, and J. Costa. In silico cancer modeling: is it ready for prime time? *Nat. Clin. Practice Oncol.*, 6(1):34–42, 2009.

T. S. Deisboeck, Z. Wang, P. Macklin, and V. Cristini. Multiscale cancer modeling. *Annu. Rev. Biomed. Eng.*, 13, 2011. (in press).

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.

M. E. Edgerton, Y.-L. Chuang, P. Macklin, W. Yang, E. L. Bearer, and V. Cristini. A novel, patient-specific mathematical pathology approach for assessment of surgical volume: Application to ductal carcinoma in situ of the breast. *Anal. Cell. Pathol.*, 2011. URL `http://mathcancer.org/Publications.php#edgerton11_ACP`. (in press).

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.

B. Erbas, E. Provenzano, J. Armes, and D. Gertig. The natural history of ductal carcinoma *in situ* of the breast: a review. *Breast Canc. Res. Treat.*, 97(2):135–44, 2006.

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.

R. A. Gatenby, K. Smallbone, P. K. Maini, F. Rose, J. Averill, R. B. Nagle, L. Worrall, and R. J. Gillies. Cellular adaptations to hypoxia and acidosis during somatic evolution of breast cancer. *Br. J. Cancer*, 97(5):646–53, 2007.

F. G. Giancotti and E. Ruoslahti. Integrin signaling. *Science*, 285(5430): 1028–32, 1999.

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.

J. J. Going and T. J. Mohun. Human breast duct anatomy, the 'sick lobe' hypothesis and intraductal approaches to breast cancer. *Breast. Canc. Res. and Treat.*, 97(3):0167–6806, 2006.

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.

R. K. Hansen and M. J. Bissell. Tissue architecture and breast cancer: the role of extracellular matrix and steroid hormones. *Endocrine-Related Cancer*, 7 (2):95–113, 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.

D. Ilić, E. A. Almeida, D. D. Schlaepfer, P. Dazin, S. Aizawa, and C. H. Damsky. Extracellular matrix survival signals transduced by focal adhesion kinase suppress p53-mediated apoptosis. *J. Cell Biol.*, 143(2):547–60, 1998.

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.

T. X. Jiang and C. M. Chuong. Mechanism of skin morphogenesis I: Analyses with antibodies to adhesion molecules tenascin, NCAM, and integrin. *Dev. Biol.*, 150(1):82–98, 1992.

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.

S. Khan, M. Rogers, K. Khurana, M. Meguid, and P. Numann. Estrogen receptor expression in benign breast epithelium and breast cancer risk. *J. Natl. Canc. Inst.*, 90(1):37–42, 1998.

S. Khan, A. Sachdeva, S. Naim, M. Meguid, W. Marx, H. Simon, et al. The normal breast epithelium of women with breast cancer displays an aberrant response to estradiol. *Canc. Epidemiol. Biomarkers Prev.*, 8:867–72, 1999.

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. J. Lucio, M. T. Faria, A. M. Pinto, M. R. da Silva, M. E. Correia Jr., R. J. da Costa, and A. B. Parreira. Expression of adhesion molecules in chronic B-cell lymphoproliferative disorders. *Haematologica*, 83(2):104–11, 1998.

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 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 adaptation 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.

D. F. Moffat and J. J. Going. Three dimensional anatomy of complete duct systems in human breast: pathological and developmental implications. *J. Clin. Pathol.*, 49(1):48–52, 1996.

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.

T. Ohtake, I. Kimijima, T. Fukushima, M. Yasuda, K. Sekikawa, S. Takenoshita, and R. Abe. Computer-assisted complete three-dimensional reconstruction of the mammary ductal/lobular systems. *Cancer*, 91(12): 2263–72, 2001.

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.

D. L. Page. The clinical significance of mammary epithelial hyperplasia. *Breast*, 1(1):3–7, 1992.

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.

K. Rennstam and I. Hedenfalk. High-throughput genomic technology in research and clinical management of breast cancer. molecular signatures of progression from benign epithelium to metastatic breast cancer. *Breast Canc. Res.*, 8(4):213ff, 2006.

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.

S. A. Silver and F. A. Tavassoli. Ductal carcinoma in situ with microinvasion. *Breast J.*, 4(5):344–8, 1998.

M. J. Silverstein. Ductal carcinoma in situ of the breast. *Annu. Rev. Med.*, 51(1):17–32, 2000.

P. T. Simpson, J. S. Reis-Filho, T. Gale, and S. R. Lakhani. Molecular evolution of breast cancer. *J. Pathol.*, 205(2):248–54, 2005.

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, 2007a.

K. Smallbone, D. J. Gavaghan, P. K. Maini, and J. M. Brady. Quiescence as a mechanism for cyclical hypoxia and acidosis. *J. Math. Biol.*, 55(5–6): 767–79, 2007b.

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`.

L. Sontag and D. E. Axelrod. Evaluation of pathways for progression of heterogeneous breast tumors. *J. Theor. Biol.*, 232(2):179–89, 2005.

T. A. Springer. Adhesion receptors of the immune system. *Nature*, 346(6283): 425–34, 1990.

P. C. Stomper and F. R. Margolin. Ductal carcinoma in situ: the mammographer's perspective. *Am. J. Roentgenology*, 162:585–91, 1994.

D. G. Stupack and D. A. Cheresh. Get a ligand, get a life: Integrins, signaling and cell survival. *J. Cell. Sci.*, 115(19):3729–38, 2002.

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.

P. J. Tannis, O. E. Nieweg, R. A. Valdés Olmos, and B. B. R. Kroon. Anatomy and physiology of lymphatic drainage of the breast from the perspective of sentinel node biopsy. *J. Am. Coll. Surg.*, 192(3):399–409, 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.

M. J. Terol, M. Tormo, J. A. Martinez-Climent, I. Marugan, I. Benet, A. Ferrandez, A. Teruel, R. Ferrer, and J. Garcia-Conde. Soluble intercellular adhesion molecule-1 (s-ICAM-1/s-CD54) in diffuse large B-cell lymphoma: association with clinical characteristics and outcome. *Ann. Oncol.*, 14(3): 467–74, 2003.

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

A. Tzur, R. Kafri, V. S. LeBleu, G. Lahav, and M. W. Kirschner. Cell growth and size homeostasis in proliferating animal cells. *Science*, 325(5937):167–71, 2009.

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

C. Wei, M. Larsen, M. P. Hoffman, and K. M. Yamada. Self-organization and branching morphogenesis of primary salivary epithelial cells. *Tissue Eng.*, 13(4):721–35, 2007.

S. R. Wellings, H. M. Jensen, and R. G. Marcum. An atlas of subgross pathology of the human breast with special reference to possible precancerous lesions. *J. Natl. Cancer Inst.*, 55(2):231–73, 1975.

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.

L. Zhang, Z. Wang, J. A. Sagotsky, and T. S. Deisboeck. Multiscale agent-based cancer modeling. *J. Math. Biol.*, 58(4–5):545–59, 2009.