

# **Mechanochemical and experimental models in mathematical biology**



**Ryan John Murphy**

MMath (Oxon)

Principal Supervisor: Professor Matthew Simpson

Associate Supervisor: Doctor Pascal Buerzli

School of Mathematical Sciences

Faculty of Science

Queensland University of Technology (QUT)

2021

Submitted in fulfillment of the requirements of the degree *Doctor of Philosophy*



*We're not playing games tonight, here we go*

*Peter Kay*



# Abstract

Experiments that probe epithelial tissue dynamics, cell competition, and tumour growth are fundamental to understand processes in developmental biology, cancer progression and cancer treatment. However, interpreting complex biological experiments is challenging. To resolve this challenge we develop and use a range of different types of mathematical models. In the first part of this thesis, we develop a mathematical framework for describing epithelial tissues that incorporates cell motion due to mechanical interactions, cell proliferation, cell death, epithelial-mesenchymal transitions via cell detachment at the tissue boundary, and mechanochemical coupling, all for heterogeneous cell populations. Our approach is to start with biologically-motivated discrete models and derive the corresponding continuum models, for both fixed and free boundary conditions. Applications to experimental studies are proposed and discussed. In the second part of this thesis, we collect novel experimental data from tumour spheroid experiments that we perform over a range of experimental designs. Importantly, we use real-time cell cycle imaging to reveal proliferation-inhibited and necrotic regions inside growing tumour spheroids. We then revisit the seminal Greenspan model describing tumour growth and use statistical analysis to reveal the experimental design choices that are important and lead to reliable biological insight. In doing so we connect Greenspan's model to data for the first time since its publication in 1972. All key code is freely available on GitHub repositories.



# **Statement of original authorship**

The work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

**Signature:**

A handwritten signature in black ink, appearing to read "R. Murphy".

**Date:** 3 September 2021.



# Acknowledgements

There are many who I would like to acknowledge and thank for their contribution to this work. Firstly, I would like to thank my supervisor **Professor Matthew Simpson** for giving me the opportunity to commence this PhD, his patience, thoughtful guidance and support throughout all of this work, providing opportunities for me to work in a experimental lab, reading of drafts, providing comments, and general advice. Secondly, I would like to thank **Doctor Pascal Buenzli** for his supervision, support throughout this work, reading many drafts, providing comments, and interesting discussions. Thirdly, I would like to thank **Professor Ruth Baker** for her long-standing support, reading drafts and providing comments. I would like to thank **Professor Rik Thompson** and his group for encouraging me join lab group meetings, gain further insights into experimental biology, providing collaborative opportunities, and his encouragement. **Doctor Honor Hugo** for interesting biology discussions and collaboration. I would like to thank **Professor Nikolas Haass** for enabling me to work in his experimental laboratory, providing great insights into biology and melanoma and interesting discussions. I would like to thank **Gency Gunesingh** for training me to perform experiments in the laboratory, from no experience to being able to perform experiments independently, and for her encouragement and interesting discussions. I would like to thank **Alex Browning** for his support throughout the PhD and help in the lab. **Tamara Tambyah** for interesting discussions extending this research. I would like to thank **Doctor Liam Morrow** for trips and distractions outside the PhD room and maths discussions and **Morgan** for rock climbing trips. I would like to thank **Jesse Sharp** for interesting discussions. **Doctor Eamon Conway** for coding and maths discussions. I would also like to thank those who have been in the PhD room throughout my degree for the interesting discussions. **Doctor Nabil Fadai** for interesting mathematical discussions. **John Blake** for IncuCyte support. **Doctor David Warne** for interesting discussions and saving me on the HPC. **Gordon Shaw** for first suggesting a life in academic research and setting high expectations.

Finally I would like to thank my family and friends. I know it was a surprise when I walked in one weekday evening in the cold of winter and said I was moving to the other side of the world. Thank you for your support. **Mum** for saying do what makes you happy. **Dad** for encouraging my interest in science. **Lauren** for reminding me that maths + sun = dream. **Grandma** for asking why maths is useful and hopefully I have a good answer now. **Steve** and **Freddie** for their encouragement. The McManus and Edmunds family for welcoming me. And lastly for her love, support and general awesomeness my love **Jeorgia**. For all I am hugely grateful.



# List of Publications

1. **Murphy RJ**, Buenzli PR, Baker RE, Simpson MJ (2019). A one-dimensional individual-based mechanical model of cell movement in heterogeneous tissues and its coarse-grained approximation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences.* 475:20180838. doi:10.1098/rspa.2018.0838. bioRxiv preprint. [157]  
Selected for the cover of July 2019 issue of the journal
2. **Murphy RJ**, Buenzli PR, Baker RE, Simpson MJ (2020). Mechanical cell competition in heterogeneous epithelial tissues. *Bulletin of Mathematical Biology.* 82:130. doi:10.1007/s11538-020-00807-x. bioRxiv preprint [158]
3. **Murphy RJ**, Buenzli PR, Baker RE, Simpson MJ (2021). Travelling waves in a free boundary mechanobiological model of an epithelial tissue. *Applied Mathematics Letters.* 111:106636. doi: 10.1016/j.aml.2020.106636. arxiv preprint [159]
4. **Murphy RJ**, Buenzli PR, Tambyah TA, Thompson EW, Hugo H, Baker RE, Simpson MJ (2021). The role of mechanical interactions in EMT. *Physical Biology.* 18:046001. doi: 10.1088/1478-3975/abf425. bioRxiv preprint [160]
5. **Murphy RJ**, Browning AP, Gunasingh G, Haass NK, Simpson MJ (2021). Designing and interpreting 4D tumour spheroid experiments. *Under consideration at Nature Communications.* bioRxiv preprint [156]



# Contents

<b>Abstract</b>	v
<b>Statement of original authorship</b>	vii
<b>Acknowledgements</b>	ix
<b>List of Publications</b>	xi
<b>1 Introduction</b>	1
1.1 Overview . . . . .	2
1.2 Research questions . . . . .	13
1.3 Aims, objectives and outcomes . . . . .	14
1.4 Structure of thesis . . . . .	20
1.5 Statement of joint authorship . . . . .	22
<b>I Mathematical model development</b>	27
<b>2 Mechanical relaxation in heterogeneous populations</b>	29
2.0 Preamble . . . . .	30
2.1 Abstract . . . . .	32
2.2 Introduction . . . . .	33
2.3 Model description . . . . .	34
2.3.1 Discrete model . . . . .	34
2.3.2 Derivation of continuum model . . . . .	37
2.4 Results and discussion . . . . .	41
2.4.1 Homogeneous cell population . . . . .	41
2.4.2 Heterogeneous cell population . . . . .	43
2.4.3 Case study: breast cancer detection . . . . .	49
2.5 Conclusion and future work . . . . .	51

<b>2A Supplementary Material for Chapter 2</b>	<b>53</b>
2A.1 Discrete model simulation . . . . .	55
2A.1.1 Discrete equations . . . . .	55
2A.1.2 Converting cell density into initial positions . . . . .	55
2A.1.3 Assigning spring properties . . . . .	55
2A.1.4 Numerical methods . . . . .	56
2A.2 Continuum model simulation . . . . .	57
2A.2.1 Discretisation scheme . . . . .	57
2A.3 Steady state analysis for two tissue model . . . . .	59
2A.4 Breast cancer detection case study: model implementation . . . . .	60
2A.4.1 Choosing the resting spring length to choose the steady state . . . . .	60
2A.5 Supplementary Figures . . . . .	61
<b>3 Mechanical cell competition</b>	<b>67</b>
3.0 Preamble . . . . .	68
3.1 Abstract . . . . .	70
3.2 Introduction . . . . .	71
3.3 Model formulation . . . . .	73
3.3.1 Discrete model . . . . .	73
3.3.2 Derivation of continuum model . . . . .	76
3.4 Numerical results . . . . .	84
3.4.1 A homogeneous tissue . . . . .	84
3.4.2 Mechanical cell competition . . . . .	88
3.4.3 Importance of the discrete to continuum approach . . . . .	92
3.5 Conclusion . . . . .	95
<b>3A Supplementary Material for Chapter 3</b>	<b>97</b>
3A.1 Model formulation . . . . .	99
3A.1.1 Discrete model with $m > 1$ springs per cell . . . . .	99
3A.1.2 Death at boundaries . . . . .	102
3A.1.3 Derivation of proliferation with $m > 1$ springs per cell . . . . .	103
3A.1.4 Mechanical relaxation . . . . .	111
3A.2 Numerical methods . . . . .	112
3A.2.1 Discrete model . . . . .	112
3A.2.2 Continuum model . . . . .	114
3A.3 Homogeneous population . . . . .	119
3A.3.1 Extinction for constant proliferation and death . . . . .	119
3A.3.2 Reduced variance with cell-length-dependent mechanisms . . . . .	121

3A.3.3 Homogeneous population: logistic proliferation and death . . . . .	123
3A.3.4 Piecewise proliferation: varying mechanical relaxation rate . . . . .	125
3A.4 Mechanical cell competition . . . . .	128
3A.4.1 Two populations: cell size at mechanical equilibrium . . . . .	128
3A.4.2 Mechanical cell competition: robustness to initial conditions . . . . .	129
3A.4.3 Mechanical cell competition: logistic combination . . . . .	133
3A.4.4 Varying mechanical relaxation rate . . . . .	136
<b>4 Travelling waves</b>	<b>137</b>
4.0 Preamble . . . . .	138
4.1 Abstract . . . . .	140
4.2 Introduction . . . . .	141
4.3 Mathematical model . . . . .	142
4.4 Travelling waves . . . . .	143
<b>4A Supplementary Material for Chapter 4</b>	<b>147</b>
S1 Additional Figures . . . . .	148
<b>5 Mechanical interactions in EMT</b>	<b>153</b>
5.0 Preamble . . . . .	154
5.1 Abstract . . . . .	156
5.2 Introduction . . . . .	157
5.3 Model description . . . . .	159
5.3.1 Discrete model . . . . .	159
5.3.2 Continuum model . . . . .	162
5.4 Results and discussion . . . . .	165
5.4.1 Cell-length-independent proliferation and chemically-independent EMT .	165
5.4.2 Cell-length-dependent proliferation and chemically-independent EMT .	169
5.4.3 Chemically-dependent EMT with small diffusivity . . . . .	171
5.4.4 Chemically-dependent EMT with higher diffusivity . . . . .	172
5.5 Conclusion . . . . .	174
<b>5A Supplementary Material for Chapter 5</b>	<b>177</b>
5A.1 Continuum model: Evolution of free boundary equation derivation . . . . .	179
5A.2 Numerical methods: Discrete model . . . . .	180
5A.3 Numerical methods: Continuum model . . . . .	189
5A.3.1 Initial conditions . . . . .	194
5A.4 Additional results . . . . .	195

5A.4.1 Counting the total number of cells that detach . . . . .	195
5A.4.2 Cell-length-independent proliferation . . . . .	197
5A.4.3 Cell-length-dependent proliferation . . . . .	201
5A.4.4 Differences between the cell-based and continuum models: low $N(t)$ . . . . .	203
5A.4.5 Chemically-dependent EMT: cell-length-independent proliferation . . . . .	205
5A.4.6 Sensitivity to $\phi$ . . . . .	209
5A.4.7 Chemically-dependent EMT: linear proliferation . . . . .	211
5A.5 Diffusive equilibrium at all times . . . . .	213
<b>II Experimental design and mathematical modelling</b>	<b>215</b>
<b>6 Designing and interpreting 4D tumour spheroid experiments</b>	<b>217</b>
6.0 Preamble . . . . .	218
6.1 Abstract . . . . .	221
6.2 Introduction . . . . .	222
6.3 Results . . . . .	225
6.3.1 Outer radius measurements are not sufficient to predict inhibited and necrotic radii . . . . .	225
6.3.2 Cell cycle data are informative . . . . .	229
6.3.3 Information gained using spheroids of different sizes is consistent . . . . .	232
6.4 Discussion . . . . .	234
6.5 Methods . . . . .	236
6.5.1 Mathematical model . . . . .	236
6.5.2 Practical parameter identifiability analysis . . . . .	238
6.5.3 Experimental methods . . . . .	240
<b>6A Supplementary Material for Chapter 6</b>	<b>243</b>
6A.1 Mathematical model . . . . .	245
6A.1.1 Nutrient only . . . . .	245
6A.1.2 Nutrient and waste . . . . .	252
6A.2 Profile likelihood further details . . . . .	255
6A.2.1 Numerical method . . . . .	255
6A.2.2 Parameter bounds . . . . .	259
6A.2.3 Pooled sample variances . . . . .	260
6A.3 Experimental data . . . . .	261
6A.3.1 Outer radius experimental measurements and images . . . . .	261
6A.3.2 Confocal microscopy . . . . .	269

6A.3.3 Confocal microscopy supplementary experimental images . . . . .	270
6A.4 WM793b additional results . . . . .	279
6A.4.1 Results in tables . . . . .	279
6A.4.2 Measurement times and experimental duration . . . . .	281
6A.4.3 Profile likelihoods for $R_o(0)$ . . . . .	286
6A.4.4 Outer radius measurements are not sufficient to predict inhibited and necrotic radii . . . . .	287
6A.4.5 Cell cycle data are informative . . . . .	291
6A.5 Synthetic data: WM793b . . . . .	295
6A.5.1 Outer radius measurements are not sufficient to predict inhibited and necrotic radii . . . . .	295
6A.5.2 Cell cycle data are informative . . . . .	298
6A.5.3 Role of initial spheroid size and experiment duration . . . . .	300
6A.5.4 Increasing number of measurements . . . . .	302
6A.6 Parameter identifiability analysis for WM983b . . . . .	304
6A.6.1 Outer radius measurements are not sufficient to predict inhibited and necrotic radii . . . . .	305
6A.6.2 Cell cycle data are informative . . . . .	309
6A.6.3 Information gained across spheroid sizes is consistent . . . . .	312
6A.7 Parameter identifiability analysis for WM164 . . . . .	313
<b>7 Conclusion</b>	<b>317</b>
7.1 Summary of the research . . . . .	318
7.2 Future work . . . . .	323
7.2.1 Epithelial tissue dynamics . . . . .	323
7.2.2 Tumour growth . . . . .	325
7.3 Final remarks . . . . .	326
<b>Bibliography</b>	<b>327</b>



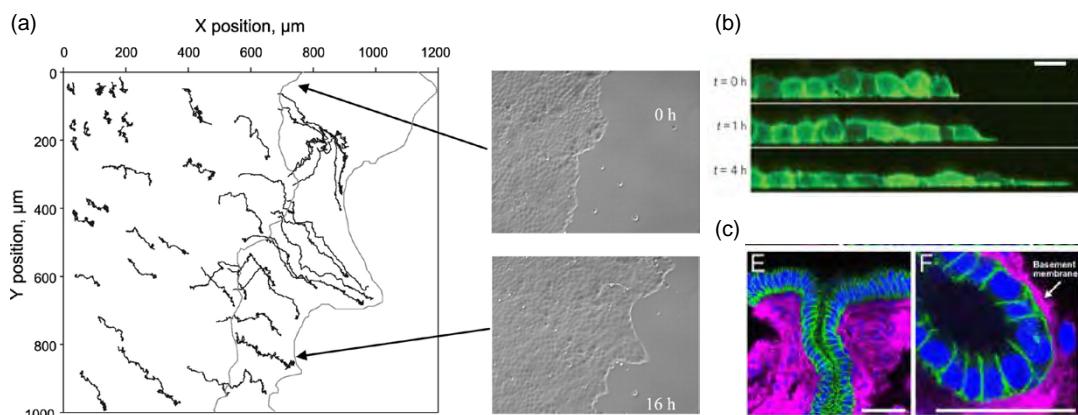
# **Chapter 1**

## **Introduction**

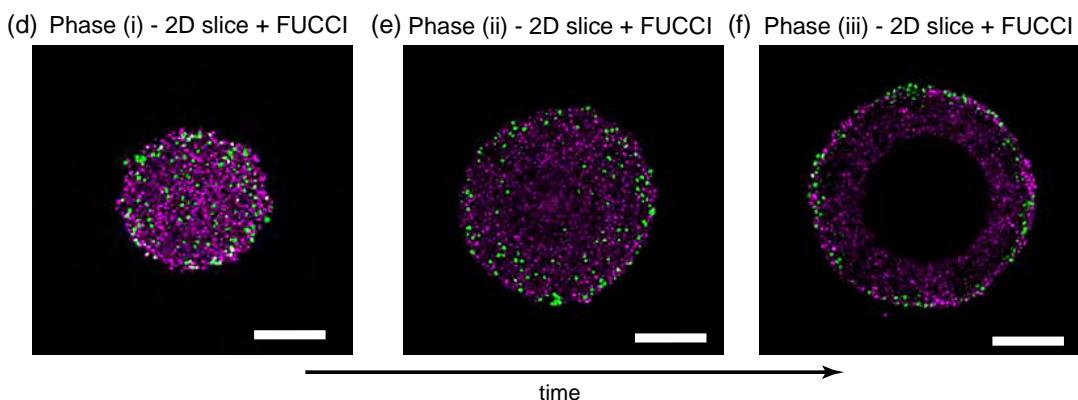
## 1.1 Overview

Mechanochemical processes, whereby mechanical and chemical processes are coupled, are of great interest in experimental biology with applications to epithelial tissue dynamics and cancer tumour growth [28, 79, 91, 98, 171, 228]. Improved understanding of these fundamental biological processes in development and disease will help combat diseases such as cancer [91]. Epithelial tissues are widespread throughout the body and cover all body surfaces, line body cavities and hollow organs [107, 171] (Figure 1.1a-c), and experiments highlight the importance of mechanochemical processes in morphogenesis, homeostasis and maintaining tissue function [28, 98, 171]. Meanwhile tumour spheroid experiments, performed since the 1970s, are routinely used to understand cancer progression and to test and develop cancer treatments [46, 71, 99, 119, 150, 169, 197, 215, 217, 245] (Figure 1.1d-f). However, interpreting experimental data from these studies is challenging without a suitable framework.

### Epithelial tissue dynamics



### Tumour spheroid experiments



**Figure 1.1:** Experimental images from epithelial tissue and tumour spheroid experiments. (a) Cell migration and tracking in an epithelial monolayer experiment wounded by scratching (Figure 4 from [171] reproduced with permission). (b) Epithelial cell migration and invasion in an *in vitro* experiment (Figure 1 from [198] reproduced with permission). (c) Mouse uterine endometrial epithelium (Figure 1 from [107] reproduced with permission). (d)-(f) Tumour spheroid growth experiments (Figure 1 from Chapter 6).

Mathematical modelling provides a powerful framework to interpret experimental data and to pose, form, and test hypotheses. Thereby, reducing experimental effort, both in time and expense. In the following we focus first on building a mechanochemical model for epithelial tissue dynamics. Second, we explore mathematical models for cancer tumour spheroid growth, discuss current tumour spheroid experimental protocols, and how statistical analysis in combination with mathematical modelling may provide great insights. By studying both epithelial tissue dynamics and tumour spheroid experiments, we develop understanding of mechanochemical and experimental models in mathematical biology from two approaches: (i) those where model development is required to catch-up to the latest experiments; and (ii) those where many mathematical models have already been posed to study experiments but few have been experimentally tested. Specifically, case (i) refers primarily to epithelial tissue dynamics and case (ii) to tumour spheroid experiments.

To study epithelial tissues at the cell and tissue scale two mathematical modelling approaches naturally arise: (i) a discrete model, often referred to as an individual based model, reviewed in [175, 179] and including cellular automata models, cellular Potts models, cell-centre models [179], vertex models, subcellular-element models [196], and tensegrity models [104]; (ii) a continuum model, for example derived using the theory of continuum mechanics [78]. Both modelling approaches are widely used and have their own advantages and disadvantages. Discrete models explicitly describe cellular-level interactions but often lack macroscopic intuition. Continuum models on the other hand often provide no cellular-level information [17] but can be more adept at including concepts of macroscopic stiffness [45, 181] and, for large numbers of cells, as in epithelial tissues, tend to be less computationally expensive. Hybrid intermediate models also exist which consider the multi-scale nature of the problem [13, 176, 236]. One approach to get the best of both worlds is to develop a mathematical modelling framework that allows a modeller to switch between both modelling approaches with an in-depth understanding of when this is appropriate and advantageous. Current approaches to implement such a framework use a variety of coarse-graining techniques and assumptions, including the use of slowly varying and periodic assumptions on the heterogeneity in the model [68, 174], correlation functions [147, 153], and interaction forces from potentials [27]. For epithelial tissue dynamics, key features that should be included in a mathematical modelling framework are cell-cell interaction forces, cell proliferation, cell death, fixed and free boundaries, and mechanochemical processes [28, 198].

At the tissue scale a natural starting point is to use a reaction-diffusion continuum model. Reaction-diffusion models are widely used in mathematical biology for a wide range of ap-

plications [30, 56, 161, 162]. In one spatial dimension, these models typically take the form of

$$\frac{\partial q(x, t)}{\partial t} = \frac{\partial}{\partial x} \left( D(q(x, t)) \frac{\partial q(x, t)}{\partial x} \right) + R(q(x, t)), \quad (1.1.1)$$

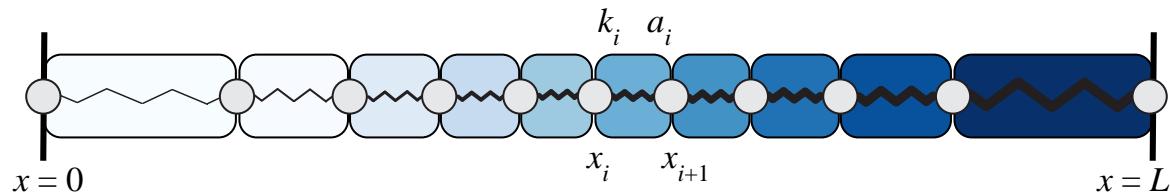
where  $q(x, t)$  is a density, for example of cells;  $x$  represents spatial position;  $t$  represents time; the diffusivity is commonly assumed to be a constant,  $D(q(x, t)) = D$ , which describes a population of cells each of which moves randomly [140, 141]; and, the reaction term can take many forms such as a logistic growth term. Many analytical and computational tools have been developed to study reaction-diffusion type problems. However, it is not immediately clear whether reaction-diffusion models can or should be used to study cells that, rather than moving randomly, move due to mechanical interactions. Two key approaches have been used to demonstrate that cell movement driven by mechanical interactions can be studied with reaction-diffusion equations. The first approach, and the approach taken in this thesis, is very fruitful, namely prescribing a discrete cell-based model, carefully deriving the corresponding continuum model, and comparing the prescribed discrete model and the derived continuum model [164, 165]. The second approach, that which is commonly applied in continuum mechanics, is to consider conservation of mass and momentum to determine the continuum model and proceed only with the continuum model [78]. We show that this second approach may lead to incorrect results.

At the cell scale, key starting points for this thesis are the studies by Murray et al. [164, 165]. By considering an epithelial tissue comprising of a homogeneous population of cells, they first describe a discrete model consisting of a chain of cells (Figure 1.2 where every cell has the same cellular properties). Each cell is thought to act like a mechanical spring with cell stiffness,  $k > 0$ , and resting cell length,  $a > 0$ . The motion of cells is assumed to occur in an overdamped and viscous environment with mobility coefficient,  $\eta > 0$ . Then the time-evolution of the position of a cell boundary,  $x_i(t)$ , is given by (Figure 1.2a),

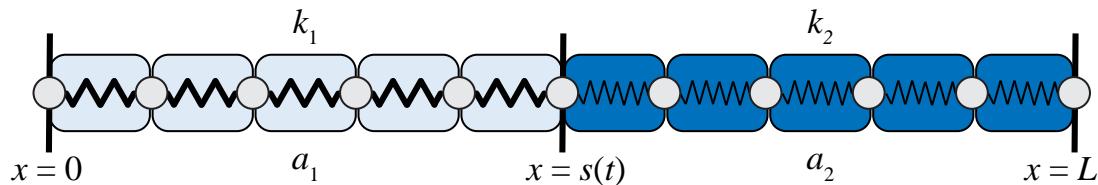
$$\eta \frac{dx_i(t)}{dt} = F_{i+1,i} - F_{i,i-1}, \quad i = 1, 2, \dots, N. \quad (1.1.2)$$

where  $F_{i+1,i}$  represents the cell-cell interaction force experienced on the  $i^{\text{th}}$  cell boundary due to cell  $i + 1$  (Figure 1.2). Equations (1.1.2) form a system of deterministic ordinary differential equations that are solved to obtain the full time-evolution for the positions of all cells within the tissue. The corresponding continuum model takes the form of a reaction-diffusion equation (1.1.1), with  $D(q(x, t))$  determined by the cell-cell interaction force [165], for example

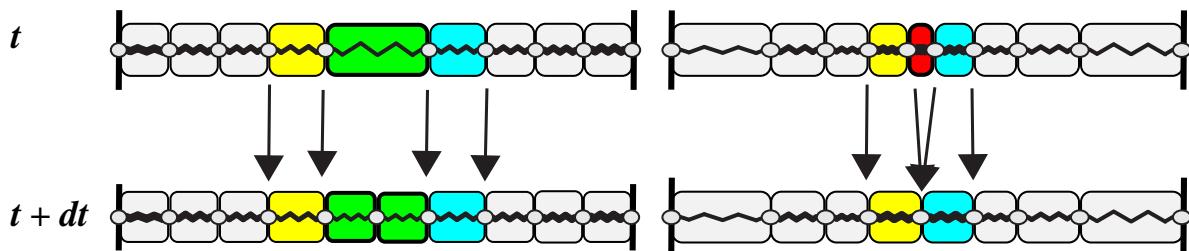
(a) Epithelial tissue dynamics for a heterogeneous cell population



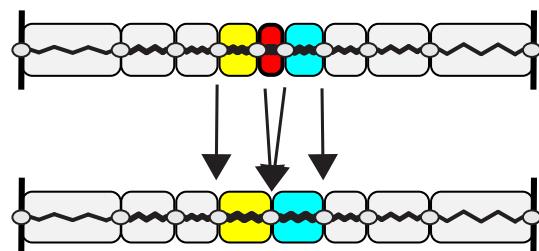
(b) Competing adjacent epithelial tissues



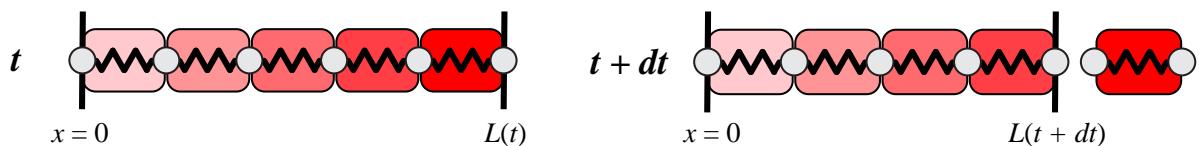
(c) Cell proliferation



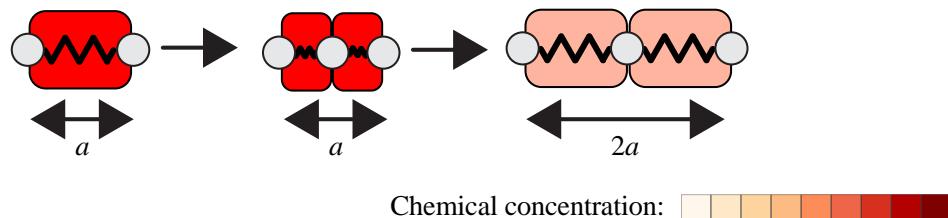
(d) Cell death



(e) Free boundary epithelial tissue dynamics with cell detachment due to EMT



(f) Impact of proliferation on size and chemical concentration



**Figure 1.2:** Mathematical modelling of epithelial tissue dynamics (adapted from schematics from Chapters 2, 3, 4, and 5). Cell-cell interactions are incorporated by considering each cell to act like a mechanical spring. Cell heterogeneity is shown in blue in (a) for a slowly-varying-in-space heterogeneous cell population, and in (b) for rapidly-varying-in-space heterogeneous cell population. (c) A new cell is formed during cell proliferation. (d) A cell is lost during cell death. (e) During epithelial-mesenchymal transitions cells undergo a phenotypic and morphological changes to detach, for example at the free boundary of the tissue where chemical concentration may be highest. (f) Impact of cell proliferation on cell size and chemical concentration.

$D(q(x,t)) = k/(\eta q(x,t)^2)$  for Hooke's linear force law, and with no reaction term when cell proliferation is neglected. Interestingly, unlike many models studied in mathematical biology, the diffusion term is nonlinear and is inversely proportional to the square of the cell density,  $q(x,t)$ . Murray et al.'s work has also been adapted to intestinal crypts [166, 167].

Murray et al.'s [164, 165] model is applicable to homogeneous cell populations. While the discrete model can be readily extended to heterogeneous cell populations (Figure 1.2a,b), it is not clear what the corresponding continuum model is nor when it is valid. Furthermore, it is not clear if the techniques used to derive the continuum model for homogeneous populations can be applied when considering a heterogeneous population of cells. Other discrete to continuum models have also been studied, for example with energy considerations [68], correlation functions [147, 153], and interaction forces from potentials [27]. However, these models do not include mechanical stiffness, which is important in cancer progression [193], cancer invasion and metastasis [170], cancer detection [10, 103, 182, 221], wound healing [61], and morphogenesis [66].

Many mathematical modelling studies assume that epithelial tissues comprise of homogeneous populations [164, 165]. However, in reality epithelial tissues are naturally heterogeneous [182]. This heterogeneity can arise at multiple scales [38, 248]. At the cellular scale fundamental properties that can vary between cells include mechanical properties, such as cell stiffness, equilibrium cell size, proliferation rate and death rate. Furthermore, heterogeneity in mechanical stiffness is an important biomarker in cancer detection [10, 103, 182, 221]. This cell-to-cell heterogeneity can be naturally incorporated into discrete models. However, Murray et al.'s [164, 165] discrete and continuum model has not been extended to heterogeneous populations. Fozard et al. [68] have introduced heterogeneity into their one-dimensional model, derived using energy considerations, but invoke slowly-varying-in-space and periodic assumptions. Heterogeneity has also been incorporated in two-dimensional and three-dimensional discrete models for epithelial monolayers, with prominent classes of models that explicitly incorporate cell-cell forces including cell-centre and vertex models [67, 175, 179]. However the corresponding continuum models are unclear with limited progress even for homogeneous populations [174]. Any insights gained at lower dimensions will be valuable to develop understanding of higher dimensional models. Developing a discrete to continuum modelling framework that incorporates slowly-vary-in-space and rapidly-vary-in-space heterogeneity for low and high cell numbers is of interest (Figure 1.2a,b), with applications to epithelial monolayers, intestinal crypts, and tissue engineering [43]. Furthermore, introducing heterogeneity into an epithelial tissue model would open up a new realm of possibilities.

One such example where including heterogeneous populations is important is in cell competition [130]. Cell competition is the mechanism responsible for the elimination of viable sub-optimal loser cells by optimal winner cells [29]. Cell competition can act as a quality control mechanism in tissue development or as a defence against precancerous cells, and harnessing cell competition has been suggested as a possible approach to enhance both cell-based cancer and regenerative therapies [184]. Many experimental studies have identified different genes and signalling pathways as potential mechanisms underlying cell competition [29]. Furthermore, early studies of cell competition suggested that cell competition and long-term population outcomes arise due to differences in intrinsic cell proliferation and cell death rates [210, 211]. Many mathematical models study the time-evolution of competing populations using ordinary differential equations [56, 161, 162, 172], but these models neglect spatial effects. Partial differential equation models incorporate spatiotemporal dynamics, and typically take the form of a system of reaction-diffusion equations (1.1.1) [56, 161, 162]. In these frameworks, the role of mechanical interactions and mechanical properties is often overlooked.

Recent experiments have demonstrated that mechanical interactions can play an important role in cell competition, leading research of mechanical cell competition [29, 125, 129, 130, 241]. In the emerging research area of mechanical cell competition, *winner* cells compress neighbouring cells promoting tissue crowding and regions of higher density, which leads to cell death (Figure 1.2d) [29, 129, 241], while cell proliferation occurs in regions of lower density (Figure 1.2c) [82]. However, a theoretical framework to study these processes which connects cell and tissue scales is unavailable at present [130]. Such a framework would be beneficial to form and test hypotheses on how cellular mechanisms and competition mechanisms impact the long-term survival of competing populations [130].

One source of experimental data for epithelial tissue dynamics is *in vitro* experiments, perhaps from wound healing scratch and barrier removal assays (Figure 1.1a) [95, 171, 183]. Such wound healing assay experiments have previously been studied with travelling wave solutions of the Fisher-KPP reaction-diffusion model, described by equation (1.1.1) with constant diffusivity and logistic type reaction term [140, 141]. Travelling wave solutions can provide analytical expressions for the shape of wavefronts and the wavespeed, the speed at which the front of the wave invades into free space. However, these solutions may not have a well-defined front, which means that it is not clear where the front of the wave is defined. For example, solutions of the Fisher-KPP equation on  $-\infty < x < \infty$  do not have compact support because the cell density,  $q(x, t)$ , is always positive, with  $q(x, t) \rightarrow 0$  as  $x \rightarrow \infty$  [58, 140, 141, 161].

A well-defined front is observed in experiments. Therefore other studies have developed

models to capture this feature, for example by reformulating the Porous-Fisher equation to include a moving boundary [194, 195, 200, 247] or by adapting the Fisher-KPP equation to a Fisher-Stefan model so that  $q(x, t) = 0$  at  $x = L(t)$ , where  $L(t)$  is the domain length [57, 58, 62, 63]. However, a physical meaning for the Stefan condition is not clear and in general travelling wave solutions for reaction-diffusion models often overlook the important role of mechanics [198]. Since nonlinear reaction-diffusion equations describing homogeneous populations often support travelling wave solutions [161, 201], a natural question to ask is whether the free boundary version of Murray et al.'s work [164, 165], proposed by Baker et al. [19], has a travelling wave solution. We might anticipate that a travelling wave solution does indeed exist and that a well-defined front may naturally arise from clearly motivated biological assumptions. Furthermore, it would be interesting to see how mechanical properties of cells influence the wavespeed [198].

Thus far we have primarily discussed the role of mechanical interactions in epithelial tissue dynamics. However, chemical processes are at least equally as important and often strongly coupled with mechanical processes [28, 91, 98, 171, 228]. Recent progress has been fast. Just under a decade ago experiments showed mechanical waves in freely expanding epithelial tissue [198]. They interpreted these results using a discrete model similar to that of Murray et al. [164, 165]. Last year, similar experiments were performed with improved technology and direct connections were made between the mechanical waves and ERK chemical waves, strongly suggesting a mechanochemical feedback loop [28, 98]. The latest results were modelled using a modified version of Murray et al.'s continuum model. However, these models tend to neglect cell proliferation and cell death, which may play an important role in experimental results and longer term studies [28, 198]. Furthermore, they do not utilise the advantages of both discrete and continuum modelling approaches.

Mechanochemical coupling has also recently been explored mathematically in relation to epithelial tissue dynamics with intracellular signalling incorporated into each cell of Murray et al.'s work [164, 165]. This work by Tambyah et al. [225], motivated by the work in this thesis and the work of Zmurchok et al. [256], nicely demonstrates how perturbations to mechanochemical coupling at the cellular scale give rise to different tissue level behaviour. This is shown for Rac-Rho pathway and for activator-inhibitor systems giving rise to Turing-like patterns in moving boundary problems [48, 49, 225]. Given that mechanochemical processes have been shown to be important for behaviour within the tissue, we ask what the role of mechanochemical processes is at the leading edge of the tissue which is largely unexplored.

Cells in epithelial tissues are characterised as moving collectively and being closely ad-

herent. These epithelial cells can undergo phenotypic and morphological changes to partially or fully transition to mesenchymal cells, typically characterised as cells that are less adherent to other cells and tend to move as individuals, in a process called epithelial-mesenchymal transitions (EMT) [110, 249]. Epithelial-mesenchymal transitions are key events in embryonic development; wound healing; and cancer development [124, 229, 249]. These transitions have been studied extensively for decades experimentally and more recently with mathematical models from the perspective of intracellular chemical processes and regulatory networks [41, 41, 110, 111, 137, 226, 229]. However, even though experimental evidence exists that demonstrates that physical signals, such as mechanical stress [77], can play a role in EMT, the role of mechanics in EMT is largely unexplored. Notable exceptions being discrete models that explore how cell adhesion is influenced by intracellular chemical processes leading to predictions of temporal cell-cell detachments [12, 187]. However, there is no current discrete-to-continuum mathematical modelling framework to explore such processes and the long-term behaviour of epithelial tissues [110]. In Figure 1.2e-f we present schematics for a free boundary model of epithelial tissue dynamics where diffusion of an epithelial-mesenchymal transition inducing chemical results in cell detachment at the tissue boundary.

Many of the key features discussed for epithelial monolayers also occur in three-dimensional cancer tumours, such as mechanical interactions, cell proliferation, cell death, mechanochemical processes, heterogeneity, and epithelial-mesenchymal transitions [46, 71, 79, 99, 119, 150, 169, 197, 215, 217, 245]. However, in comparison to two-dimensional epithelial monolayers, tumours are more appropriately studied experimentally and mathematically in three-dimensions, since behaviour within the tumour is understood to be strongly spatially-dependent [79]. Unlike mechanochemical processes in epithelial tissues, many mathematical models have been developed to study avascular tumour growth [7, 16, 17, 36, 40, 65, 79, 85, 105, 106, 108, 117, 122, 135, 148, 149, 152, 242, 243]. However, even though many mathematical models have been developed few have been experimentally validated [16]. It is therefore of interest to connect such mathematical models to experimental data.

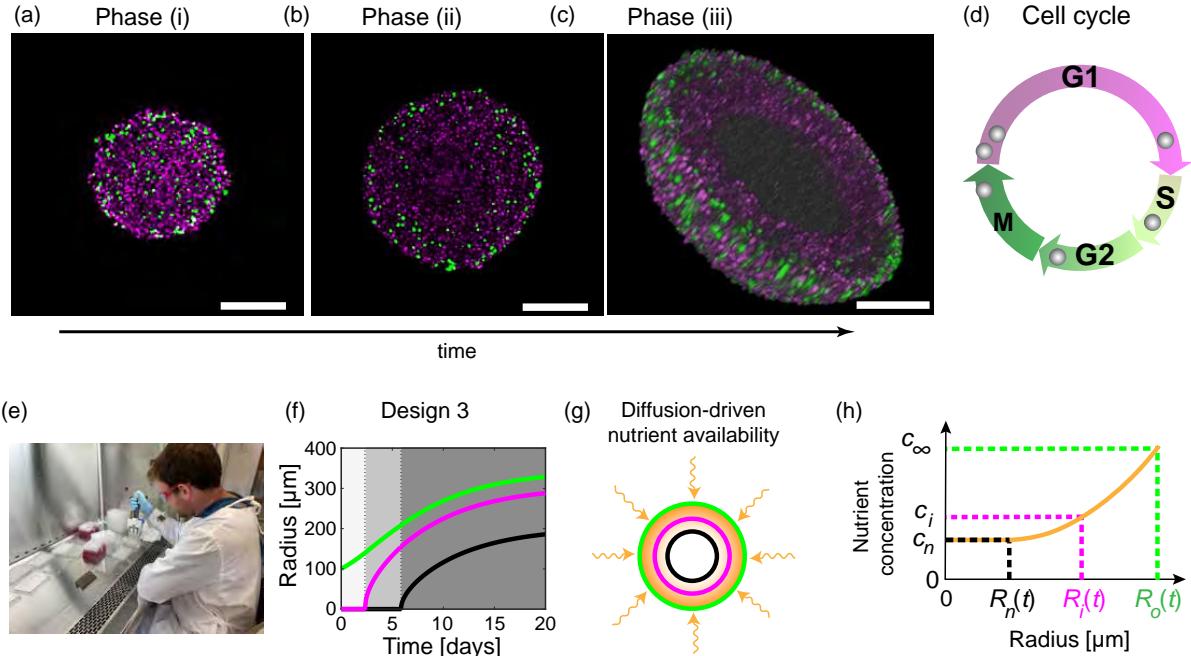
Connecting mathematical models to experimental data is also beneficial to experimental researchers. A recent powerful experimental tool is fluorescent ubiquitination-based cell cycle indicator (FUCCI) technology [87, 192, 251]. FUCCI technology enables real-time imaging of the cell cycle. These tools have been valuable in combination with mathematical modelling [39, 237, 238], for example in replicating normally-hidden inherent synchronisation in cell proliferation [238]. These two-dimensional in vitro experimental studies have been connected to mathematical modelling using different statistical methods, such as approximate Bayesian

computation [39] and statistical identifiability analysis [204]. Statistical identifiability analysis is also been useful to explore heterogeneous media [205], model misspecification [206], and tissue growth in 3D printed scaffolds [32].

Tumour spheroid experiments have been routinely used to study avascular tumour growth and to test and develop drug treatments since the 1970s [46, 71, 79, 99, 119, 150, 169, 197, 215, 217, 245]. However, experimental design is inconsistent and there is no consensus on the optimal experimental design [46, 71, 99, 119, 150, 169, 197, 215, 217, 245]. These biological experiments are inconsistent in: (i) the times when measurements are taken; (ii) the duration of experiments, which from a few days to over a month [8, 18, 83, 121, 168, 177]; (iii) the initial number of cells used to form spheroids [8, 18, 54, 83, 88, 121, 168, 177], for example 300 to 20,000 cells [54, 168]; and (iv) the type of experimental measurements that are taken [8, 18, 54, 83, 88, 121, 168, 177]. Furthermore, it is challenging to interpret and compare experimental data and results across experimental designs. Therefore, it is of interest to investigate and potentially identify which experimental designs lead to reliable biological insight. Mathematical modelling and statistical analysis could provide an insightful approach to quantitatively compare experimental designs and in doing so test mathematical models of tumour growth with experimental data.

Given that there are many mathematical models of tumour growth with few experimentally validated [16], it seems most sensible at this time to use an existing mathematical model rather than develop a new mathematical model. Furthermore, at this stage choosing a mathematical model with few parameters, that have a relatively straightforward biological interpretation, should gain mechanistic insight and avoid overfitting to experimental data. Therefore, we now discuss the seminal Greenspan model [79]. The Greenspan model was published in 1972 and has been highly influential since it was the first mathematical model to describe the three phases of avascular tumour spheroid growth: (i) all cells throughout the spheroid proliferate (Figure 1.3a), (ii) cells close to the periphery proliferate while cells at the centre of the spheroid are arrested and unable to proliferate (Figure 1.3b); (iii) cells close to periphery proliferate while an intermediate region of cells are arrested and a central region is composed of dead cells in various stages of dissolution (Figure 1.3c). These mechanisms are understood to be the result of nutrient availability [79] (Figure 1.3f-h).

Greenspan's mathematical model is a mechanochemical model describing avascular tumour growth, and distinct to the mechanochemical models previously discussed for epithelial tissue dynamics. In Greenspan's model mechanical interactions are assumed to maintain the tumour as a compact solid mass, rather than being explored explicitly as in the models of Mur-



**Figure 1.3:** Avascular tumour spheroid growth and the Greenspan mathematical model (adapted from Figure 1 of Chapter 6). Tumour spheroids experience three phases of growth. (a)-(c) Confocal microscopy reveals different phases of tumour growth, including (a)-(b) 2D equatorial plane images of WM793b human melanoma tumour spheroids on days 7, and 14 after seeding, and (c) 3D view of half of a WM793b human melanoma tumour spheroid on day 21 after seeding. Confocal microscopy of fluorescent ubiquitination-based cell cycle indicator (FUCCI) transduced cells allow visualisation of each cells stage in the cell cycle. (d) Cell cycle schematic coloured with respect to FUCCI signal. (e) To perform the tumour spheroid experiments to collect the experimental data for this chapter I was trained, starting from no experience in a laboratory prior to this PhD. I performed tumour spheroid experiments from start to finish. Image shows me in the laboratory during the spheroid formation stage of an experiment (further details are shown in Section 6.5.3 Experimental methods and the experimental protocol is detailed in [218]). (f) Simulation of Greenspan's mathematical model with Design 3, from Chapter 6, measuring the the outer (green), necrotic (black), and inhibited (magenta) radius. (g) Schematic for Greenspan mathematical model. Nutrient diffuses within the tumour spheroid and is consumed by living cells. (h) Snapshot of nutrient concentration,  $c(r, t)$  for  $0 < r < R_o(t)$ , for a tumour spheroid in phase (iii) and where  $R_o(t)$  is the tumours outer radius. External nutrient concentration is  $c_\infty$ . Inhibited radius,  $R_i(t)$ , and necrotic radius,  $R_n(t)$ , are defined as the radius where the nutrient concentration first reaches thresholds  $c_i$  and  $c_n$ , respectively.

ray et al [164, 165] and those we will consider in Chapters 2-5. Furthermore, as described previously and in Figure 1.3, chemical diffusion is important in Greenspan's model to determine the time-evolution of the tumour internal structure, whereas in Chapter 5 we will explore chemical diffusion in relation to epithelial-mesenchymal transitions. Greenspan's model, due to a spherical symmetry assumption, describes the evolution of the tumour size with equations with one-spatial dimension, namely the tumour radius. Note that models of Murray et al. [164, 165] and those we will consider in Chapters 2-5 also have one spatial dimension. It is of interest to experimentally test Greenspan's seminal publication for the first time and explore if we can use this model to improve tumour spheroid experimental protocols.

To experimentally test Greenspan's seminal publication for the first time and explore if we can use this model to improve tumour spheroid experimental protocols, we require experimental data. I collect this experimental data in the wet-laboratory (Figure 1.3e). I started this PhD with no laboratory experience and was trained to perform experiments in preparation for the final results chapter of this thesis. Then for the final results chapter of this thesis I perform tumour spheroid experiments from start to finish including: cell culturing; spheroid formation; spheroid harvesting; spheroid fixing and mounting in preparation for imaging; confocal microscopy; and, image acquisition, processing, and analysis. By designing and performing these experiments myself I was able to collect the experimental data to interpret with mathematical modelling and statistical analysis. Furthermore, the opportunity to perform wet-laboratory experiments provided a wider appreciation of the questions explored in this thesis and for future work.

## 1.2 Research questions

The work presented in this thesis is a combination of mathematical model development, new experimental data, and statistical analysis. Following the discussion in the overview (Section 1.1), we address the the following five questions:

### **Part 1: Mathematical model development**

1. Can we extend existing mathematical models of cell movement due to mechanical interactions in homogeneous epithelial tissues to heterogeneous cell populations?
2. Can we extend the model for cell movement due to mechanical interactions in heterogeneous epithelial tissues to incorporate cell proliferation and cell death, and what is impact of the mechanical cell competition?
3. Do travelling wave solutions exist for mathematical models of free boundary epithelial tissue dynamics and if so what are their properties?
4. What is the role of mechanical interactions on epithelial-mesenchymal transitions?

### **Part 2: Experimental design and mathematical modelling**

5. Can we connect Greenspan's mathematical model for avascular tumour spheroid growth to experimental data, and in doing so can we improve the experimental design of future experiments and demonstrate how to quantitatively compare data collected across different experimental designs?

### 1.3 Aims, objectives and outcomes

This thesis has two parts corresponding to two key aims. The first key aim is to develop a mathematical modelling framework to study epithelial tissue dynamics incorporating key biological processes, such as cell movement due to mechanical interactions, mechanical relaxation, cell proliferation, cell death, mechanical cell competition, mechanochemical coupling, and cell detachment at tissue boundaries due to epithelial-mesenchymal transitions. This framework will be applicable to heterogeneous populations on free and fixed domains, and where the discrete model will be prescribed based on biological observations and the corresponding continuum-limit model derived. The second key aim is to use mathematical models to quantitatively compare experimental designs for tumour spheroid experiments to reveal those design choices that are important and lead to reliable biological insight. To achieve these aims it is required that we develop new mathematical models, perform tumour spheroid experiments, and use statistical analysis. This thesis addresses the following five objectives that have a direct one-to-one correspondence with the five research questions:

#### **Part 1: Mathematical model development**

1. Develop a discrete mathematical model to describe cell movement due to mechanical interactions in heterogeneous epithelial tissues and derive and compare to the corresponding continuum-limit model,
2. Extend objective 1 to develop a discrete mathematical model that includes cell proliferation and cell death in order to describe mechanical cell competition in heterogeneous epithelial tissues and derive and compare to the corresponding continuum model,
3. Extend objectives 2 and 3 to examine the travelling wave behaviour of the free boundary continuum model, incorporating cell movement due to mechanical interactions, cell proliferation, and cell death,
4. Extend objectives 2, 3, and 4 to develop a discrete mathematical model to describe the role of mechanical interactions in epithelial-mesenchymal transitions and derive and compare to the corresponding continuum model,

#### **Part 2: Experimental design and mathematical modelling**

5. Perform tumour spheroid experiments in the laboratory and use statistical analysis with the Greenspan model to identify experimental design choices that are important and lead to reliable biological insight. Provide recommendations for future studies and demonstrate how to quantitatively compare data collected across different experimental designs.

This thesis is by publication, consisting of four published articles and one manuscript currently under consideration. The PhD candidate contributed significantly to all five manuscripts as recognised by first authorship of all five manuscripts. The work presented in this thesis fulfils the requirements for the award of a thesis by published papers at the Queensland University of Technology. This thesis incorporates the publications listed, with abstracts, on the next five pages.

1. **Murphy RJ**, Buenzli PR, Baker RE, Simpson MJ (2019). A one-dimensional individual-based mechanical model of cell movement in heterogeneous tissues and its coarse-grained approximation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 475:20180838. doi:10.1098/rspa.2018.0838 biorxiv preprint

### **Abstract**

Mechanical heterogeneity in biological tissues, in particular stiffness, can be used to distinguish between healthy and diseased states. However, it is often difficult to explore relationships between cellular-level properties and tissue-level outcomes when biological experiments are performed at a single scale only. To overcome this difficulty we develop a multi-scale mathematical model which provides a clear framework to explore these connections across biological scales. Starting with an individual-based mechanical model of cell movement, we subsequently derive a novel coarse-grained system of partial differential equations governing the evolution of the cell density due to heterogeneous cellular properties. We demonstrate that solutions of the individual-based model converge to numerical solutions of the coarse-grained model, for both slowly-varying-in-space and rapidly-varying-in-space cellular properties. We discuss applications of the model, such as determining relative cellular-level properties and an interpretation of data from a breast cancer detection experiment.

2. **Murphy RJ**, Buenzli PR, Baker RE, Simpson MJ (2020). Mechanical cell competition in heterogeneous epithelial tissues. *Bulletin of Mathematical Biology*. 82:130. doi:10.1007/s11538-020-00807-x biorxiv preprint

### Abstract

Mechanical cell competition is important during tissue development, cancer invasion, and tissue ageing. Heterogeneity plays a key role in practical applications since cancer cells can have different cell stiffness and different proliferation rates than normal cells. To study this phenomenon, we propose a one-dimensional mechanical model of heterogeneous epithelial tissue dynamics that includes cell-length-dependent proliferation and death mechanisms. Proliferation and death are incorporated into the discrete model stochastically and arise as source/sink terms in the corresponding continuum model that we derive. Using the new discrete model and continuum description, we explore several applications including the evolution of homogeneous tissues experiencing proliferation and death, and competition in a heterogeneous setting with a cancerous tissue competing for space with an adjacent normal tissue. This framework allows us to postulate new mechanisms that explain the ability of cancer cells to outcompete healthy cells through mechanical differences rather than an intrinsic proliferative advantage. We advise when the continuum model is beneficial and demonstrate why naively adding source/sink terms to a continuum model without considering the underlying discrete model may lead to incorrect results.

3. **Murphy RJ**, Buenzli PR, Baker RE, Simpson MJ (2021). Travelling waves in a free boundary mechanobiological model of an epithelial tissue. *Applied Mathematics Letters*. 111: 106636. doi: 10.1016/j.aml.2020.106636 arxiv preprint

### Abstract

We consider a free boundary model of epithelial cell migration with logistic growth and nonlinear diffusion induced by mechanical interactions. Using numerical simulations, phase plane and perturbation analysis, we find and analyse travelling wave solutions with negative, zero, and positive wavespeeds. Unlike classical travelling wave solutions of reaction-diffusion equations, the travelling wave solutions that we explore have a well-defined front and are not associated with a heteroclinic orbit in the phase plane. We find leading order expressions for both the wavespeed and the density at the free boundary. Interestingly, whether the travelling wave solution invades or retreats depends only on whether the carrying capacity density corresponds to cells being in compression or extension.

4. Murphy RJ, Buenzli PR, Tambyah TA, Thompson EW, Hugo H, Baker RE, Simpson MJ (2021). The role of mechanical interactions in epithelial mesenchymal transitions. *Physical Biology*. 18:046001. doi:10.1088/1478-3975/abf425 biorxiv preprint

### **Abstract**

The detachment of cells from the boundary of an epithelial tissue and the subsequent invasion of these cells into surrounding tissues is important for cancer development and wound healing, and is strongly associated with the epithelial-mesenchymal transition (EMT). Chemical signals, such as TGF- $\beta$ , produced by surrounding tissue can be uptaken by cells and induce EMT. In this work, we present a novel cell-based discrete mathematical model of mechanical cellular relaxation, cell proliferation, and cell detachment driven by chemically-dependent EMT in an epithelial tissue. A continuum description of the model is then derived in the form of a novel nonlinear free boundary problem. Using the discrete and continuum models we explore how the coupling of chemical transport and mechanical interactions influences EMT, and postulate how this could be used to help control EMT in pathological situations.

5. **Murphy RJ**, Browning AP, Gunasingh G, Haass NK, Simpson MJ (2021). Designing and interpreting 4D tumour spheroid experiments using mathematical models. *Under consideration at Nature Communications*. bioRxiv preprint

## Abstract

Tumour spheroid experiments are routinely used to study cancer progression and treatment. However, experimental designs are inconsistent, leading to challenges in interpretation and reproducibility. Using live-dead cell staining, and real-time cell cycle imaging, we measure necrotic and proliferation-inhibited regions in many tumour spheroids using various experimental designs that intentionally vary the initial spheroid size across multiple cell lines, and involve making various measurements of the internal spheroid structure using various temporal sampling frequencies. These data are difficult to compare and interpret. However, using an objective mathematical modelling framework and statistical identifiability analysis we quantitatively compare experimental designs and identify design choices that produce reliable biological insight. Measurements of internal spheroid structure provide the most insight, whereas varying initial spheroid size and temporal measurement frequency is less important. Our general framework applies to spheroids grown in different conditions and with different cell types

## 1.4 Structure of thesis

This thesis is presented as a thesis by published papers. Each of the main chapters (Chapters 2, 3, 4, 5, and 6) include the publications listed in Section 1.3, respectively. Each chapter includes a preamble, an abstract and an introduction describing the problem of interest, relevant background and relevant literature. All chapters are independent publications so there is partial overlap in these sections. Each chapter then includes mathematical model development and discussion of new mathematical and/or experimental results. Chapters conclude with a summary of the research findings and suggestions for future work. Chapters 2A, 3A, 4A, 5A, and 6A present supplementary material for each of Chapters 2, 3, 4, 5, and 6, respectively. Supplementary material includes additional results, such as additional details of derivations, mathematical and experimental results. Numerical methods are also presented in supplementary material. The structure of this thesis is then as follows.

**Chapter 1** is the introduction. This includes an overview of mechanochemical and experimental models in mathematical biology focusing on the two parts of this thesis: (i) mathematical model development for mechanochemical models, and (ii) experimental design and mathematical modelling of tumour spheroid experiments. Research questions, aims, objectives and outcomes, structure of the thesis, and statements of joint authorship are stated.

**Chapter 2** includes Publication 1, addresses objective 1 and research question 1. The key results include a new discrete model describing cell movement due to mechanical interactions in heterogeneous epithelial tissues, a derivation to obtain the corresponding continuum model, and the continuum model. Using these new models we examine under what conditions the discrete and continuum model show good agreement, explore slowly-varying-in-space and rapidly-varying-in-space epithelial tissues, and discuss applications including an interpretation of data from a breast cancer detection experiment. Supplementary material, such as additional results and numerical methods, associated with publication 1, is presented in Chapter 2A.

**Chapter 3** includes Publication 2, addresses objective 2 and research question 2. This chapter extends the work of Chapter 2, by incorporating cell proliferation and cell death into the model of Chapter 1. The key results include a discrete model describing mechanical cell competition in heterogeneous epithelial tissues, a derivation to obtain the corresponding the continuum model, and the continuum model. Using these models we examine under what conditions the discrete and continuum model show good agreement, explore the evolution of homogeneous tissues and heterogeneous tissues with a cancerous cells competing for space with healthy cells. We advise why naively adding source/sink terms to a continuum model without considering the underlying discrete model may lead to incorrect results. Supplemen-

tary material, such as additional results for homogeneous and heterogeneous populations and numerical methods, associated with publication 2, is presented in Chapter 3A.

**Chapter 4** includes Publication 3, addresses objective 3 and research question 3. This chapter extends the work of Chapter 2 and 3, by considering a free boundary rather than a fixed boundary and analysing the continuum model derived in earlier chapters. The key results include extending the continuum model to a free boundary and examining travelling wave behaviour. Using this we show travelling wave solutions that may invade or retreat depending on whether the carrying capacity density corresponds to cells being in compression or extension. Further, travelling wave solutions have well-defined fronts and are not associated with heteroclinic orbits in the phase plane. Supplementary material, including additional results, associated with publication 3, is presented in Chapter 4A.

**Chapter 5** includes Publication 4, addresses objective 4 and research question 4. This chapter extends the work of Chapter 2, 3, and 4, by incorporating diffusion of a chemical that influences the rate at which cells detach from the tissue boundary, in a process called epithelial-mesenchymal transition (EMT). In this new mechanochemical model, we focus on exploring the role of mechanical interactions in epithelial-mesenchymal transitions (EMT). The key results include a new discrete model describing the role of mechanical interactions in epithelial-mesenchymal transitions, a derivation to obtain the corresponding the continuum model, and the continuum model. Using this novel nonlinear free boundary problem we explore how mechanochemical coupling influences epithelial-mesenchymal-transitions. Supplementary material, such as numerical methods and additional results, associated with publication 4, is presented in Chapter 5A.

**Chapter 6** transitions to Part 2 of this thesis, addresses objective 5 and research question 5. This chapter uses Greenspan's mathematical model [79] which is a mechanochemical model, but a different mathematical model to that derived in previous chapters. The key results include: performing tumour spheroid experiments with real-time cell cycle imaging where I collect an abundance of experimental data across a range of experimental designs; verification of the Greenspan's mathematical model to that experimental data; and development of an objective mathematical modelling framework with statistical identifiability analysis to quantitatively compare experimental designs and identify design choices that produce reliable biological insight to provide recommendations for future studies. Supplementary material, such as additional experimental images, experimental designs and numerical methods, is presented in Chapter 6A.

**Chapter 7** presents a summary of the research and suggestions for future research.

## 1.5 Statement of joint authorship

Here we outline the contributions of the Ph.D. candidate and the co-authors to each article. All co-authors consent to the presentation of this material in this thesis. Each chapter also includes a signed statement of contribution of co-authors for thesis by published paper.

### Chapter 2: Mechanical relaxation in heterogeneous populations

The corresponding article is:

**Murphy RJ**, Buenzli PR, Baker RE, Simpson MJ (2019). A one-dimensional individual-based mechanical model of cell movement in heterogeneous tissues and its coarse-grained approximation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 475:20180838. doi:10.1098/rspa.2018.0838 biorxiv preprint

#### Statement of joint authorship:

- Ryan J. Murphy (Candidate): Conceived and designed the study, derived the continuum model, developed the codes for numerical simulation of the discrete and continuum models, performed numerical simulations, generated results, interpreted results, drafted the manuscript, and revised the manuscript during the peer-review process.
- Pascal R. Buenzli: Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research.
- Ruth E. Baker: Conceived and designed the study and provided comments and gave final approval for publication.
- Matthew J. Simpson: Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research. Acted as corresponding author.

## Chapter 3: Mechanical cell competition

The corresponding article is:

**Murphy RJ**, Buerzli PR, Baker RE, Simpson MJ (2020). Mechanical cell competition in heterogeneous epithelial tissues. *Bulletin of Mathematical Biology*. 82:130. doi:10.1007/s11538-020-00807-x biorxiv preprint

### Statement of joint authorship:

- Ryan J. Murphy (Candidate): Conceived and designed the study, derived the continuum model, developed the codes for numerical simulation of the discrete and continuum models, performed numerical simulations, generated results, interpreted results, drafted the manuscript, and revised the manuscript during the peer-review process. Acted as corresponding author.
- Pascal R. Buerzli: Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research.
- Ruth E. Baker: Conceived and designed the study and provided comments and gave final approval for publication.
- Matthew J. Simpson: Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research.

## Chapter 4: Travelling waves

The corresponding article is:

**Murphy RJ**, Buenzli PR, Baker RE, Simpson MJ (2021). Travelling waves in a free boundary mechanobiological model of an epithelial tissue. *Applied Mathematics Letters*. 111: 106636. doi: 10.1016/j.aml.2020.106636 arxiv preprint

### **Statement of joint authorship:**

- Ryan J. Murphy (Candidate): Conceived and designed the study, derived the continuum model, developed the codes for numerical simulation of the continuum models, performed numerical simulations, generated results, interpreted results, drafted the manuscript, and revised the manuscript during the peer-review process. Acted as corresponding author.
- Pascal R. Buenzli: Supervised the research, and provided comments and gave final approval for publication.
- Ruth E. Baker: Provided comments and gave final approval for publication.
- Matthew J. Simpson: Supervised the research, and provided comments and gave final approval for publication.

## Chapter 5: Mechanical interactions in epithelial-mesenchymal-transitions

The corresponding article is:

**Murphy RJ, Buenzli PR, Tambyah TA, Thompson EW, Hugo H, Baker RE, Simpson MJ (2021).**

The role of mechanical interactions in epithelial mesenchymal transitions. *Physical Biology*.

18:046001. doi:10.1088/1478-3975/abf425 biorxiv preprint

### Statement of joint authorship:

- Ryan J. Murphy (Candidate): Conceived and designed the study, derived the continuum model, developed the codes for numerical simulation of the discrete and continuum models, performed numerical simulations, generated results, interpreted results, drafted the manuscript, and revised the manuscript during the peer-review process. Acted as corresponding author.
- Pascal R. Buenzli: Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research.
- Tamara A. Tambyah: Provided comments and gave final approval for publication.
- Erik W. Thompson: Provided advice on cancer and EMT. Provided comments and gave final approval for publication.
- Honor Hugo: Provided advice on cancer and EMT. Provided comments and gave final approval for publication.
- Ruth E. Baker: Provided comments and gave final approval for publication.
- Matthew J. Simpson: Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research.

## Chapter 6: Tumour spheroid modelling and parameterisation

The corresponding article is:

**Murphy RJ**, Browning AP, Gunasingh G, Haass NK, Simpson MJ (2021). Designing and interpreting 4D tumour spheroid experiments. *Under consideration at Nature Communications.* biorxiv preprint

### Statement of joint authorship:

- Ryan J. Murphy (Candidate): Conceived and designed the study, performed biological experiments, developed the codes for numerical simulation continuum model and statistical analysis, performed numerical simulations, generated results, interpreted results, drafted the manuscript.
- Alexander P. Browning: Performed biological experiments. Provided comments and gave final approval for publication.
- Gency Gunasingh: Performed biological experiments. Provided comments and gave final approval for publication.
- Nikolas K. Haass: Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research.
- Matthew J. Simpson: Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research.

## **Part I**

# **Mathematical model development**



## **Chapter 2**

# **A one-dimensional individual-based mechanical model of cell movement in heterogeneous tissues and its coarse-grained approximation**

## 2.0 Preamble

An article published in *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*

**Murphy RJ**, Buenzli PR, Baker RE, Simpson MJ (2019). A one-dimensional individual-based mechanical model of cell movement in heterogeneous tissues and its coarse-grained approximation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 475:20180838. doi:10.1098/rspa.2018.0838. bioRxiv preprint.

Selected for the cover of July 2019 issue of the journal.

This chapter includes Publication 1, addresses objective 1 and research question 1. The key results include a new discrete model describing cell movement due to mechanical interactions in heterogeneous epithelial tissues, a derivation to obtain the corresponding continuum model, and the continuum model. Using these new models we examine under what conditions the discrete and continuum model show good agreement, explore slowly-varying-in-space and rapidly-varying-in-space epithelial tissues, and discuss applications including an interpretation of data from a breast cancer detection experiment. Supplementary material, such as additional results and numerical methods, associated with publication 1, is presented in Chapter 2A.

### Statement of Contribution of Co-Authors for Thesis by Published Paper

The authors listed below have certified that:

1. they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
2. they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
3. there are no other authors of the publication according to these criteria;
4. potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit, and
5. they agree to the use of the publication in the student's thesis and its publication on the QUT's ePrints site consistent with any limitations set by publisher requirements.

In this case of this chapter:

Murphy RJ, Buerzli PR, Baker RE, Simpson MJ (2019). A one-dimensional individual-based mechanical model of cell movement in heterogeneous tissues and its coarse-grained approximation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 475:20180838. doi:10.1098/rspa.2018.0838.

Contributor	Statement of Contribution
Ryan J. Murphy (Candidate) <i>R. Murphy</i> 02/09/2021	Conceived and designed the study, derived the continuum model, developed the codes for numerical simulation of the discrete and continuum models, performed numerical simulations, generated results, interpreted results, drafted the manuscript, and revised the manuscript during the peer-review process.
Pascal R. Buerzli	Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research.
Ruth E. Baker	Conceived and designed the study and provided comments and gave final approval for publication.
Matthew J. Simpson	Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research. Acted as corresponding author.

#### Principal Supervisor Confirmation

I have sighted email or other correspondence from all Co-authors confirming their certifying authorship.

Matthew Simpson

2/9/2021

Name

Signature

Date

## 2.1 Abstract

Mechanical heterogeneity in biological tissues, in particular stiffness, can be used to distinguish between healthy and diseased states. However, it is often difficult to explore relationships between cellular-level properties and tissue-level outcomes when biological experiments are performed at a single scale only. To overcome this difficulty we develop a multi-scale mathematical model which provides a clear framework to explore these connections across biological scales. Starting with an individual-based mechanical model of cell movement, we subsequently derive a novel coarse-grained system of partial differential equations governing the evolution of the cell density due to heterogeneous cellular properties. We demonstrate that solutions of the individual-based model converge to numerical solutions of the coarse-grained model, for both slowly-varying-in-space and rapidly-varying-in-space cellular properties. We discuss applications of the model, such as determining relative cellular-level properties and an interpretation of data from a breast cancer detection experiment.

## 2.2 Introduction

Biological tissues are heterogeneous and multi-scale by their very nature (Figure 6.1(a)). This heterogeneity exists at all scales from sub-cellular to cellular, and from cellular to tissue levels [11, 72, 227]. We focus on cellular interactions driven by mechanical stiffness which is of great importance in a variety of applications including epithelial tissue mechanics, cancer progression [193], cancer invasion and metastasis [170], stiffness as a biomarker in cancer detection [10, 103, 182, 221], wound healing [61], and morphogenesis [66]. Tissue-level stiffness information [131] has been available for much longer than cellular-level stiffness data which requires advanced technology, such as atomic force microscopy [52, 126–128]. However, difficulties arise in relating cellular-level data with tissue-level information when experiments are conducted and analysed at a single scale only. Mathematical modelling with *in silico* simulations provides a clear framework to explore these connections across biological scales.

Mathematical models of cell populations are broadly classified as either discrete or continuum. Discrete models, reviewed in [175, 179], include cellular automata models, cellular Potts models, cell-centre models [179], vertex models, subcellular-element models [196], and tensegrity models [104]. Discrete models explicitly describe cellular-level interactions but often lack macroscopic intuition. Continuum models on the other hand often provide no cellular-level information [17] but can be more adept at including concepts of macroscopic stiffness [45, 181] and, for large numbers of cells, as in epithelial tissues, tend to be less computationally expensive. Hybrid intermediate models also exist which consider the multi-scale nature of the problem [13, 176, 236]. A range of models specifically examine the role of mechanics [186, 224]. However, in this work we focus on models which relate cellular-level details to tissue-level outcomes. These models have been developed with a variety of coarse-graining techniques and assumptions, including the use of slowly varying and periodic assumptions on the heterogeneity in the model [68, 174], correlation functions [147, 153], and interaction forces from potentials [27]. Few of these models explore the role of stiffness. The work of Murray et al. [164–167] explicitly incorporates cell stiffness; they derive a nonlinear diffusion equation governing the evolution of the cell density in space and time, however the framework focuses exclusively on homogeneous cell populations. Here, we extend this framework to heterogeneous cell populations.

The key focus of this work is to present a novel coarse-grained system of partial differential equations governing the evolution of the cell density, cell stiffness and resting cell length, from a heterogeneous cell-based model of epithelial tissue mechanics. The cell stiffness and resting cell length are constant for each cell and are simply transported in space by cell movements.

The motion in this model is governed by cell-cell interaction forces modelled with Hooke's law. In extending the work of Murray et al. [164], we provide a more general derivation of the governing equations, see Section 2.3, which is robust to the inclusion of both slowly-varying-in-space and rapidly-varying-in-space cellular properties, see Section 2.4. We show that solutions from the discrete model converge to the corresponding continuum model solution, under appropriate scalings. Additional results in Section 2.4 show the model can be applied to interpret experimental and clinical observations relating to breast cancer detection. Key algorithms used to generate results are available on GitHub.

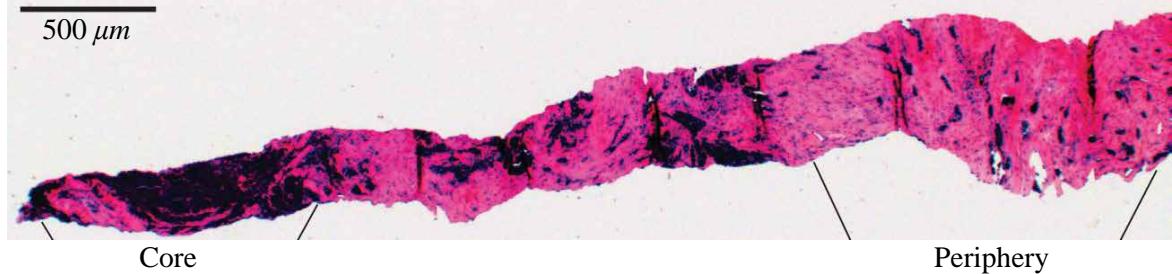
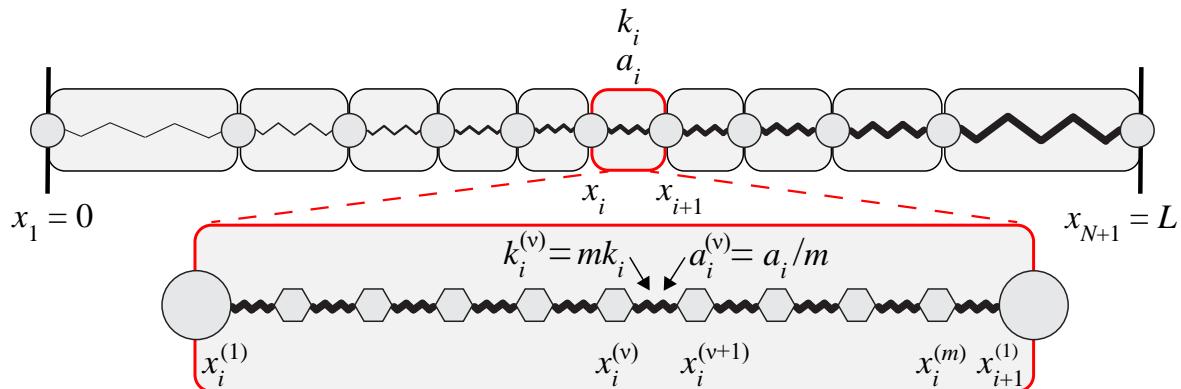
## 2.3 Model description

In this section we describe the individual-based model, which we refer to throughout this work as the discrete model, and derive a corresponding coarse-grained approximation in the form of system of partial differential equations, which we refer to as the continuum description. The continuum limit usually assumes that the number of discrete entities that makes up the system tends to infinity [60, 68], while the size of the domain also tends to infinity, as in the thermodynamic limit, or the size of a length scale tends to zero, both in such a way that the ratio of the size of length scale to the number of discrete entities is fixed. Here, to maintain a fixed total tissue length and a fixed total number of cells in the continuum limit, we instead assume that each cell is internally represented by several identical springs. We then take the continuum limit by considering that the number of springs per cell tends to infinity whilst the spring length tends to zero.

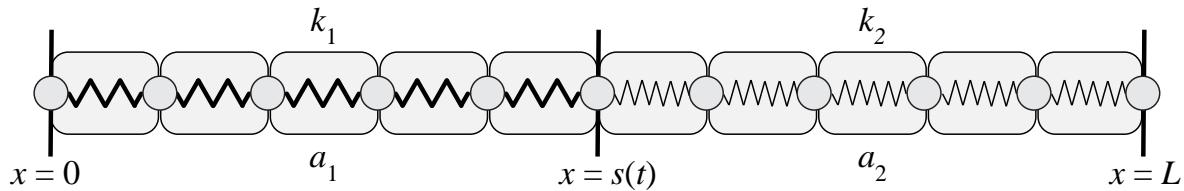
### 2.3.1 Discrete model

In this work, the discrete model describes an epithelial tissue formed by cells in contact with each other. For simplicity, we assume that the tissue can be modelled as a one-dimensional chain of  $N$  cells with fixed total length  $L$ . Tissues in the body commonly evolve in confined spaces, for example imposed by bone tissues, and are subjected to strong geometric constraints so we fix the left tissue boundary at  $x = 0$  and the right tissue boundary at  $x = L$ . This also allows us to focus on internal cellular heterogeneity. Alternate free boundary conditions are possible [19, 136, 164] but we do not focus on such free boundary conditions in this work. Each cell can have distinct mechanical properties (Figure 6.1). This model could be used to represent a single tissue with intrinsic heterogeneity or multiple adjacent tissues with different properties. Each cell interacts with its neighbour through an effective interaction force which

(a) Stiffness in a human breast cancer biopsy

(b) Model schematic for heterogeneous cells, with  $m$  springs per cell

(c) Special case of model schematic with two adjacent tissues



**Figure 2.1:** Stiffness heterogeneity in biological tissues. (a) Post atomic force microscopy histological overview of an entire breast cancer biopsy, where dark regions and pink regions are associated with low and high cell stiffness, respectively. Reproduced from [182] with permission. (b) Individual-based model schematic for arbitrarily heterogeneous tissue with  $N$  cells and  $m$  identical springs per cell. Cell  $i$  occupies the region  $x_i(t) < x < x_{i+1}(t)$  has cell stiffness  $k_i$  and resting cell length  $a_i$ . Spring  $\nu$  in cell  $i$ , occupies the region  $x_i^{(\nu)}(t) < x < x_i^{(\nu+1)}(t)$ , is prescribed with spring stiffness  $k_i^{(\nu)} = m k_i$  and resting spring length  $a_i^{(\nu)} = a_i/m$ . The first and final spring boundaries in cell  $i$  coincide with the cell boundary positions so that  $x_i^{(1)}(t) = x_i(t)$  and  $x_{i+1}^{(1)}(t) = x_{i+1}(t)$  for all time. The cell and spring boundaries are shown as discs and hexagons, respectively. (c) Individual-based model schematic for a special case with two adjacent tissues, similarly this could model a heterogeneous tissue with two cell types. Cells in tissue  $i$  are prescribed with cell stiffness  $k_i$  and resting cell length  $a_i$  for  $i = 1, 2$ . Here each cell is represented with a single spring. The position of the interface between the two tissues is at  $x = s(t)$ .

could represent cell-cell adhesion [109] or compressive stresses [231]. We consider cell  $i$ , for  $i = 1, 2, \dots, N$ , to have its left boundary at  $x_i(t)$  and its right boundary at  $x_{i+1}(t)$ , with  $x_1(t) = 0$  and  $x_{N+1}(t) = L$  at all times. The cell has a prescribed cell stiffness,  $k_i$ , and resting cell length,  $a_i$ . Inside the tissue, Newton's second law of motion governs the motion of each cell boundary such that

$$M_i \frac{d^2 x_i}{dt^2} = f_{i+1} - f_i + F_i^{\text{visc}}, \quad i = 2, 3, \dots, N, \quad (2.3.1)$$

where  $M_i$  is the mass associated with cell boundary  $i$ ,  $F_i^{\text{visc}}$  is the viscous force associated with cell boundary  $i$ , and we model interaction forces at cell boundary  $i$  using Hooke's law,

$$f_i = k_{i-1}(x_i - x_{i-1} - a_{i-1}), \quad i = 1, 2, \dots, N. \quad (2.3.2)$$

The viscous force experienced by cells, due to cell-medium and cell-matrix interactions, is modelled with  $F_i^{\text{visc}} = -\eta dx_i(t)/dt$ , where  $\eta > 0$  is the viscosity coefficient. Cells migrate in dissipative environments and this is commonly modelled by assuming that the motion is overdamped [151, 164], hence the term on the left of Equation (2.3.1) is zero, giving,

$$\eta \frac{dx_i}{dt} = f_{i+1} - f_i, \quad i = 2, 3, \dots, N. \quad (2.3.3)$$

This model, as presented thus far, considers each cell to be represented by a single spring [19, 164] which is sufficient to describe the discrete model. However, when we derive the continuum model, to maintain  $L$  and  $N$ , we represent each cell internally with  $m$  identical springs and we will later consider  $m \gg 1$ , which corresponds to the spring length tending to zero, see Section 2.32.3.2. The corresponding discrete model for  $m$  springs per cell is now described. Cell  $i$  with boundaries  $x_i$  and  $x_{i+1}$  now has  $m + 1$  spring boundaries,  $x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(m)}, x_{i+1}^{(1)}$ , with  $x_i = x_i^{(1)}$  and  $x_{i+1} = x_{i+1}^{(1)}$ , (Figure 6.1(c)). The cell length is related to the spring length through the scaling  $x_{i+1} - x_i \sim m(x_i^{(\nu+1)} - x_i^{(\nu)})$  as  $m \rightarrow \infty$ , and with equality for all  $m$  as  $t \rightarrow \infty$ . Each spring  $\nu$  in cell  $i$  is prescribed with a spring stiffness  $k_i^{(\nu)}$  and resting spring length  $a_i^{(\nu)}$  related to cell properties  $k_i$  and  $a_i$  through

$$k_i^{(\nu)} = m k_i, \quad a_i^{(\nu)} = \frac{a_i}{m}, \quad i = 1, 2, \dots, N, \quad \nu = 1, 2, \dots, m. \quad (2.3.4)$$

The viscosity coefficient for a cell boundary,  $\eta$ , is related to the viscosity coefficient for a spring boundary,  $\eta^{(\nu)}$ , through  $\eta^{(\nu)} = \eta/m$ . Then the corresponding spring boundary equations are

$$\begin{aligned}\eta^{(\nu)} \frac{dx_i^{(\nu)}}{dt} &= f_i^{(\nu+1)} - f_i^{(\nu)}, \\ f_i^{(\nu)} &= k_i^{(\nu-1)} \left[ x_i^{(\nu)} - x_i^{(\nu-1)} - a_i^{(\nu-1)} \right].\end{aligned}\tag{2.3.5}$$

The scalings in Equation (2.3.4) and for the viscosity coefficient are chosen such that the cell boundary velocities are maintained and are independent of  $m$ , i.e. such that  $dx_i^{(1)}/dt = dx_i/dt$ . These scalings are supported by results from the discrete model with varying  $m$ , see Section 2.4.

The discrete model is governed by Equation (2.3.3) with the fixed boundary conditions for a system with a single spring per cell, and by Equation (2.3.5) with fixed boundary conditions for a system with  $m$  springs per cell. In each situation the discrete model forms a deterministic coupled system of ordinary differential equations that we can solve numerically, see Supplementary Material Section 1. We can also solve each system with an eigenmode decomposition to conveniently determine the long-time steady state solution.

### 2.3.2 Derivation of continuum model

We now derive a coarse-grained system of partial differential equations describing the evolution of cell density at a larger scale. To do so we take the continuum limit by increasing the number of springs per cell,  $m$ , while maintaining the total number of cells,  $N$ , and tissue length,  $L$ , fixed, and by performing spatial averages over length scales involving a sufficiently large number of cells to define continuous densities, but sufficiently small to retain spatial heterogeneities. We first define the microscopic cell density,  $\hat{q}(x, t)$ , in terms of the spring boundary positions,  $x_i^{(\nu)}(t)$ , as

$$\hat{q}(x, t) = \frac{1}{m} \sum_{i=1}^N \sum_{\nu=1}^m \delta \left( x - x_i^{(\nu)}(t) \right),\tag{2.3.6}$$

where  $\delta$  is the Dirac delta function [60, 133]. Integrating Equation (2.3.6) over the tissue domain,  $0 < x < L$ , gives the total number of cells,  $N$ , which is independent of  $m$ . We introduce a mesoscopic length scale  $\delta x$  such that  $a_i^{(\nu)} \ll a_i \ll \delta x \ll L$  and define a local spatial average which, for the microscopic cell density,  $q(x, t) = \langle \hat{q}(x, t) \rangle$ , is

$$\langle \hat{q}(x, t) \rangle = \frac{1}{2\delta x} \int_{x-\delta x}^{x+\delta x} \hat{q}(y, t) dy.\tag{2.3.7}$$

Differentiating Equation (2.3.7) with respect to time leads to the general conservation law

$$\frac{\partial q(x, t)}{\partial t} = -\frac{\partial}{\partial x} \left\langle \frac{1}{m} \sum_{i=1}^N \sum_{\nu=1}^m \delta \left( x - x_i^{(\nu)}(t) \right) \frac{dx_i^{(\nu)}}{dt} \right\rangle, \quad (2.3.8)$$

where we use properties of the Dirac delta distribution [133] and take the spatial derivative outside of the average by making use of the fact that  $\delta x$  is small. The averaged term on the right of Equation (2.3.8) is the coarse-grained cell density flux,  $j(x, t)$ , describing spring migration at the mesoscopic scale, expressed explicitly in terms of the spring boundary positions and velocities [60]. We now introduce three field functions,  $f(x, t)$ ,  $k(x, t)$ ,  $a(x, t)$ , for the cell-cell interaction force, the cell stiffness and the resting cell length, respectively, defined such that

$$f \left( x_i^{(\nu)}(t), t \right) = f_i^{(\nu)}, \quad k \left( x_i^{(\nu)}(t), t \right) = m k_i^{(\nu)}, \quad a \left( x_i^{(\nu)}(t), t \right) = \frac{a_i^{(\nu)}}{m}, \quad (2.3.9)$$

where the scalings for  $f$ ,  $k$ , and  $a$ , with respect to  $m$ , agree with the scalings from the discrete system, see Equation (2.3.4). The field functions  $k(x, t)$  and  $a(x, t)$  capture the assumption that spring properties and respective cell properties are constant along spring boundary trajectories,  $x_i^{(\nu)}(t)$ . To represent the distribution of spring lengths across the domain, we introduce a continuously differentiable function,  $l(x, t)$ , which we define such that

$$l \left( x_i^{(\nu)}(t), t \right) = l_i^{(\nu)}(t) = x_i^{(\nu+1)}(t) - x_i^{(\nu)}(t), \quad (2.3.10)$$

where  $l_i^{(\nu)} \ll a_i \ll \delta x \ll L$ . Writing Equation (2.3.5) in terms of these continuous variables, expanding each cell-cell interaction force using the small parameter  $l_i^{(\nu)}$ , using the viscosity coefficient scaling, and simplifying to leading order gives,

$$\begin{aligned} \eta \frac{dx_i^{(\nu)}}{dt} &= m \left[ f_i^{(\nu+1)} - f_i^{(\nu)} \right] = m \left[ f \left( x_i^{(\nu+1)}(t), t \right) - f \left( x_i^{(\nu)}(t), t \right) \right] \\ &= m \frac{\partial f \left( x_i^{(\nu)}(t), t \right)}{\partial x} l_i^{(\nu)} + \mathcal{O} \left( \left[ l_i^{(\nu)} \right]^2 \right). \end{aligned} \quad (2.3.11)$$

Substituting Equation (2.3.11) into Equation (2.3.8), relating the spring length to the cell density with  $l \left( x_i^{(\nu)}(t), t \right) = 1 / \left[ mq \left( x_i^{(\nu)}(t), t \right) \right]$ , and integrating over the spatial average interval,  $(x - \delta x, x + \delta x)$  gives

$$j(x, t) = \frac{1}{\eta} \left( \frac{n}{2\delta x} \right) \frac{1}{n} \sum_{i=1}^n \frac{1}{m} \sum_{\nu=1}^m \frac{\partial f \left( x_i^{(\nu)}(t), t \right)}{\partial x} \frac{1}{q \left( x_i^{(\nu)}(t), t \right)}, \quad (2.3.12)$$

where  $n$  is the number of cells in the interval  $(x - \delta x, x + \delta x)$  and  $i$  has been reset to count these cells. Then, taking the limit as  $m \rightarrow \infty$  and performing an average over the  $m$  springs per cell, gives

$$j(x, t) = \frac{1}{\eta} \left( \frac{n}{2\delta x} \right) \frac{1}{n} \sum_{i=1}^n \frac{\partial f(x_i(t), t)}{\partial x} \frac{1}{q(x_i(t), t)}. \quad (2.3.13)$$

We apply a mean field approximation, as  $n \gg 1$  in  $(x - \delta x, x + \delta x)$  due to  $a_i \ll \delta x$ , by substituting  $q(x_i(t), t)$  and  $\partial f(x_i(t), t)/\partial x$  in the sum with the average density  $q(x, t)$  and the average interaction force gradient  $\partial f/\partial x$  in the interval  $(x - \delta x, x + \delta x)$ . The factor  $1/q$  is now independent of  $i$  and cancels with the factor  $n/(2\delta x)$  which represents the density of cells in the spatial average interval. Then the coarse-grained cell density flux is

$$j(x, t) = \frac{1}{\eta} \frac{\partial f(x, t)}{\partial x}, \quad (2.3.14)$$

which provides us with an important physical interpretation and is directly related to the velocity, net force and cell-cell interaction force gradient. By inspection of Equation (2.3.11) and Equation (2.3.14), we see that the cell density flux,  $j$ , is an advective flux  $j = qu$ , where  $u(x, t) = \langle dx_i/dt \rangle$  is the average velocity induced by the average force gradient  $\langle \partial f/\partial x \rangle$ . We also see that the net force is given by  $\eta j/q$  and the spatially averaged interaction force gradient is given by  $\eta j$ .

Substituting Equation (2.3.14) into Equation (2.3.8) gives

$$\frac{\partial q(x, t)}{\partial t} = -\frac{1}{\eta} \frac{\partial^2 f(x, t)}{\partial x^2}. \quad (2.3.15)$$

We now return to Equation (2.3.9) and differentiate with respect to time to derive an evolution equation for the cell stiffness function

$$0 = \frac{d}{dt} \left[ k \left( x_i^{(\nu)}(t), t \right) - m k_i^{(\nu)} \right] = \frac{\partial k \left( x_i^{(\nu)}(t), t \right)}{\partial t} + \frac{dx_i^{(\nu)}(t)}{dt} \frac{\partial k \left( x_i^{(\nu)}(t), t \right)}{\partial x}. \quad (2.3.16)$$

Using Equation (2.3.11) and similar developments, the evolution equations for the cell stiffness and resting cell length expressed in terms of mesoscopic variables become

$$\frac{\partial k(x, t)}{\partial t} + u(x, t) \frac{\partial k(x, t)}{\partial x} = 0, \quad (2.3.17)$$

$$\frac{\partial a(x, t)}{\partial t} + u(x, t) \frac{\partial a(x, t)}{\partial x} = 0. \quad (2.3.18)$$

Written in terms of velocity we identify the left-hand sides of Equations (2A.2.2) and (2A.2.3) as the convective derivatives of the cell properties.

In summary, the governing equations of the coarse-grained model are given by Equations (2A.2.1), (2A.2.2) and (2A.2.3) with the interaction force  $f$  given by

$$f(x, t) = k(x, t) \left( \frac{1}{q(x, t)} - a(x, t) \right). \quad (2.3.19)$$

This results in a system of four self-consistent equations for the continuous fields  $q(x, t)$ ,  $k(x, t)$ ,  $a(x, t)$ ,  $f(x, t)$  in terms of spatial position rather than particle trajectories. The initial conditions for the average cell density, cell stiffness and resting cell length are

$$q(x, 0) = q_0(x), \quad k(x, 0) = k_0(x), \quad a(x, 0) = a_0(x), \quad 0 < x < L, \quad (2.3.20)$$

together with no flux boundary conditions for the average cell density, due to the microscopic motion constraints, and Dirichlet boundary conditions for the cell stiffness and resting cell length, as cell properties are constant along cell boundary trajectories,

$$\frac{\partial f(x, t)}{\partial x} = 0, \quad k(x, t) = k_0(x), \quad a(x, t) = a_0(x), \quad x = 0, L. \quad (2.3.21)$$

These governing partial differential equations (2A.2.1), (2A.2.2), (2A.2.3), (2A.2.4) are solved numerically with the initial conditions (2.3.20) and boundary conditions (2.3.21), see Supplementary Material Section 2. With homogeneous cell populations the governing equations reduce to the single nonlinear density diffusion equation previously derived in [164],

$$\frac{\partial q}{\partial t} = \frac{\partial}{\partial x} \left( \frac{k}{\eta q^2} \frac{\partial q}{\partial x} \right). \quad (2.3.22)$$

## 2.4 Results and discussion

In this section we compare solutions of the continuum and discrete models with the expectation that as the number of springs per cell,  $m$ , increases solutions from the discrete model converge to the corresponding continuum solution.

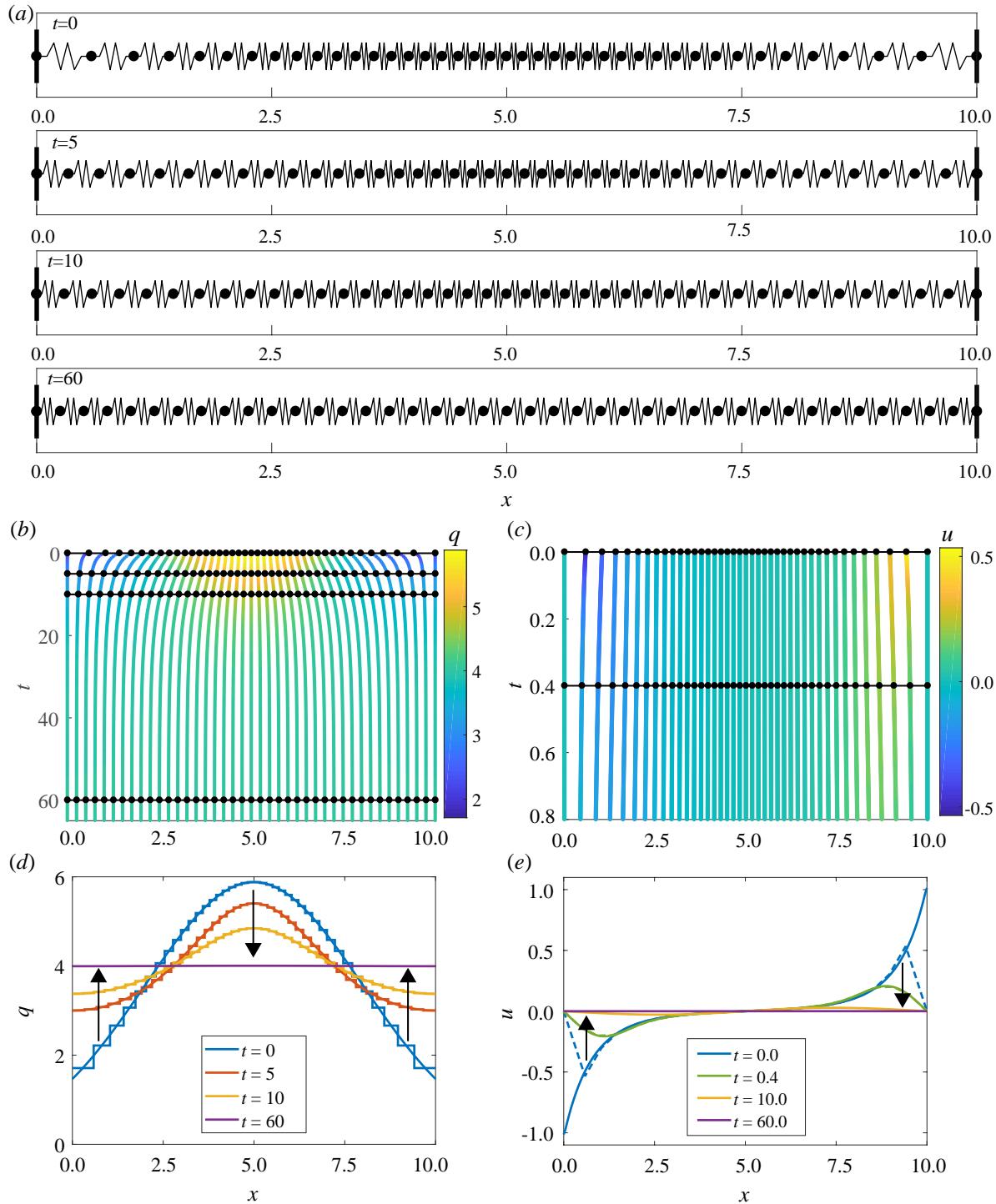
### 2.4.1 Homogeneous cell population

We first consider a homogeneous cell population, with one spring per cell,  $m = 1$ , to illustrate the time evolution of the cell density flux during mechanical relaxation of the tissue. To compare results from the discrete and continuum systems we choose the initial cell configuration (Figure 6.2(a)) to represent a normally distributed cell density,

$$q_0(x) = \frac{\lambda}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right), \quad 0 < x < L, \quad (2.4.1)$$

with mean position  $\mu = 5$  and variance  $\sigma = 3$ . We choose  $\lambda$  to satisfy  $\int_0^L q_0(x) dx = 40$  so that with  $L = 10$  the total number of cells is  $N = 40$ , see Supplementary Material Section 1. Then, using the discrete model, we observe that the system relaxes to a uniform cell distribution (Figure 6.2(a)). Figures 6.2(b) and 6.2(c) show how the density and velocity, respectively, propagate along the cell boundary characteristics and demonstrate that the system undergoes temporal relaxation to a steady state configuration. With an eigenmode decomposition of the governing equations of the discrete system, given by Equation (2.3.3) and the fixed boundary conditions, we find all eigenvalues are negative which explains the exponential decay behaviour.

We determine the discrete cell density as the inverse of the spacing between cell boundary trajectories,  $q_i = 1/(x_{i+1} - x_i)$  and we assign this value throughout the region  $x_i < x < x_{i+1}$ . We now compare this discrete information with the density from the continuum system,  $q$ , obtained by solving Equations (2A.2.1), (2A.2.2), (2A.2.3), and (2A.2.4). In Figure 6.2(d) we see that the initially normally distributed density tends to the uniform density  $\mathcal{Q}$ , given by  $\lim_{t \rightarrow \infty} q(x, t) = \mathcal{Q} = N/L$ , which is independent of  $k$  and  $a$ . From Equation (2A.2.1) we see that this motion is driven by imbalances in the local interaction force field. We relate this to the velocity,  $u = (\partial f / \partial x) / (\eta q)$  from Section 2.3, and we see that as the local imbalances tend to zero the cell boundary velocities tend to zero (Figure 6.2(e)). Due to fast dynamics followed by slow long-term dynamics, results for  $t = 10$  and  $t = 60$  are mostly overlapping with the steady state (Figure 6.2(e)). This agrees with the interpretation of the discrete system from Equation (2.3.3).



**Figure 2.2:** Results for homogeneous  $k$  and  $a$ , with  $N = 40$  and  $m = 1$ . (a) Snapshots of cell boundary positions and cell lengths at  $t = 0, 5, 15, 60$ . (b) Characteristic diagram for cell boundary position evolution for  $0 \leq t \leq 65$ . Colour denotes the cell density. Black lines with dots represent snapshots in (a) and (d). (c) Characteristic diagram for cell position evolution for  $0.0 \leq t \leq 0.8$ . Colour denotes velocity. Black lines and dots represent snapshots in (e). (d) Cell density snapshots at  $t = 0, 5, 10, 60$ . Results from discrete/continuum system displayed as stepped/solid lines. (e) Velocity snapshots at  $t = 0.0, 0.4, 10.0, 60.0$ . Results from discrete simulation and continuum system displayed as dashed/solid lines. Arrows indicate the direction of increasing time.

### 2.4.2 Heterogeneous cell population

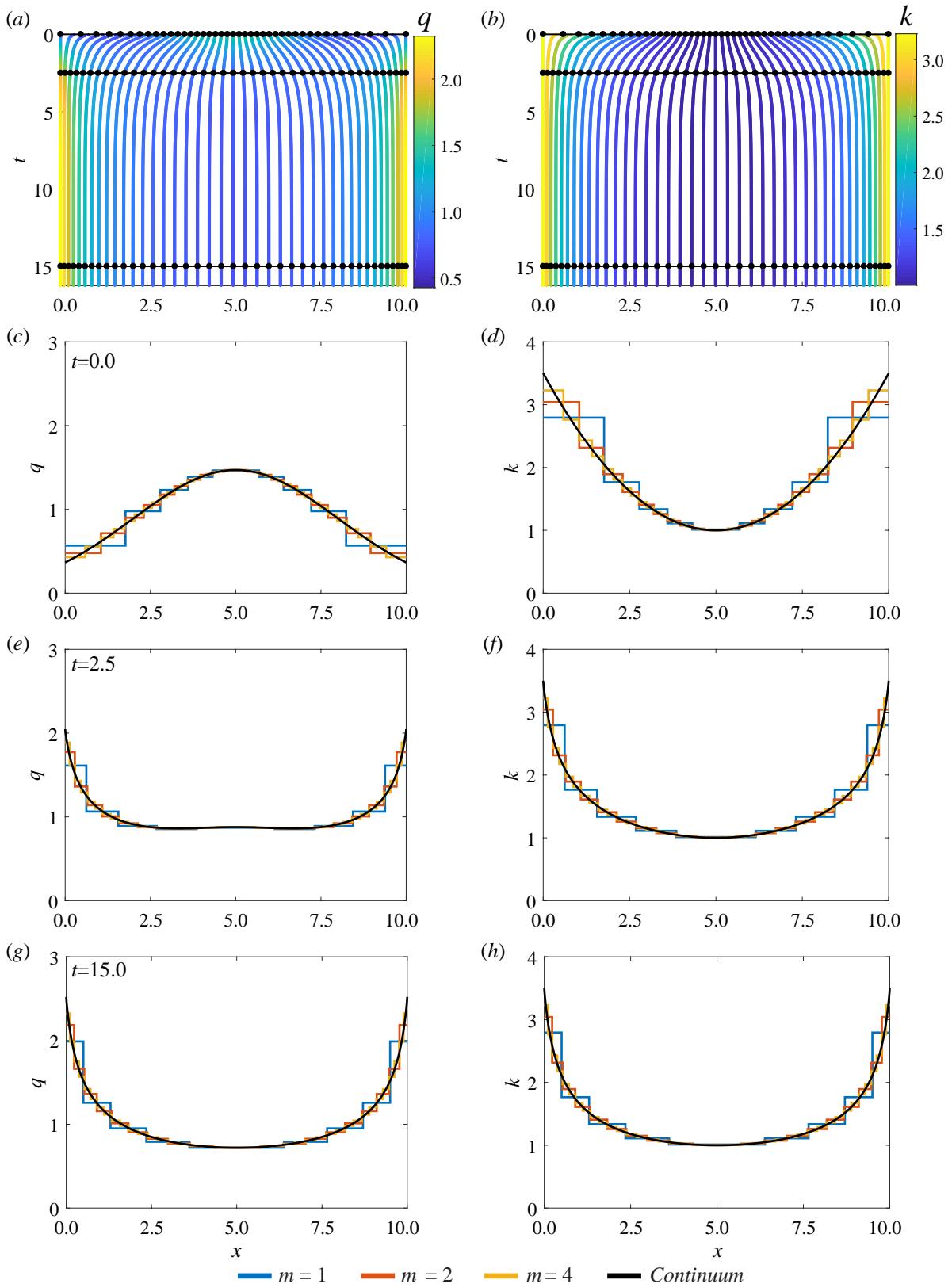
Here we present results for slowly-varying-in-space and piecewise constant heterogeneous cell populations.

#### Slowly varying cell population

For slowly-varying-in-space cellular properties, we explore how solutions of the discrete system converge to the solution of the continuum system as  $m$  increases. We consider heterogeneity in  $k$  and homogeneous  $a$  so that, on average, cells are in compression. Figure 6.3 depicts how the system relaxes to a non-uniform density distribution, due to cell stiffness heterogeneity, as the velocity field  $u$  tends to zero. From this simulation, we observe higher density in regions of higher  $k$ . This prediction agrees with the steady state solution to the coarse-grained model, governed by Equations (2A.2.1), (2A.2.2), (2A.2.3), and (2A.2.4),

$$\mathcal{Q}(x) = \frac{\mathcal{K}(x)}{b\eta + \mathcal{K}(x)\mathcal{A}(x)}, \quad (2.4.2)$$

where  $\mathcal{Q}(x) = \lim_{t \rightarrow \infty} q(x, t)$ ,  $\mathcal{K}(x) = \lim_{t \rightarrow \infty} k(x, t)$  and  $\mathcal{A}(x) = \lim_{t \rightarrow \infty} a(x, t)$ , are steady-state solutions and  $b$  is a constant of integration that is related to  $N$ . We also observe that, as cell properties are constant along trajectories, the cell stiffness evolves at a fixed location in space. We see in Figure 6.3(d-h) that there is close agreement between the discrete model and the continuum solutions as  $m$  increases. It is notable that even for low  $m$  we have excellent agreement between the discrete density and the continuum density at the centre of each spring. However, at spring boundaries the agreement does not hold as well for low  $m$ . We see similar discrete-continuum agreement when we consider other examples with heterogeneous  $k$  and homogeneous  $a$ , with homogeneous  $k$  and heterogeneous  $a$ , and heterogeneous  $k$  and heterogeneous  $a$  (Supplementary Figures S3-S6).



**Figure 2.3:** Results for heterogeneous  $k$  and homogeneous  $a$  with  $N = 10$ ,  $k_0(x) = 1 + 0.1(x - 5)^2$ , and  $a_0(x) = 0$ . (a,b) Characteristic diagram for spring boundary position evolution for  $0.00 \leq t \leq 16.25$ , with  $m = 4$  so that every fourth trajectory represents a cell boundary. Colour denotes (a) cell density, (b) cell stiffness. In (a,b) black lines and dots represent times for snapshots in (c-h). (c,e,g) Cell density snapshots at  $t = 0.0, 2.5, 15.0$ . (d,f,h) Cell stiffness snapshots at  $t = 0.0, 2.5, 15.0$ . In (c-h) lines display results for  $N = 10$  with  $m = 1, 2, 4$ , and continuum system.

### Piecewise constant cell population

In this section, we consider a simple scenario with two adjacent tissues, modelled by assuming sharp inhomogeneities in cellular properties. This may represent the boundary between a malignant tissue and a normal tissue. We first explore how solutions from the discrete system converge to the corresponding continuum solution as  $m$  increases, under these rapidly-varying-in-space conditions. Each tissue has homogeneous cell properties given by cell stiffnesses  $k_1, k_2$  and resting cell lengths  $a_1, a_2$  in the left and right tissue, respectively, with interface position  $s(t)$  (Figure 6.1(b)). For initial conditions, we choose a uniform density,  $q_0(x) = 1$ , cell properties  $k_1 = 1/2, k_2 = 1, a_1 = a_2 = 0, L = 10$  and  $s(0) = 5$ , respectively. The cell stiffness discontinuity rapidly induces a sharp change in the density at  $s(t)$  followed by slower dynamics until reaching a piecewise constant steady state as  $t \rightarrow \infty$  (Figure 6.4). Even with these sharp inhomogeneities we again observe close agreement between solutions of the discrete and continuum models, especially for the cell stiffness,  $k$ , where it is difficult to distinguish between the discrete model with different  $m$  and the solution of the continuum model. For the cell density,  $q$ , we again see that agreement at the spring boundaries improves as we increase  $m$ . This holds especially well given that the numerical discretisation of the continuum model does not explicitly follow the location of the interface, see Supplementary Material Section 2. It could however be determined by evaluating the velocity,  $ds(t)/dt = u$ , at the interface position.

This simple mechanical relaxation scenario between two tissues enables us to infer some information on the cellular-level properties  $k_i$  and  $a_i$  by considering the evolution of the interface position,  $s(t)$ . The steady state interface position,  $\mathcal{S} = \lim_{t \rightarrow \infty} s(x, t)$ , is given by

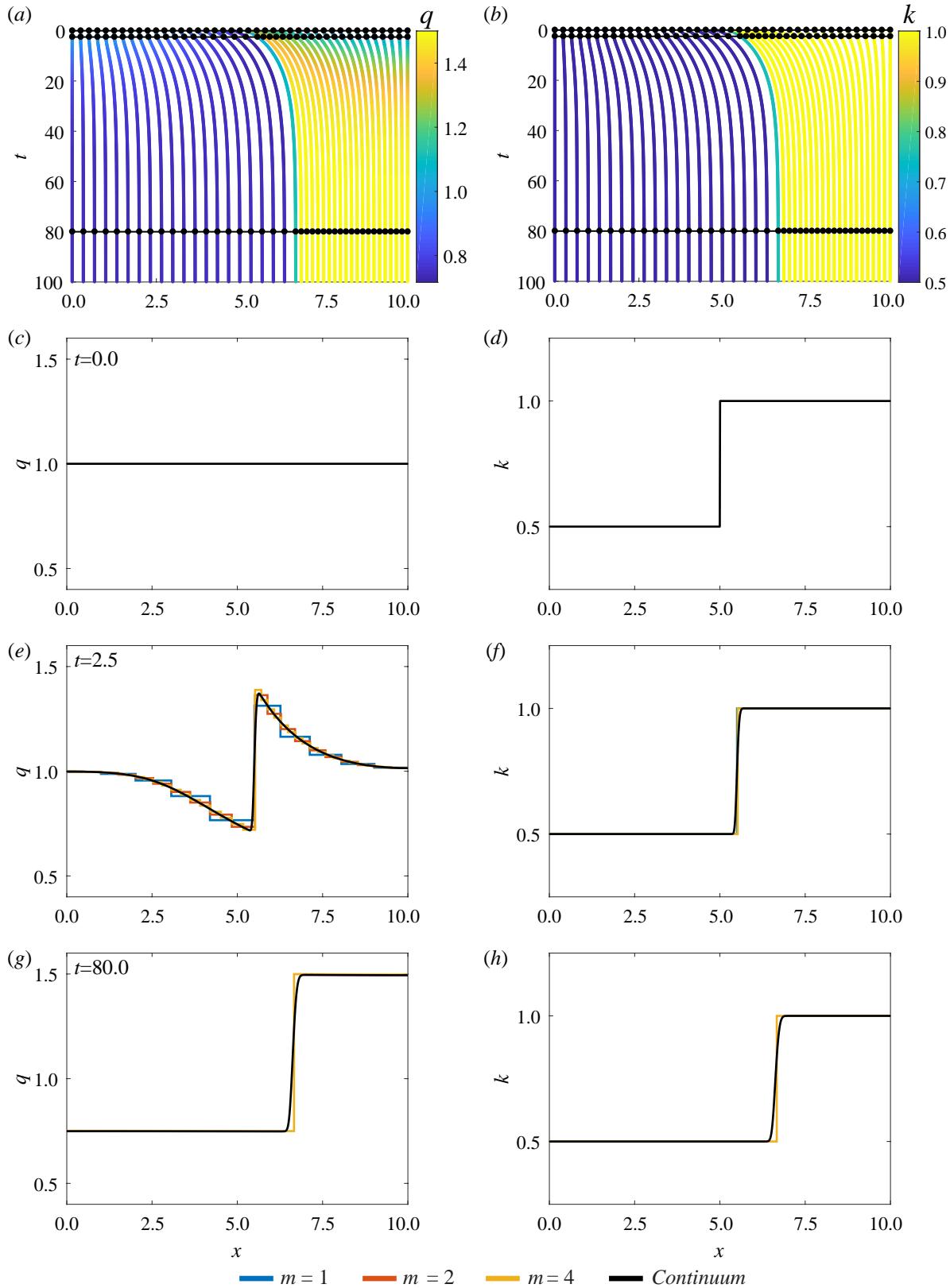
$$\mathcal{S} = \frac{\frac{k_1 a_1}{k_2} + \frac{L}{N_2} - a_2}{\frac{k_1}{k_2 N_1} + \frac{1}{N_2}}, \quad (2.4.3)$$

which depends on  $k_1/k_2$ ,  $a_1$  and  $a_2$ . Here  $N_1$  and  $N_2$  represent the total number of cells in the left and right tissues, respectively, see Supplementary Material Section 3. We can identify  $\mathcal{S}$  and  $L - \mathcal{S}$  as the lengths of the left and right tissues, respectively, after their mechanical relaxation.

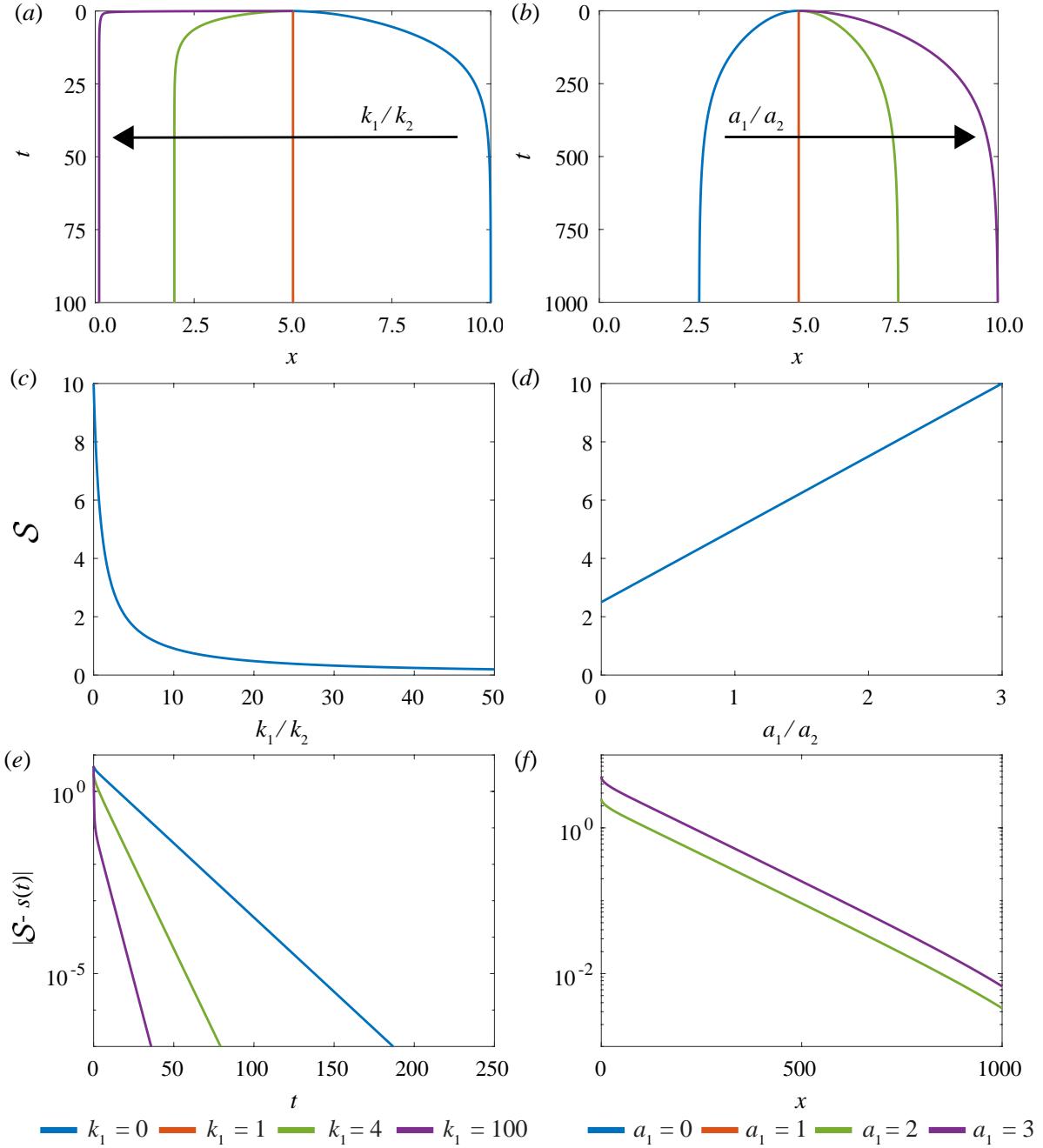
To investigate the influence of  $k_1/k_2$  we vary  $k_1$  and set  $k_2 = 1$ . As we have fixed boundaries at  $x = 0$  and  $x = L$ , we set  $a_1 = a_2 = 0$  to emphasise properties when we vary  $k_1$ , and choose a uniform density initial condition and  $N_1 = N_2 = 5$ . Evaluating  $s(t)$  numerically, for efficiency with the discrete model from Equation (2.3.3), and  $\mathcal{S}$  from Equation (2A.3.4), shows that if  $k_1 = 0$  then  $\mathcal{S} = L$  and the left tissue occupies the entire domain. As  $k_1 \rightarrow \infty$  then  $\mathcal{S} \rightarrow 0$  the length of the left tissue decreases (Figures 2.5(a,c)).

Similarly, to investigate the influence of  $a_1, a_2$  we consider  $a_1/a_2$ , vary  $a_1$  and set  $a_2 = 1$ . We set  $k_1 = k_2 = 1$  which only impacts the rate at which we reach the long-time solution. In contrast to varying  $k_1/k_2$ , steady state results depend on the choice of  $a_2$ , not just the ratio  $a_1/a_2$ , see Equation (2A.3.4). For example, when  $a_1 = 0$  then  $\mathcal{S} = 2.5$  which corresponds to a non-zero minimum left tissue length and a maximum length for the right tissue. We also observe that  $\mathcal{S}$  is proportional to  $a_1$  (Figures 2.5(b,d)).

We find that we can use the interface boundary velocities to infer cellular-level properties. Plotting  $|\mathcal{S} - s(t)|$  on a logarithmic scale against time shows that we can determine  $k_1/k_2$  from the gradient of the linear section and we can determine  $a_1/a_2$  from the  $y$ -intercept (Figure 2.5(e,f)). We find that it is easier to distinguish the ratio  $k_1/k_2$  than it is to distinguish the ratio  $a_1/a_2$ . If the second tissue was a reference material with known  $k_2, a_2$  we could then determine  $k_1, a_1$ .



**Figure 2.4:** Results for piecewise constant cell properties, with  $N = 10$ . (a,b) Characteristic diagram for spring boundary position evolution for  $0 \leq t \leq 100$ , with  $m = 4$  so that every fourth trajectory represents a cell boundary. Colour denotes (a) cell density, (b) cell stiffness. In (a,b) black lines and dots represent times for snapshots in (c-h). (c,e,g) Cell density snapshots at  $t = 0.0, 2.5, 80.0$ . (d,f,h) Cell stiffness snapshots at  $t = 0.0, 2.5, 80.0$ . In (c-h) lines display results from  $N = 10$  with  $m = 1, 2, 4$ , and continuum system.



**Figure 2.5:** Variation of relative cell stiffness,  $k_1/k_2$ , and relative resting cell length,  $a_1/a_2$ , in a model with two adjacent tissues and a constant density initial condition. (a) Characteristics of the interface position for varying  $k_1/k_2$ . The right tissue has fixed cell stiffness  $k_2 = 1$  while the left tissue cell stiffness is varied. (b) Characteristics of the interface particle for varying  $a_1/a_2$ . The right tissue has fixed resting cell length  $a_2 = 1$  while the resting cell length of the left tissue is varied. Analytical solution for the steady state position of the interface position with given (c) relative cell stiffness and (d) relative resting cell length. (e,f) Absolute difference between position and steady state for interface position for increasing time for varying (e) relative cell stiffness and (f) relative resting cell length.

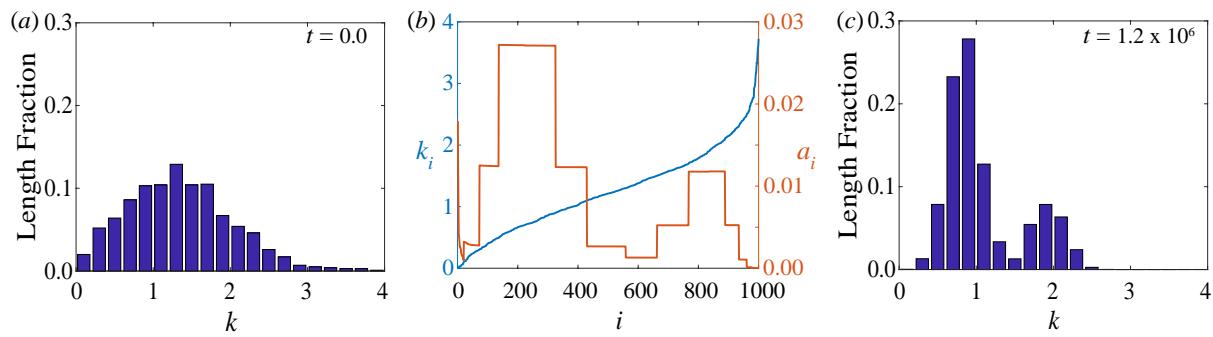
### 2.4.3 Case study: breast cancer detection

Recent experiments have proposed a new method to classify breast biopsies in situations where standard histological analysis is inconclusive [126, 128, 182]. The method is based on determining the stiffness histogram distribution of the tissue using atomic force microscopy. Normal tissues are associated with a single, well-defined unimodal stiffness peak, whereas malignant tissues are associated with a bimodal distribution with a prominent low-stiffness peak. Using our mathematical model, we are able to gain more insight into the differences in mechanical properties of normal and malignant tissues at the cellular level, in particular, the role of the resting cell length, which is not an easy quantity to measure experimentally. It would be impossible to interpret this experimental data with previous models that deal only with homogeneous cell populations.

For this case study, as the experimental data is relatively discrete, we use the discrete model, which we consider to be a sufficiently simple yet insightful portrayal of the biological details. We set the initial state of the system by assuming a uniform initial density distribution and by assigning the cell stiffness of the  $i^{\text{th}}$  cell,  $k_i$ , so as to reconstruct the unimodal stiffness profile from Figure 1b (top) in [182]. To do so, we normalise the experimental stiffness histogram and interpret the normalised value as the length fraction of the tissue containing stiffness in the given histogram bin (Figure 2.6(a)). This is consistent with the experimental method which implicitly assumes that the probability a cell is examined during a biopsy is proportional to its size [182]. To estimate  $k_i$ , we randomly sample the unimodal stiffness distribution and arbitrarily assign them to cells  $i = 1, 2, \dots, N$  in ascending order. Note that the ordering of the cells does not affect our results or the interpretation of our results in any way. We assume  $N = 1000$ ,  $m = 1$  and  $L = 10$  for illustration purposes. In order for this initial setup to be in equilibrium despite the heterogeneity in stiffness in the tissue, the resting cell lengths  $a_i$  must be chosen heterogeneously, per the steady state system of discrete equations, see Supplementary Material Section 4.

We proceed to consider how a bimodal stiffness distribution, associated with malignant tissues, could arise from such an initial state with a unimodal stiffness distribution. The simplest explanation is that a bimodal stiffness distribution may arise as a result of changes to individual cell stiffnesses,  $k_i$ , [81] e.g. due to some pre-cancerous biological mechanisms. This model provides an alternative interpretation where the bimodal distribution may arise from changes not solely to the individual cell stiffnesses but to the resting cell lengths also. We now present an extreme case where the bimodal distribution may arise from changes in the resting cell lengths only. Specifically, when we simulate the discrete model with the initial conditions as

above, but modify the heterogeneity in the resting cell lengths,  $a_i$ , to a bimodal profile with high  $a_i$  for very low  $k_i$ , without changing their stiffnesses,  $k_i$ , the cells redistribute themselves in the tissue in such a way that the tissue stiffness histogram develops a bimodal distribution at mechanical equilibrium (Figure 2.6(c)). We note that this result is not surprising due to the coupling of cell stiffness and resting cell length in the mathematical model. However, this intuitive result may not have been clear had we relied upon experimental data and experimental observations alone. In addition, this approach assumes that cells may have very different lengths which is consistent with biological observations. Specifically, it is understood that breast cancer cells are less stiff and, in general, have a larger diameter in comparison to normal breast cells [230]. This is consistent with other areas of biology, for example, in the context of melanoma biology it is well accepted that cancer cells can be smaller than healthy cells [92, 93]. We also note here that changes in the resting cell lengths have been assumed in other works [256] to model two-way feedback between mechanical tensions and signalling and here could similarly represent some unknown underlying pre-cancerous biological mechanisms.



**Figure 2.6:** Breast cancer detection case study. (a) Initial unimodal stiffness distribution, normalised by tissue length fraction, associated with normal tissues. (b) Initial cell stiffness  $k_i$  and modified resting cell length  $a_i$  for each cell  $i = 1, 2, \dots, 1000$ , leading to a bimodal stiffness distribution. (c) Steady-state stiffness distribution obtained with the modified resting cell lengths, exhibiting a bimodal distribution associated with malignant tissues.

## 2.5 Conclusion and future work

In this work, we present a one-dimensional cell-based model with heterogeneous cell properties, and its coarse-grained continuum approximation. The motion of cells is driven by cell-cell interaction forces which could represent cell-cell adhesion [109] or compressive stresses [231]. Heterogeneous cell properties, cell stiffness and resting cell length, are constant for each cell and are transported in space by cell movements. The continuum limit is taken by increasing the number of springs per cell, while maintaining the number of cells in the tissue and its fixed total length, and by considering spatial averages over length scales involving a large enough number of cells to define continuous densities but small enough to retain spatial heterogeneities.

Our results shows that solutions of the discrete model approach the solution of the continuum model as the number of springs per cell increases whilst the spring length tends to zero, even for rapidly varying spatial cell properties. Excellent agreement is observed even for few springs per cell at the centre of each cell. For the examples presented in this work, we find that the solution of the discrete model can be obtained much faster than the solution of the continuum model. However, the time required to simulate the discrete model increases rapidly with the number of cells. In contrast, the time required to simulate the continuum model is independent of the number of cells. Therefore, when we have large numbers of cells, as in an epithelial tissue, the continuum model is advantageous. Another advantage of the continuum model is that we can quickly develop exact closed form expressions for the long-time interface position which are more difficult to establish with the discrete model. Furthermore, the continuum model allows us to understand macroscale phenomena which are not obvious from microscopic interactions. The fact that the cell density flux in the continuum model, a macroscopic quantity, is explicitly related to the gradient of the cell-cell interaction force may have been anticipated, but it is not obvious from the microscopic interactions that this leads to an effective non-linear diffusive transport. Finally, because the continuum model exhibits explicit relationships between macroscopic quantities, it will be more useful for inverse problems.

By dealing explicitly with heterogeneous cell populations, this model has many potential applications. The first application we consider is a simple tissue relaxation simulation, where we track the position of the interface between two distinct adjacent tissues as the system mechanically relaxes, to infer cellular-level properties. Results suggest it is easier to determine the relative cell stiffnesses than it is to determine the relative resting cell lengths. Results also show that when cells are, on average, in tension a tissue with lower stiffness extends and compresses a tissue with higher stiffness. In the second application, we use the model to interpret recent experiments in breast cancer detection which reveal distinct stiffness profiles

associated with normal, benign and malignant tissues [182]. We show that a bimodal stiffness distribution, associated with a malignant tissue, could arise from a unimodal stiffness distribution, associated with a normal tissue, from changes not just in cell stiffnesses but from changes in the resting cell length's only. The resting cell length is not an easily measured experimental quantity and these results suggest that this could be an important variable to consider.

Many extensions of this work are possible, both mathematically and biologically. Important extensions will be to introduce cell proliferation (Chapters 3, 4, and 5), apoptosis (Chapters 3, 4, and 5), and free boundaries (Chapters 4, and 5) where the continuum limit is less obvious [19, 136, 164]. Another interesting extension will be to generalise the cell-cell interaction force law to include nonlinear effects for large separations [19, 165]. Finally, the model's ability to relate cellular-level stiffness data and tissue-level information has many potential extensions biologically including applying the model to particular scenarios such as epithelial tissue mechanics, cancer progression [170, 193], cancer detection [103, 182, 221], wound healing [61], and morphogenesis [66].

## **Chapter 2A**

# **Supplementary Material for Chapter 2**

<b>Contents</b>	<b>Page No.</b>
S1 Discrete model simulation	55
S1.1 Discrete equations	55
S1.2 Converting cell density into initial positions	55
S1.3 Assigning spring properties	55
S1.4 Numerical methods	56
S2 Continuum model simulation	57
S2.1 Discretisation scheme	57
S3 Steady state analysis for two tissue model	59
S4 Breast cancer detection case study: model implementation	60
S4.1 Choosing the resting spring length to choose the steady state	60
S5 Supplementary Figures	61

## 2A.1 Discrete model simulation

In this section, we describe how to perform simulations with the discrete model.

### 2A.1.1 Discrete equations

Here, we describe the method to solve Equation (2A.1.1), governing the discrete model, with  $m$  springs per cell, for the position of each spring boundary,  $x_i^{(\nu)}$ , for  $i = 1, 2, \dots, N$  and  $\nu = 1, 2, \dots, m$  and  $x_{N+1}^{(1)}$ . The equations are included here for convenience:

$$\eta^{(\nu)} \frac{dx_i^{(\nu)}}{dt} = f_i^{(\nu+1)} - f_i^{(\nu)}; \quad (2A.1.1)$$

$$f_i^{(\nu)} = k_i^{(\nu-1)} \left( x_i^{(\nu)} - x_i^{(\nu-1)} - a_i^{(\nu-1)} \right). \quad (2A.1.2)$$

### 2A.1.2 Converting cell density into initial positions

We now explain how to convert an initial distribution of density,  $q_0(x)$ , into an initial condition for the discrete model by determining the initial spring boundary positions,  $x_i^{(\nu)} = x_i^{(\nu)}(0)$ , for  $i = 1, 2, \dots, N$  and  $\nu = 1, 2, \dots, m$ , and  $x_{N+1}^{(1)}$ . To solve for these positions we use the Matlab `fsolve` function [5] applied to the system of the equations

$$x_1^{(1)} = 0, \quad (2A.1.3)$$

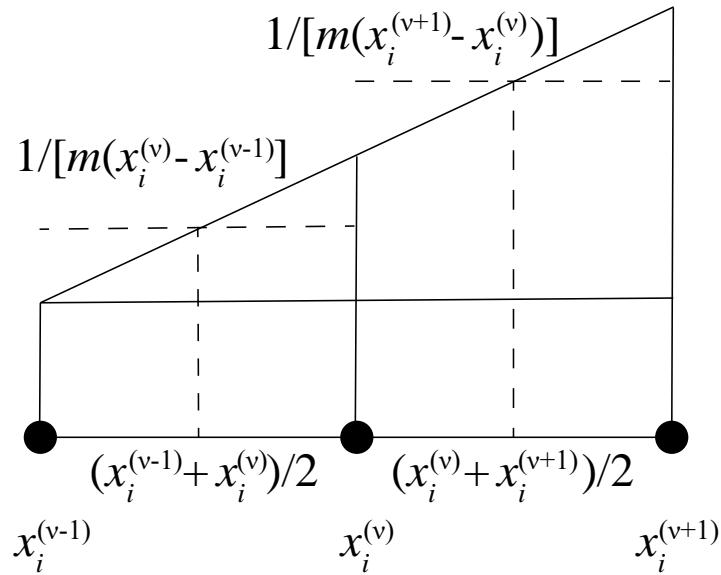
$$\frac{\frac{1}{x_i^{(\nu-1)} - x_i^{(\nu)}} + \frac{1}{x_i^{(\nu)} - x_i^{(\nu+1)}}}{\frac{x_i^{(\nu+1)} - x_i^{(\nu-1)}}{2}} - m \frac{\partial q_0(x)}{\partial x} \Big|_{x=x_i^\nu} = 0, \quad (2A.1.4)$$

$$x_{N+1}^{(1)} = L. \quad (2A.1.5)$$

Equations (2A.1.3) and (2A.1.5) arise from the fixed boundary conditions. Equation (2A.1.4) arises from equating the approximate numerical gradient of the density from the discrete system at a position  $x_i^{(\nu)}$  with the gradient of  $q_0(x)$  at the same position. To evaluate the numerical gradient we use the midpoints of the domains  $x_i^{(\nu-1)} < x < x_i^{(\nu)}$  and  $x_i^{(\nu)} < x < x_{i+1}^{(\nu+1)}$ . The density in the first domain is given by  $1 / [m (x_i^{(\nu)} - x_i^{(\nu-1)})]$  and similarly  $1 / [m (x_{i+1}^{(\nu+1)} - x_i^{(\nu)})]$  for the second domain (Figure S1).

### 2A.1.3 Assigning spring properties

In this section we explain our approach to assigning spring properties assuming that we know the initial cell boundary positions,  $x_i^{(\nu)}$ , the initial cell stiffness distribution,  $k_0(x)$ , and initial



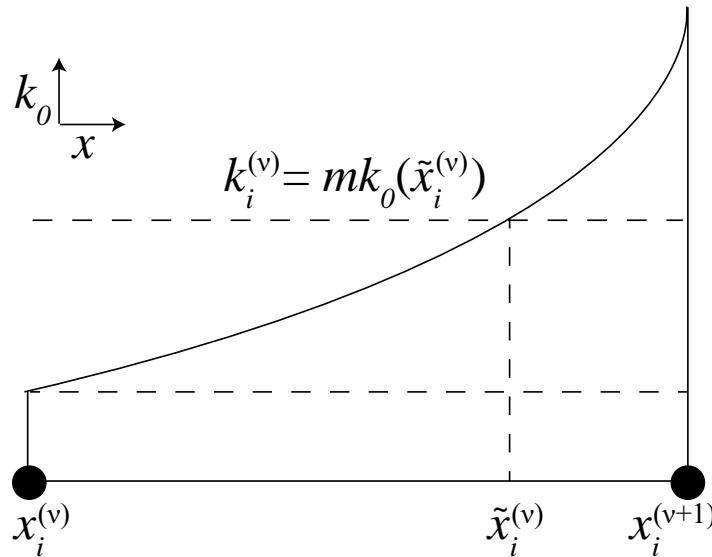
**Figure 2A.1:** Schematic to determine the approximate numerical gradient of the density from the discrete system at position  $x_i^v$ .

resting cell length distribution,  $a_0(x)$ . Here we consider the cell stiffness initial condition and note that the same ideas apply to assign the resting cell length. We consider spring  $\nu$  in cell  $i$ , located at  $x_i^{(\nu)} < x < x_i^{(\nu+1)}$ ; let  $\tilde{x}_i^{(\nu)}$  be the position of the median of  $k_0(x)$  in this domain, and use this to define the discrete spring stiffness as  $k_i^{(\nu)} = m k_0(\tilde{x}_i^{(\nu)})$  (Figure S2). For the resting cell length this would be  $a_i^{(\nu)} = a_0(\tilde{x}_i^{(\nu)})/m$ . This median position,  $\tilde{x}_i^{(\nu)}$ , is solved for by equating the integral of the initial condition spring stiffness function in  $x_i^{(\nu)} < x < \tilde{x}_i^{(\nu)}$  and  $\tilde{x}_i^{(\nu)} < x < x_i^{(\nu+1)}$ ,

$$\int_{x_i^{(\nu)}}^{\tilde{x}_i^{(\nu)}} k_0(x) dx = \int_{\tilde{x}_i^{(\nu)}}^{x_i^{(\nu+1)}} k_0(x) dx. \quad (2A.1.6)$$

## 2A.1.4 Numerical methods

This discrete model, with  $m$  springs per cell, is governed by Equations (2A.1.1) with the fixed boundary conditions  $x_1^{(1)} = 0$ ,  $x_{N+1}^{(1)} = L$ . Appropriate scalings of the cell properties are required to determine the spring properties,  $k_i^{(\nu)} = m k_0(\tilde{x}_i^{(\nu)})$  and  $a_i^{(\nu)} = a_0(\tilde{x}_i^{(\nu)})/m$ . The viscosity coefficient must also be scaled appropriately through  $\eta^{(\nu)} = \eta/m$ . These form a system of  $Nm - 1$  ordinary differential equations, with the two boundary conditions, and initial conditions for the positions,  $x_i^{(\nu)}$ , spring stiffnesses,  $k_i^{(\nu)}$ , and resting spring lengths,  $a_i^{(\nu)}$ , of each spring  $\nu$  in cell  $i$ , and viscosity coefficient for the system,  $\eta^{(\nu)}$ . We solve this system using Matlab ode15s [199]. Alternatively, this system can be analysed with an eigenmode decomposition which is convenient to determine the steady state.



**Figure 2A.2:** Schematic to determine the position  $\tilde{x}_i^{(\nu)}$  used to define the discrete spring stiffness value for spring  $\nu$  in cell  $i$ ,  $k_i^{(\nu)} = m k_0(\tilde{x}_i^{(\nu)})$ .

## 2A.2 Continuum model simulation

Here we describe the numerical method used to solve the partial differential equations associated with the continuum description.

### 2A.2.1 Discretisation scheme

For convenience we re-state the governing equations:

$$\frac{\partial q(x, t)}{\partial t} = -\frac{1}{\eta} \frac{\partial^2 f(x, t)}{\partial x^2}; \quad (2A.2.1)$$

$$\frac{\partial k(x, t)}{\partial t} = -\frac{1}{\eta} \frac{1}{q(x, t)} \frac{\partial f}{\partial x} \frac{\partial k(x, t)}{\partial x}; \quad (2A.2.2)$$

$$\frac{\partial a(x, t)}{\partial t} = -\frac{1}{\eta} \frac{1}{q(x, t)} \frac{\partial f}{\partial x} \frac{\partial a(x, t)}{\partial x}; \quad (2A.2.3)$$

$$f(x, t) = k(x, t) \left( \frac{1}{q(x, t)} - a(x, t) \right). \quad (2A.2.4)$$

The first step is to substitute the interaction force from Equation (2A.2.4) into Equations (2A.2.1), (2A.2.2) and (2A.2.3). We solve Equation (2A.2.1) implicitly, and we solve Equations (2A.2.2) and (2A.2.3) for the cell stiffness and resting cell length, respectively, explicitly [50]. First, we uniformly discretise the domain with nodes spaced  $\Delta x$  apart. The nodes are indexed  $j = 1, \dots, R$ , where  $R$  is the total number of spatial nodes. We apply an upwinding scheme using a numerically determined velocity at each node, which is defined for node  $j$  and time

step  $n$  as  $v_j^n$ , given by

$$v_j^n = \frac{1}{\eta} \frac{1}{q_j^n} \left[ k_{j+1}^n \left( \frac{1}{q_{j+1}^n} - a_{j+1}^n \right) - k_{j-1}^n \left( \frac{1}{q_{j-1}^n} - a_{j-1}^n \right) \right]. \quad (2A.2.5)$$

We use non-constant time stepping for efficiency with the timestep to advance the numerical solution from timestep  $n$  to timestep  $n+1$  denoted  $\Delta t^n$ . The value of this timestep is determined based on the maximum numerical velocity across all nodes at time  $n$ ,  $\max_j(v_j^n)$ , and is chosen as  $\Delta t^n = \min(0.001(\Delta x)^2, 0.00001(\Delta x)^2 / \max_j(v_j^n))$ , where  $\Delta x = 0.01$  to produce Figure 4 and Supplementary Figures S3, S4 and S6, and  $\Delta x = 0.05$  otherwise.

We solve Equation (2A.2.1) using a Crank-Nicolson approximation. At the central nodes we have

$$\begin{aligned} \frac{q_j^{n+1} - q_j^n}{\Delta t^n} &= -\frac{1}{2} \frac{1}{(\Delta x)^2} \left[ \frac{k_{j+1}^n}{\eta} \left( \frac{1}{q_{j+1}^n} - a_{j+1}^n \right) - 2 \frac{k_j^n}{\eta} \left( \frac{1}{q_j^n} - a_j^n \right) + \frac{k_{j-1}^n}{\eta} \left( \frac{1}{q_{j-1}^n} - a_{j-1}^n \right) \right] \\ &\quad - \frac{1}{2} \frac{1}{(\Delta x)^2} \left[ \frac{k_{j+1}^{n+1}}{\eta} \left( \frac{1}{q_{j+1}^{n+1}} - a_{j+1}^{n+1} \right) - 2 \frac{k_j^{n+1}}{\eta} \left( \frac{1}{q_j^{n+1}} - a_j^{n+1} \right) + \frac{k_{j-1}^{n+1}}{\eta} \left( \frac{1}{q_{j-1}^{n+1}} - a_{j-1}^{n+1} \right) \right], \\ j &= 3, \dots, R-2, \quad n = 0, \dots, T, \end{aligned} \quad (2A.2.6)$$

where we use the approximation

$$\frac{1}{q_j^{n+1}} = \frac{q_j^{n+1}}{\left( q_j^n \right)^2}, \quad (2A.2.7)$$

for the terms at timestep  $n+1$  on the right-hand side of Equation (2A.2.6). This approximation allows us to write the discretised system of equations in tridiagonal form. For the boundary condition at  $x = 0$ , corresponding to node  $j = 1$ , we apply a second order forward difference stencil so that

$$q_1^n = 1 \Bigg/ \left[ a_1^n + \frac{4k_2^n}{3k_1^n} \left( \frac{1}{q_2^n} - a_2^n \right) - \frac{k_3^n}{3k_1^n} \left( \frac{1}{q_3^n} - a_3^n \right) \right]. \quad (2A.2.8)$$

To obtain the equation for node  $j = 2$  we set  $j = 2$  in Equation (2A.2.6) and replace  $q_1^n$  with Equation (2A.2.8). Similarly, to obtain an equation for node  $j = R$  at the right boundary  $x = L$  we apply a second order backwards difference stencil,

$$q_R^n = 1 \Bigg/ \left[ a_R^n - \frac{k_{R-2}^n}{3k_R^n} \left( \frac{1}{q_{R-2}^n} - a_{R-2}^n \right) + \frac{4k_{R-1}^n}{3k_R^n} \left( \frac{1}{q_{R-1}^n} - a_{R-1}^n \right) \right]. \quad (2A.2.9)$$

This allows us to form an equation for node  $j = R-1$  also. Now we use the Thomas algorithm

[254] to advance one time step for the equations governing nodes  $j = 2, \dots, R - 1$ . Using the results for nodes  $j = 2, \dots, R - 1$  and Equations (2A.2.8) and (2A.2.9) we can update the boundary nodes  $q_1^{n+1}, q_R^{n+1}$ .

While performing each temporal step we must also update the cell stiffnesses,  $k_j^n$ , and resting cell lengths,  $a_j^n$ , at each node. At the boundaries,  $x = 0$  and  $x = L$ , corresponding to nodes  $j = 1$  and  $j = R$ , these cell properties are fixed so  $k_1^n = k_1^0$  and  $a_1^n = a_1^0$  for  $n = 1, \dots, T$ . For the interior nodes  $j = 2, \dots, R - 1$ , we apply an explicit method with upwinding. The sign of  $\max_j(v_j^n)$  determines whether we apply forward or backward difference stencils. For example, if  $\max_j(v_j^n) > 0$  then we apply a backward first order difference to the cell stiffness, Equation (2A.2.2),

$$\frac{k_j^{n+1} - k_j^n}{\Delta t^n} = -\frac{1}{\eta} \frac{k_j^n - k_{j-1}^n}{(\Delta x)^2} \left[ k_j^n \left( \frac{1}{q_j^n} - a_j^n \right) - k_{j-1}^n \left( \frac{1}{q_{j-1}^n} - a_{j-1}^n \right) \right], \quad j = 2, \dots, R - 1. \quad (2A.2.10)$$

Cell properties are constant along cell boundary trajectories. Therefore, if we have a cell property which is initially homogeneous it will remain homogeneous. The numerical method can then be simplified by not simulating the related cell property equation and replacing its appearance in the other equations with its constant value.

### 2A.3 Steady state analysis for two tissue model

We obtain an analytical expression for the steady state position of the interface,  $\mathcal{S}$ , between two distinct adjacent tissues. Suppose that the left-most tissue is characterised by  $N_1$  cells with stiffness  $k_1$  and resting spring length  $a_1$ , for  $0 < x < \mathcal{S}$ . Similarly, suppose that the right-most tissue is characterised by  $N_2$  cells with stiffness  $k_2$  and resting spring length  $a_2$  for  $\mathcal{S} < x < L$ . Considering the continuum system given by Equations (2A.2.1), (2A.2.2), (2A.2.3), and (2A.2.4) for each tissue at steady state, the equations governing the evolution of the cellular properties are trivially solved as we have a homogeneous cell population in each tissue. However, solving Equation (2A.2.1) and applying the no flux density boundary conditions gives, for the first tissue,

$$\frac{k_1}{\eta} \left( \frac{1}{q_1} - a_1 \right) = c_1, \quad 0 < x < \mathcal{S}, \quad (2A.3.1)$$

where  $c_1$  is an arbitrary constant. At steady state the forces at the interface,  $x = \mathcal{S}$ , are at equilibrium, giving,

$$\frac{k_1}{\eta} \left( \frac{1}{q_1} - a_1 \right) = \frac{k_2}{\eta} \left( \frac{1}{q_2} - a_2 \right). \quad (2A.3.2)$$

Relating this mesoscopic density to the microscopic density for the tissue we have

$$q_1 = \frac{N_1}{\mathcal{S}}, \quad q_2 = \frac{N_2}{L - \mathcal{S}}. \quad (2A.3.3)$$

Substituting Equation (2A.3.3) into Equation (2A.3.2) and rearranging gives

$$\mathcal{S} = \frac{\frac{k_1 a_1}{k_2} + \frac{L}{N_2} - a_2}{\frac{k_1}{k_2 N_1} + \frac{1}{N_2}}. \quad (2A.3.4)$$

## 2A.4 Breast cancer detection case study: model implementation

Here we present our method to obtain a user specified steady state cell stiffness histogram consistent with the initial  $k_i$  by choosing the initial condition for the resting spring length,  $a_i$ .

### 2A.4.1 Choosing the resting spring length to choose the steady state

The possible final steady state spring stiffness histogram distributions must be consistent with the initial  $k_i$  as spring properties are constant along cell trajectories. For illustrative purposes we choose the steady state spring stiffness histogram distribution we wish to obtain as guided by experimental results [182]. With this choice we can read off the histogram frequencies that describe the coverage of the each histogram interval at steady state. For simplicity, we assume that cells with  $k_i$  in the same histogram interval are of equal length at steady state. Then the length of each cell in a histogram interval is given by the total length of cells in the histogram interval divided by the number of cells in the interval. As we have an initial ordering of the cells and each cell length at steady state we can now readily determine the steady state cell boundary positions. Returning to Equation (2A.1.1), with a single spring per cell, at steady state and the fixed boundary conditions  $x_1 = 0, x_{N+1} = L$ , we now know every  $x_i$  and  $k_i$  and we can solve this system of  $N + 1$  nonlinear equations to find each  $a_i$ , using `fsolve` in MATLAB [5]. A simulation is then initiated with these  $a_i$ . This simulation reaches the steady state spring stiffness histogram distribution we chose to obtain.

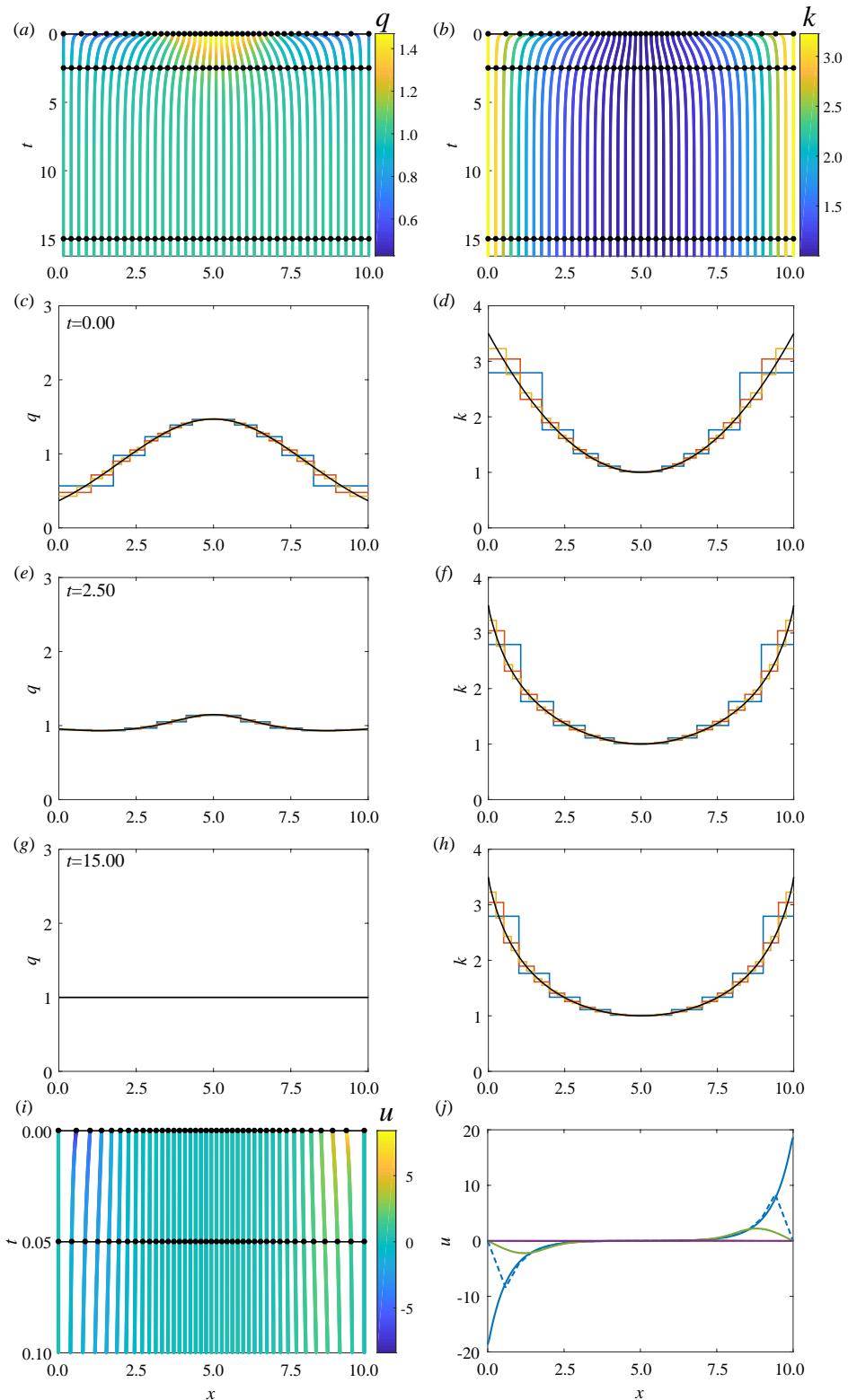
## 2A.5 Supplementary Figures

In this section we present the supplementary figures which are referenced in the main text.

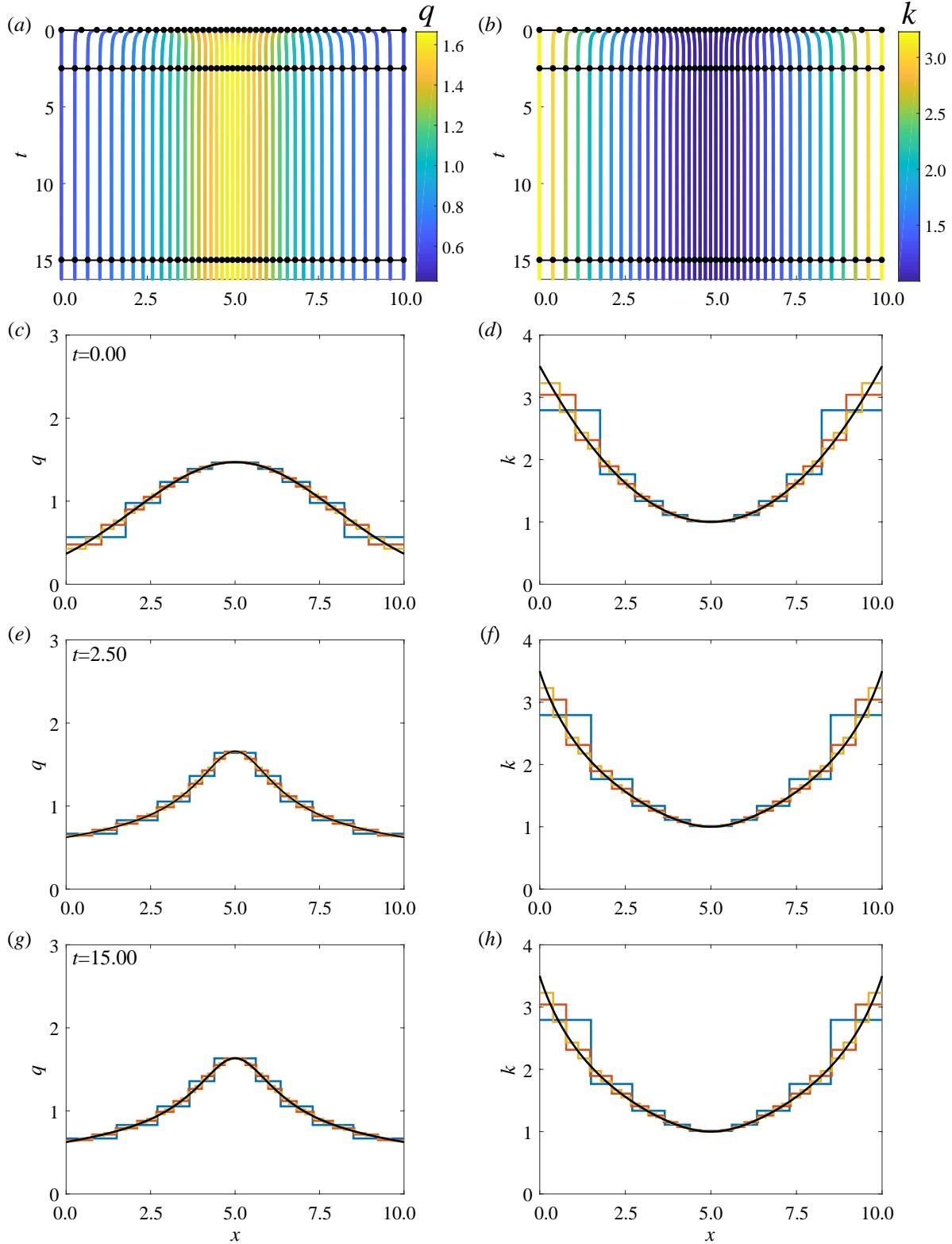
*Heterogeneous  $k$  and homogeneous  $a$ .* Even though  $a$  is homogeneous, the resting cell length is still a key factor as it determines whether the system is, on average, in extension or in compression. This is determined by comparing the resting spring length,  $a_i^{(\nu)}$ , with the critical value,  $a_{\text{crit}} = L/(mN)$ , which is the average length of a spring. In Figure 3 we consider  $a < a_{\text{crit}}$  so that on average cells are in extension. We present other cases, where  $a = a_{\text{crit}}$  and  $a > a_{\text{crit}}$ , in Supplementary Figures S3 and S4. We choose the initial cell stiffness distribution to be  $k_0(x) = 1 + 0.1[x - (L/2)]^2$ . Figure S3, depicts how the system relaxes to a uniform density distribution, due to the combination of  $a = a_{\text{crit}}$  and the cell stiffness heterogeneity, as the velocity field  $u$  tends to zero.

It is notable that even for low  $m$  we have excellent agreement between the discrete density and the continuum density especially at the centre of each spring. However, agreement at the spring boundaries does not hold as well for low  $m$ .

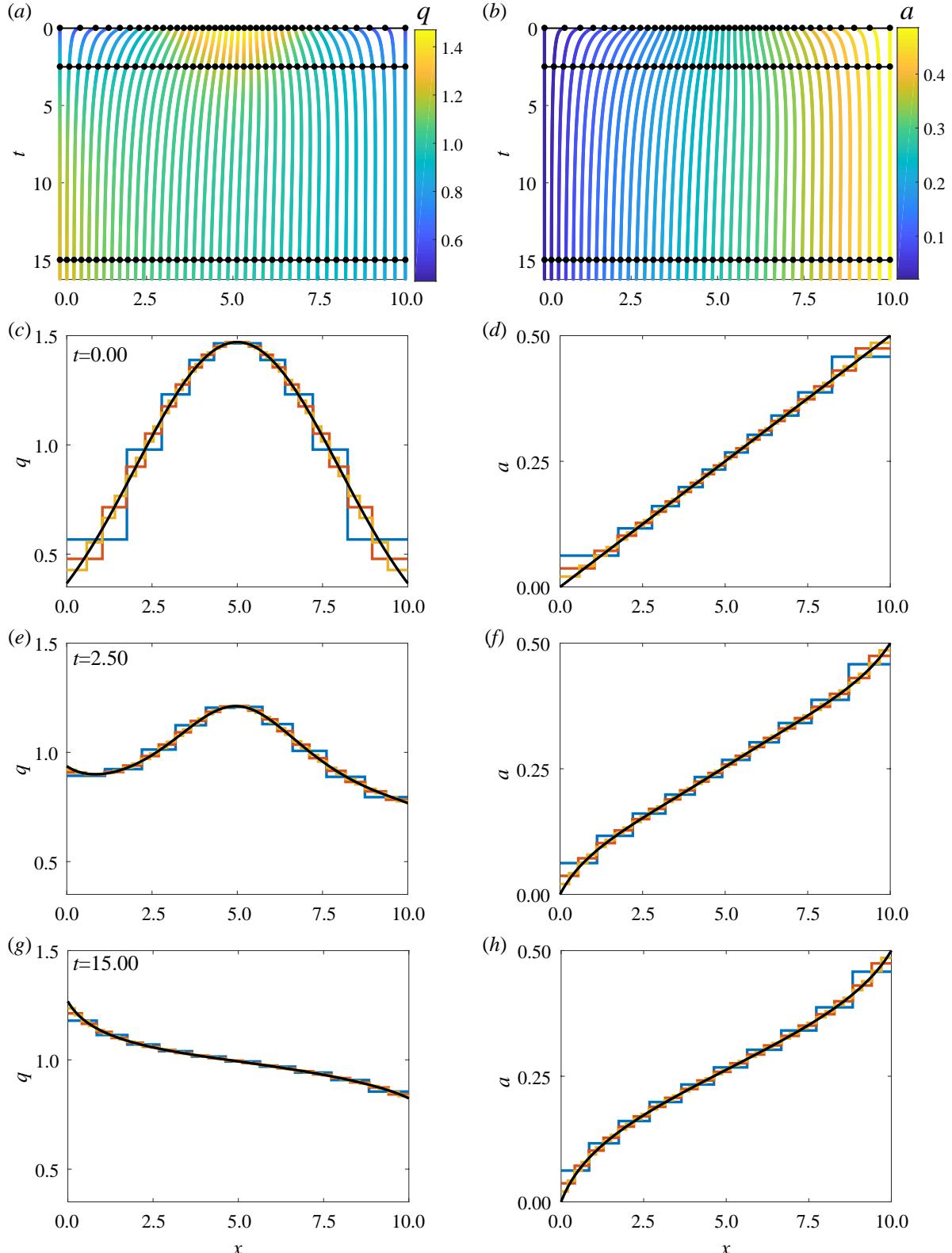
*Heterogeneous  $a$ .* We see similar discrete-continuum agreement when we consider examples with homogeneous  $k$  and heterogeneous  $a$ , and heterogeneous  $k$  and heterogeneous  $a$ , see Supplementary Figures S5 and S6, respectively. We observe higher cell density in regions of lower  $a$  which agrees with the steady state solution to the coarse-grained model, Equation (3.2) in the main paper.



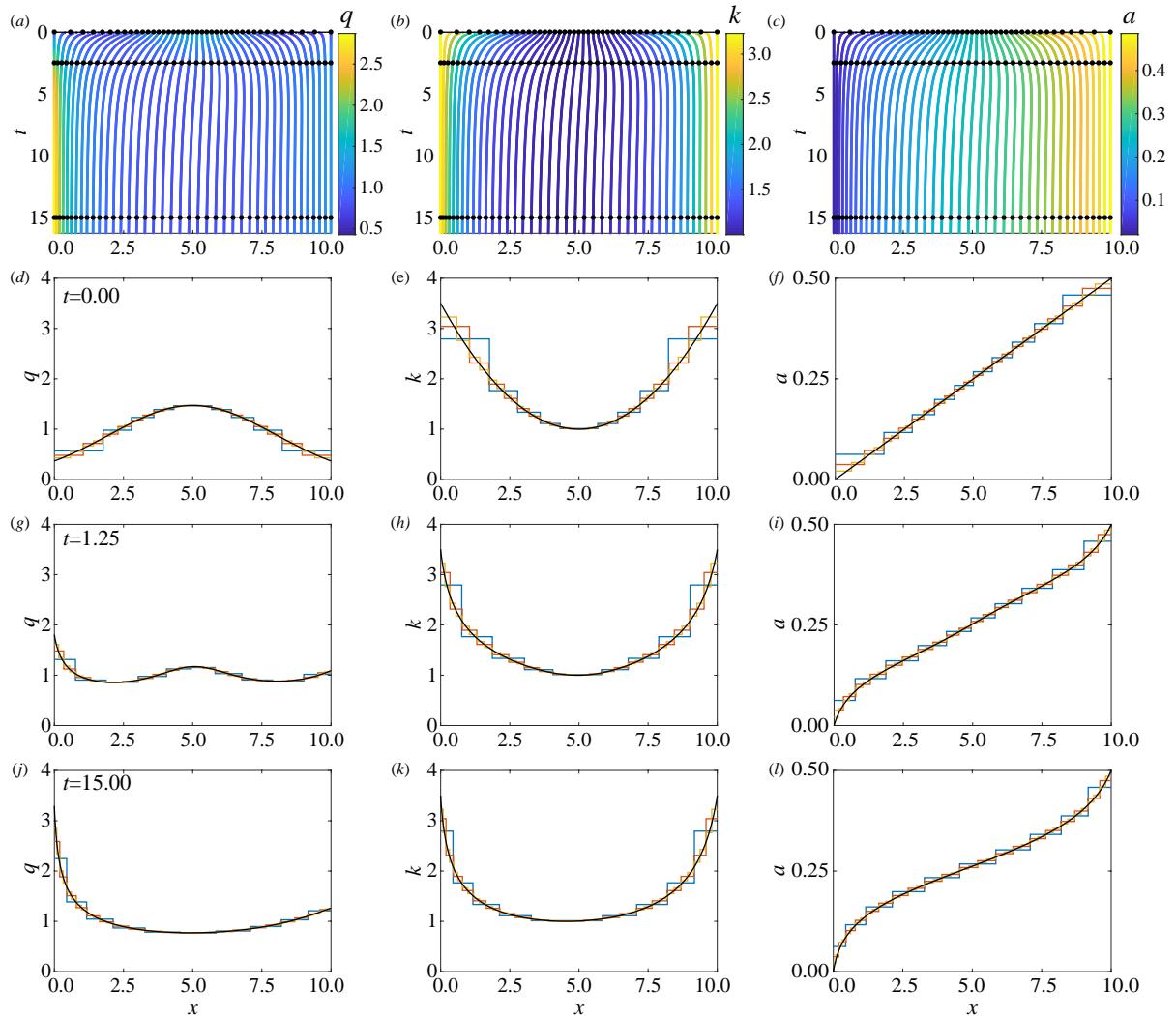
**Figure 2A.3:** Results for heterogeneous cell stiffness and homogeneous cell spring length for cells in extension, with  $N = 10$ ,  $k_0(x) = 1 + 0.1(x - 5)^2$ , and  $a_0(x) = 1$ . (a,b) Characteristic diagram for spring boundary position evolution for  $0.00 \leq t \leq 16.25$  with  $m = 4$  so that every fourth trajectory represents a cell boundary. Colour denotes (a) cell density, (b) cell stiffness. In (a,b) black lines and dots represent times for snapshots in (c-h). (c,e,g) Cell density snapshots at  $t = 0.00, 2.50, 15.00$ . (d,f,h) Cell stiffness snapshots at  $t = 0.00, 2.50, 15.00$ . In (c-h) lines display results from  $m = 1$  (blue),  $2$  (red),  $4$  (yellow), and continuum system (black). (i) Characteristic diagram for spring boundary position evolution for  $0.00 \leq t \leq 0.10$ . Colour denotes velocity. (j) Velocity snapshots at  $t = 0.00$  (blue),  $0.05$  (green),  $0.25$  (yellow),  $0.15$  (magenta). Dashed/solid line represent solutions from discrete model with  $m = 4$  and continuum model, respectively.



**Figure 2A.4:** Results for heterogeneous cell stiffness and homogeneous resting cell length, for cells on average in compression, with  $N = 10$ ,  $k_0(x) = 1 + 0.1(x - 5)^2$ , and  $a_0(x) = 2$ . (a,b) Characteristic diagram for spring boundary position evolution for  $0.00 \leq t \leq 16.25$  with  $m = 4$  so that every fourth trajectory represents a cell boundary. Colour denotes (a) cell density, (b) cell stiffness. In (a,b) black lines and dots represent times for snapshots in (c-h). (c,e,g) Cell density snapshots at  $t = 0.00, 2.50, 15.00$ . (d,f,h) Cell stiffness snapshots at  $t = 0.00, 2.50, 15.00$ . In (c-h) lines display results from  $m = 1$  (blue), 2 (red), 4 (yellow), and continuum system (black).



**Figure 2A.5:** Results for homogeneous  $k$  and heterogeneous  $a$  for cells on average not in extension or compression, with  $N = 10$ ,  $k_0(x) = 1$ , and  $a_0(x) = 0.05x$ . (a,b) Characteristic diagram for spring boundary position evolution for  $0.00 \leq t \leq 16.25$ , with  $m = 4$  so that every fourth trajectory represents a cell boundary. Colour denotes (a) cell density, (b) resting cell length. In (a,b) black lines and dots represent times for snapshots in (c-h). (c,e,g) Cell density snapshots at  $t = 0.00, 2.50, 15.00$ . (d,f,h) Resting cell length snapshots at  $t = 0.00, 2.50, 15.00$ . In (c-h) lines display results from  $N = 10$  with  $m = 1$  (blue), 2 (red), and continuum system (black).



**Figure 2A.6:** Results for heterogeneous cell stiffness and heterogeneous resting cell length, with  $N = 10$ ,  $k_0(x) = 1 + 0.1(x - 5)^2$  and  $a_0(x) = 0.05x$ . (a,b,c) Characteristic diagram for spring boundary position evolution for  $0.00 \leq t \leq 16.25$  with  $m = 4$  so that every fourth trajectory represents a cell boundary. Colour denotes (a) cell density, (b) cell stiffness, (c) resting cell length. In (a,b,c) black lines and dots represent times for snapshots in (d-l). (d,g,j) Cell density snapshots at  $t = 0.00, 1.25, 15.00$ . (e,h,k) Cell stiffness snapshots at  $t = 0.00, 1.25, 15.00$ . (f,i,l) Resting cell length snapshots at  $t = 0.00, 1.25, 15.00$ . In (d-l) lines display results from  $m = 1$  (blue), 2 (red), 4 (yellow), and continuum system (black).



## **Chapter 3**

# **Mechanical cell competition in heterogeneous epithelial tissues**

### 3.0 Preamble

An article published in *Bulletin of Mathematical Biology*

**Murphy RJ**, Buenzli PR, Baker RE, Simpson MJ (2020). Mechanical cell competition in heterogeneous epithelial tissues. *Bulletin of Mathematical Biology*. 82:130. doi:10.1007/s11538-020-00807-x. bioRxiv preprint

This chapter includes Publication 2, addresses objective 2 and research question 2. This chapter extends the work of Chapter 2, by incorporating cell proliferation and cell death into the model of Chapter 1. The key results include a discrete model describing mechanical cell competition in heterogeneous epithelial tissues, a derivation to obtain the corresponding the continuum model, and the continuum model. Using these new models we examine under what conditions the discrete and continuum model show good agreement, explore the evolution of homogeneous tissues and heterogeneous tissues with a cancerous cells competing for space with healthy cells. We advise why naively adding source/sink terms to a continuum model without considering the underlying discrete model may lead to incorrect results. Supplementary material, such as additional results for homogeneous and heterogeneous populations and numerical methods, associated with publication 2, is presented in Chapter 3A.

### Statement of Contribution of Co-Authors for Thesis by Published Paper

The authors listed below have certified that:

1. they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
2. they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
3. there are no other authors of the publication according to these criteria;
4. potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit, and
5. they agree to the use of the publication in the student's thesis and its publication on the QUT's ePrints site consistent with any limitations set by publisher requirements.

In this case of this chapter:

Murphy RJ, Buerzli PR, Baker RE, Simpson MJ (2020). Mechanical cell competition in heterogeneous epithelial tissues. *Bulletin of Mathematical Biology*. 82:130. doi:10.1007/s11538-020-00807-x.

Contributor	Statement of Contribution
Ryan J. Murphy (Candidate)  <i>R. Murphy</i> 02/09/2021	Conceived and designed the study, derived the continuum model, developed the codes for numerical simulation of the discrete and continuum models, performed numerical simulations, generated results, interpreted results, drafted the manuscript, and revised the manuscript during the peer-review process. Acted as corresponding author.
Pascal R. Buerzli	Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research.
Ruth E. Baker	Conceived and designed the study and provided comments and gave final approval for publication.
Matthew J. Simpson	Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research.

#### Principal Supervisor Confirmation

I have sighted email or other correspondence from all Co-authors confirming their certifying authorship.

Matthew Simpson

2/9/2021

Name

Signature

Date

### 3.1 Abstract

Mechanical cell competition is important during tissue development, cancer invasion, and tissue ageing. Heterogeneity plays a key role in practical applications since cancer cells can have different cell stiffness and different proliferation rates than normal cells. To study this phenomenon, we propose a one-dimensional mechanical model of heterogeneous epithelial tissue dynamics that includes cell-length-dependent proliferation and death mechanisms. Proliferation and death are incorporated into the discrete model stochastically and arise as source/sink terms in the corresponding continuum model that we derive. Using the new discrete model and continuum description, we explore several applications including the evolution of homogeneous tissues experiencing proliferation and death, and competition in a heterogeneous setting with a cancerous tissue competing for space with an adjacent normal tissue. This framework allows us to postulate new mechanisms that explain the ability of cancer cells to outcompete healthy cells through mechanical differences rather than an intrinsic proliferative advantage. We advise when the continuum model is beneficial and demonstrate why naively adding source/sink terms to a continuum model without considering the underlying discrete model may lead to incorrect results.

## 3.2 Introduction

In cell biology, epithelial tissues are continuously experiencing forces and replacing cells, through cell proliferation and death, to maintain homeostasis. These tissues can be naturally heterogeneous or heterogeneous due to cancer development and progression [89, 182]. This heterogeneity is observed at multiple scales, from sub-cellular to cellular to the tissue scale [227], and can result in cell competition. Cell competition can act as a quality control mechanism in tissue development or as a defence against precancerous cells, and harnessing cell competition has been suggested as a possible approach to enhance both cell-based cancer and regenerative therapies [184]. Therefore, gaining a greater understanding of the mechanisms underlying cell competition is very desirable. In mathematical models of cell competition the classical hypothesis is that cells compete due to differences in their intrinsic proliferation rates. However, different mechanisms, such as mechanical cell competition, may play a role [125]. We will explore mechanical cell competition.

In the emerging field of mechanical cell competition, *winner* cells compress neighbouring cells promoting tissue crowding and regions of higher density, which leads to cell death [29, 129, 241], while cell proliferation occurs in regions of lower density [82]. In this work, we focus on mechanical cell competition arising from the coupling of a cell-based model of epithelial tissue mechanics with cell-length-dependent proliferation and death mechanisms. We consider mechanical forces to be driven by cell stiffness which is important for cancer progression [193], cancer detection [182], morphogenesis [66], and wound healing [61]. A grand challenge in cell biology is to understand how tissue-level outcomes are related to cell-based mechanisms, especially when experiments are performed by focusing on a single scale, and many cellular processes occur over multiple overlapping timescales [38, 248]. Therefore, we apply mathematical modelling with *in silico* simulations to develop a framework to quantitatively connect cell-level mechanisms with tissue-level outcomes.

Many mathematical modelling frameworks, including both discrete models and continuum models, have been used to study cell migration and cell proliferation. In discrete models individual cell properties and inter-cellular interactions can be prescribed [175, 179]. However, discrete models often lack macroscopic intuition and can be computationally intensive, especially with proliferation and death included, which are commonly stochastic and require many realisations to understand the average behaviour. Continuum models commonly include proliferation and death through source/sink terms and may require constitutive equations to close the system [14, 21, 78, 129, 144, 154, 189, 202]. In general, continuum models do not make the underlying cell-level processes clear [75]. However, continuum models can be less com-

putationally expensive than discrete models and can be analysed with well-established mathematical techniques such as stability analysis [17], phase plane analysis and perturbation techniques [58, 123, 159].

We are most interested starting with discrete descriptions of individual cell dynamics and properties and then deriving corresponding continuum models [27, 68, 147, 174, 181, 235, 253] because this allows us to switch between the two spatial scales and take advantage of both. Further, this approach is very insightful as it can be used to demonstrate conditions when continuum models are valid and when they are not valid. Having a continuum model which is more computationally efficient to solve than the discrete model, and which well-established mathematical techniques can be applied to, is only beneficial if the continuum model accurately represents the underlying discrete behaviour. In this work, we start with the model of mechanical relaxation in heterogeneous epithelial tissues from Murphy et al. (2019) (Chapter 2, [157]) and now incorporate cell-length-dependent proliferation and death mechanisms. This framework allows us to explore mechanical cell competition, which was not previously possible when considering only homogeneous populations [19, 164–167] or two populations without cell death [136].

This work is structured as follows. In Section 5.3, we present a new discrete mechanical model that includes cell-length-dependent proliferation and death mechanisms. We then derive the corresponding novel continuum model that takes the form of a system of coupled nonlinear partial differential equations with both hyperbolic and parabolic properties. In Section 3.4.1, we explore our novel model by considering the evolution of a homogeneous tissue where cells are undergoing both proliferation and death. In Section 3.4.2, we explore mechanical cell competition in the context of cancer invasion by considering a heterogeneous tissue composed of both cancerous and normal cells that compete for space. Using the model we explore whether cancer cells will eventually replace all of the healthy cells or can the cancer cells coexist with the healthy cells? In Section 3.4.3, we demonstrate the importance of the discrete to continuum approach.

### 3.3 Model formulation

In this section, we focus on how we stochastically implement cell proliferation and death for heterogeneous cell populations within the discrete mechanical framework and the derivation of the continuum description.

#### 3.3.1 Discrete model

We start by describing the mechanical model and then include proliferation and death. We consider a one-dimensional cross-section of an epithelial tissue, and represent it as a one-dimensional chain of cells, connected at cell boundaries, in a fixed domain of length  $L$ . The cells experience cell-cell interaction forces at their cell boundaries, for example cell-cell adhesion [109] or compressive stresses [231]. For a system of  $N$  cells, cell  $i$  has left and right cell boundaries at positions  $x_i^N(t)$ ,  $x_{i+1}^N(t)$ , respectively. Fixed boundary conditions at  $x = 0$  and  $x = L$  are imposed  $x_1^N(t) = 0$  and  $x_{N+1}^N(t) = L$ . To allow for heterogeneous tissues, each cell  $i$ , which can be thought of here as a mechanical spring, is prescribed with intrinsic cell properties including a cell stiffness,  $k_i^N$ , and resting cell length,  $a_i^N$  (Figure 3.1a). We assume cell motion occurs in a viscous environment such that cell boundaries experience a drag force with mobility coefficient  $\eta > 0$  (Chapter 2, [67, 147, 157]). In the overdamped regime, where inertia effects are neglected, the evolution of cell boundary  $i$  in a system of  $N$  cells is

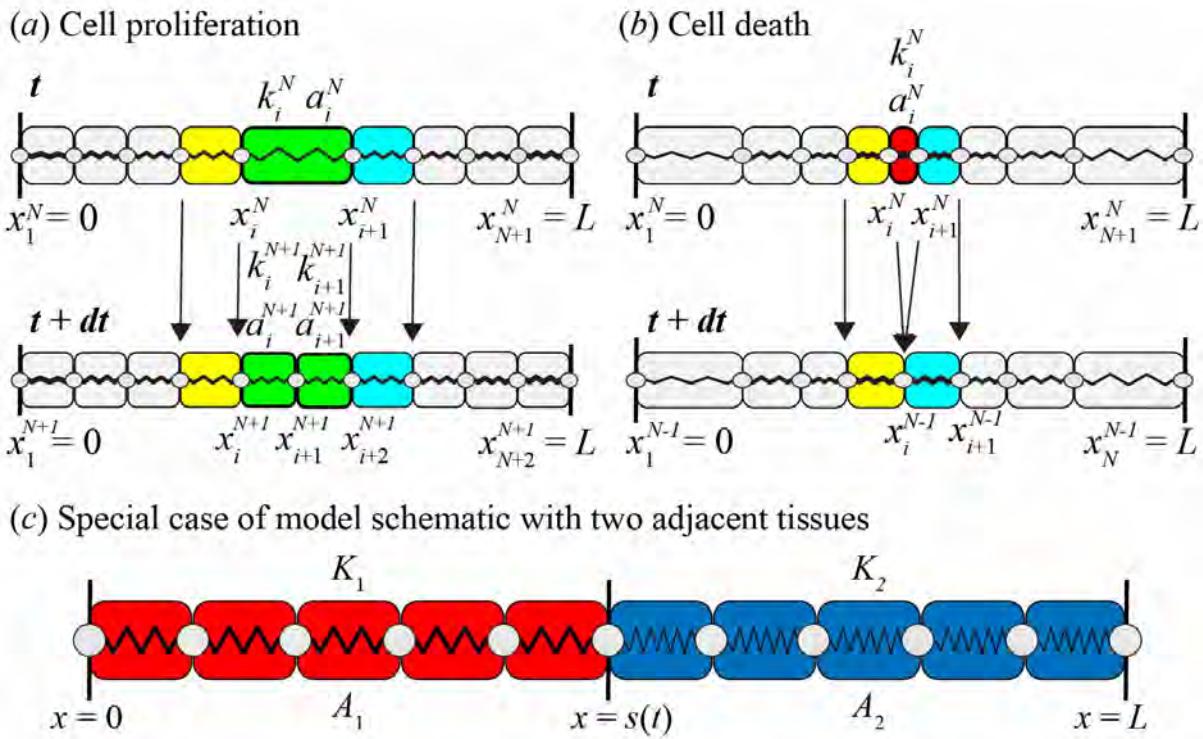
$$\eta \frac{dx_i^N(t)}{dt} = f_i^N(l_i^N(t)) - f_{i-1}^N(l_{i-1}^N(t)), \quad i = 2, \dots, N, \quad (3.3.1)$$

where  $f_i^N(l_i(t))$  is the force exerted on cell  $i - 1$  by cell  $i$  (Chapter 2, [157]). When  $f_i^N(l_i(t)) > 0$  cell  $i$  contracts and pulls cell  $i - 1$ . When  $f_i^N(l_i(t)) < 0$  cell  $i$  extends and pushes cell  $i - 1$ . This cell-cell interaction force law may be given by, for example, a cubic, Hertz, Lennard-Jones, or Johnson-Kendall-Roberts law [19, 136, 165]. However, for simplicity, we choose a Hookean force law,

$$f_i^N(l_i^N(t)) = k_i^N [l_i^N(t) - a_i^N], \quad (3.3.2)$$

where cell  $i$  has length  $l_i^N(t) = x_{i+1}^N(t) - x_i^N(t) > 0$ .

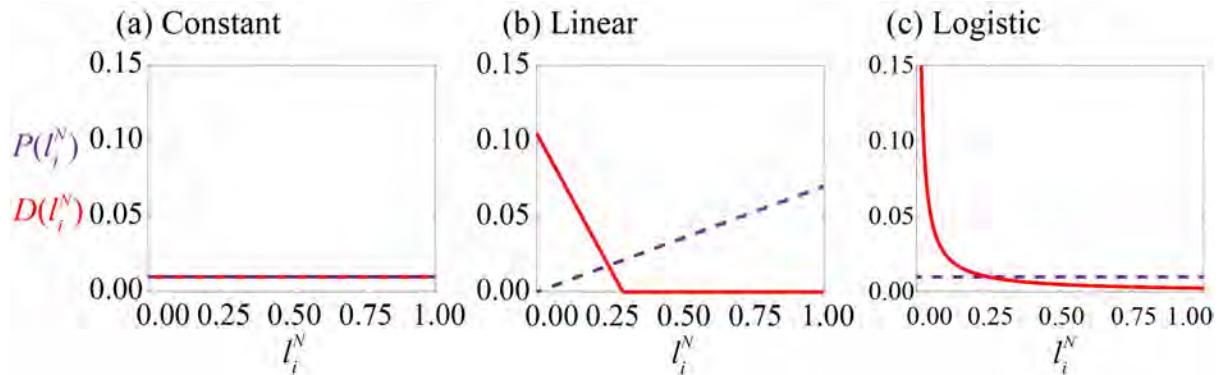
We include cell proliferation stochastically, by considering that cell  $i$  proliferates with probability  $P(l_i^N(t))dt$  in the small time interval  $[t, t + dt]$ , that depends on the current cell length,  $l_i^N(t)$ , and proliferation mechanism  $P(\cdot)$  [19, 185]. When cell  $i$  proliferates we increase the number of cells by one by introducing a new cell boundary,  $x_{i+1}^{N+1}$ , at the midpoint of the original cell,



**Figure 3.1:** Discrete model schematic for a heterogeneous epithelial tissue with cell proliferation and death. Cell  $i$  in a system of  $N$  cells has left and right cell boundaries  $x_i^N(t)$ ,  $x_{i+1}^N(t)$ , with  $x_i^N(t) < x_{i+1}^N(t)$ , respectively, and is prescribed with a cell stiffness  $k_i^N > 0$ , and a resting cell length  $a_i^N \geq 0$ . (a) Cell proliferation. Cell  $i$ , coloured green, is selected to proliferate at time  $t$ . At time  $t + dt$ , the cell has proliferated with a new cell boundary introduced at the midpoint of the original cell. Cell properties of the daughter cell are prescribed from the parent cell. (b) Cell death. Cell  $i$ , coloured red, is selected to die at time  $t$ . At time  $t + dt$ , the cell is removed and the cell boundaries of cell  $i$  at time have coalesced at midpoint of the original cell. For both proliferation and death cells are re-indexed at time  $t + dt$ . (c) Special case with two adjacent tissues. The left tissue (tissue 1) is coloured red and the right tissue (tissue 2) is coloured blue. The interface position between the left and right tissues is  $x = s(t)$ . Each cell in tissue  $i$  has cell stiffness  $K_i$  and resting cell length  $A_i$ . Proliferation and death rates remain dependent on the length of each cell. This could also represent a single tissue with internal heterogeneity.

**Table 3.1:** Proliferation and death mechanisms written in terms of cell length,  $l_i^N$ , proliferation parameter,  $\beta$ , and death parameters,  $\gamma, l_d$ .

	<b>Constant</b>	<b>Linear</b>	<b>Logistic</b>
$P(l_i^N)$	$\beta$	$\beta l_i^N$	$\beta$
$D(l_i^N)$	$\gamma$	$\begin{cases} \gamma(l_d - l_i^N), & 0 \leq l_i^N \leq l_d \\ 0, & l_d < l_i^N \end{cases}$	$\frac{\gamma}{l_i^N}$



**Figure 3.2:** Proliferation and death mechanisms considered in this work. Proliferation rates,  $P(l_i^N)$  (dashed), and death rates,  $D(l_i^N)$  (solid), shown as a function of cell length,  $l_i^N$ . Parameters used in this work: (a)  $\beta = 0.01, \gamma = 0.01$ , (b)  $\beta = 0.07, \gamma = 0.35, l_d = 0.3$ , (c)  $\beta = 0.01, \gamma = 0.0025$ .

$x_{i+1}^{N+1} = (x_i^N + x_{i+1}^N)/2$ , and relabel indices accordingly (Figure 3.1a). Daughter cells take the same intrinsic cell properties as the parent cell. Cell death is included similarly to cell proliferation with a cell-length-dependent death mechanism,  $D(l_i^N(t))$ . In a system of  $N+1$  cells, when cell  $i$  dies, with cell boundaries  $x_i^{N+1}$  and  $x_{i+1}^{N+1}$ , the number of cells is reduced by one. The two cell boundaries are set to instantly coalesce at the midpoint of the dying cell (Figure 3.1b). Cell death at the tissue boundaries needs to be considered separately (Supplementary Material SM1.2). In this work, we consider constant, linear, and logistic models of proliferation and death (Table 3.1, Figure 3.2). We solve discrete Equations (3.3.1) together with a stochastic implementation of proliferation and death numerically (Supplementary Material SM2.1).

### 3.3.2 Derivation of continuum model

To understand the mean behaviour of the discrete model we must average over many identically prepared stochastic realisations. However, this can be computationally intensive, especially for large  $N$ . The corresponding continuum model, which we first present and then derive, represents the average behaviour and unlike the discrete model the computational time required to solve the continuum model is independent of  $N$ .

The continuum model for the evolution of the cell density,  $q(x, t)$ , in terms of the continuous cell-cell interaction force,  $f(x, t)$ , proliferation rate,  $P(1/q(x, t))$ , and death rate,  $D(1/q(x, t))$ , is the conservation of mass equation

$$\frac{\partial q(x, t)}{\partial t} = \underbrace{-\frac{1}{\eta} \frac{\partial^2 f(x, t)}{\partial x^2}}_{\text{mechanical relaxation}} + \underbrace{q(x, t)P\left(\frac{1}{q(x, t)}\right)}_{\text{proliferation}} - \underbrace{q(x, t)D\left(\frac{1}{q(x, t)}\right)}_{\text{death}}, \quad (3.3.3)$$

where the continuous cell-cell interaction force which corresponds to Equation (3.3.2) is given by

$$f(x, t) = k(x, t) \left( \frac{1}{q(x, t)} - a(x, t) \right), \quad (3.3.4)$$

with cell stiffness,  $k(x, t)$ , and the resting cell length,  $a(x, t)$ , also being described by continuous fields. From Equation (3.3.3), we know that the cell density flux,  $j(x, t) = q(x, t)u(x, t)$ , is equal to the spatial gradient of the cell-cell interaction force,  $(1/\eta)\partial f/\partial x$ . Therefore, the cell velocity,  $u(x, t)$ , is related to the cell density and gradient of the cell-cell interaction force through

$$u(x, t) = \frac{1}{\eta q(x, t)} \frac{\partial f(x, t)}{\partial x}. \quad (3.3.5)$$

Note that Equation (3.3.5) corresponds to the discrete linear momentum equation in Equation (3.3.2). Intrinsic mechanical cell properties are constant for each cell and transported by the motion of cells. The proliferation and death functions,  $P(\cdot)$  and  $D(\cdot)$ , respectively, (Table 3.1) are evaluated at  $1/q(x, t)$ . Depending on the choice of proliferation and death mechanisms we may have additional intrinsic cellular properties,  $\beta(x, t)$ ,  $\gamma(x, t)$ , and  $l_d(x, t)$ . All intrinsic cellular properties evolve according to the following transport equation,

$$\frac{\partial \chi(x, t)}{\partial t} + u(x, t) \frac{\partial \chi(x, t)}{\partial x} = 0, \quad \chi = k, a, \beta, \gamma, l_d, \quad (3.3.6)$$

where  $u(x, t)$  is the cell velocity. The left hand side of Equation (3.3.6) corresponds to the

material derivative, expressing the fact that there is no change in cellular properties along cell trajectories. We solve the system of Equations (3.3.3)-(3.3.6) together with initial conditions and boundary conditions numerically (Supplementary Material SM2.2).

We now systematically derive Equation (3.3.3). We take care to explicitly state and make clear all approximations made in this derivation. We incorporate proliferation and then death into the modelling framework, under the assumption that the two processes are independent. The previously derived mechanical relaxation term and transport of cellular property equations (3.3.6) are only briefly discussed here, please refer to Murphy et al. (2019) (Chapter 2, [157]) for full details. For clarity, the derivation is shown for one spring per cell. However, this analysis can be extended to  $m > 1$  springs per cell which, for sufficiently small  $N$ , is a more appropriate method to define the continuous field functions (Chapter 2, [157], Supplementary Material SM1.3).

## Proliferation

As cell proliferation is included stochastically (Sections 5.3.1, SM2.1), we consider an infinitesimal time interval  $[t, t + dt]$  and condition on the possible proliferation events that could occur and influence the position of cell boundary  $i$  in a system of  $N$  cells. Choosing  $dt$  sufficiently small so that at most one proliferation event can occur in  $[t, t+dt)$ , there are four possibilities: i) there is no proliferation, in which case the cell boundary position  $x_i^N$  only changes by mechanical relaxation; ii) there is proliferation to the right of cell  $i - 1$ ; iii) there is proliferation to the left of cell  $i - 1$ ; and iv) cell  $i - 1$  proliferates. This leads to the following infinitesimal evolution law for the position of cell boundary  $x_i^N$ , accounting for cell relabelling when a new cell is added:

$$\begin{aligned}
 x_i^N(t + dt) = & \left[ x_i^N(t) + \frac{dt}{\eta} \{ f_i^N(l_i^N) - f_{i-1}^N(l_{i-1}^N) \} \right] \times \mathbb{1}\{\text{no proliferation}\} \\
 & + \left[ x_{i-1}^{N-1}(t) + \frac{dt}{\eta} \{ f_i^{N-1}(l_i^{N-1}) - f_{i-1}^{N-1}(l_{i-1}^{N-1}) \} \right] \\
 & \quad \times \mathbb{1}\{\text{proliferation right of cell } i - 1\} \\
 & + \left[ x_{i-1}^{N-1}(t) + \frac{dt}{\eta} \{ f_{i-1}^{N-1}(l_{i-1}^{N-1}) - f_{i-2}^N(l_{i-2}^N) \} \right] \\
 & \quad \times \mathbb{1}\{\text{proliferation left of cell } i - 1\} \\
 & + \left[ \frac{x_i^{N-1}(t) + x_{i-1}^{N-1}(t)}{2} + \frac{dt}{2\eta} \{ f_i^{N-1}(l_i^{N-1}) - f_{i-2}^N(l_{i-2}^N) \} \right] \\
 & \quad \times \mathbb{1}\{\text{proliferation of cell } i - 1\}.
 \end{aligned} \tag{3.3.7}$$

Each term in square brackets is the resulting force from neighbouring cells due to mechanical relaxation, given by Equations (3.3.1), for each potential event. In addition, we include Boolean random variables expressed as indicator functions,  $\mathbb{1}\{\cdot\}$ , defined as

$$\mathbb{1}\{\text{event}\} = \begin{cases} 1, & \text{if event occurs in } [t, t + dt), \\ 0, & \text{otherwise,} \end{cases} \quad (3.3.8)$$

whose expectations in the context of Equation (3.3.7) can be interpreted as proliferation probabilities. For a system of  $N$  cells, where  $dt$  is sufficiently small, these proliferation probabilities are given by

$$\mathbb{P}(\text{no proliferation in } [t, t + dt)) = 1 - dt \sum_{j=1}^N P(l_j^N), \quad (3.3.9a)$$

$$\mathbb{P}(\text{proliferation to the right of cell } i-1 \text{ in } [t, t + dt)) = dt \sum_{j=i}^N P(l_j^N), \quad (3.3.9b)$$

$$\mathbb{P}(\text{proliferation to the left of cell } i-1 \text{ in } [t, t + dt)) = dt \sum_{j=1}^{i-2} P(l_j^N), \quad (3.3.9c)$$

$$\mathbb{P}(\text{proliferation of cell } i-1 \text{ in } [t, t + dt)) = dt P(l_{i-1}^N). \quad (3.3.9d)$$

Taking a statistical expectation, denoted  $\langle \cdot \rangle$ , of Equation (3.3.7),  $\langle x_i^N(t) \rangle$  now represents the expected position of cell boundary  $i$  at time  $t$  in a system of  $N$  cells. We use the proliferation probabilities with the following simplifying assumptions: i)  $\langle x_i^N(t) \mathbb{1}\{\text{event}\} \rangle = \langle x_i^N(t) \rangle \langle \mathbb{1}\{\text{event}\} \rangle$ , namely independence of the position of the cell boundary in space and proliferation propensity, and a mean-field approximation as proliferation propensities depend on cell length; ii)  $\langle f_i(l_i^N(t)) \mathbb{1}\{\text{event}\} \rangle = \langle f_i(l_i^N(t)) \rangle \langle \mathbb{1}\{\text{event}\} \rangle$ , namely independence of the force and the propensity to proliferate, and a mean-field approximation as force depends on cell length; iii) a statistical mean-field approximation for force,  $\langle f_j^N(l_j^N) \rangle = f_j^N(\langle l_j^N \rangle)$ , and proliferation propensities,  $\langle P(l_j^N) \rangle = P(\langle l_j^N \rangle)$ . For simplicity we now drop the  $\langle \cdot \rangle$  notation. Then,

$$\begin{aligned} \frac{x_i^N(t + dt) - x_i^N(t)}{dt} &= \frac{1}{\eta} [f_i^N(l_i^N) - f_{i-1}^N(l_{i-1}^N)] \\ &- x_i^N(t) \sum_{j=1}^N P(l_j^N) + x_{i-1}^{N-1}(t) \sum_{j=i}^{N-1} P(l_j^{N-1}) \\ &+ x_{i-1}^{N-1}(t) \sum_{j=1}^{i-2} P(l_j^{N-1}) + \left( \frac{x_i^{N-1}(t) + x_{i-1}^{N-1}(t)}{2} \right) P(l_{i-1}^{N-1}) + \mathcal{O}(dt). \end{aligned} \quad (3.3.10)$$

We also assume: iv) the total propensity to proliferate is not significantly changed due to single a proliferation event,  $\sum_{j=1}^{N-1} P(l_j^{N-1}) dt = \sum_{j=1}^N P(l_j^N) dt + \mathcal{O}(dt^2, N^{-1})$ ; v) a single proliferation event does not significantly alter the position of a cell boundary,  $x_i^{N-1}(t) = x_i^N(t) + \mathcal{O}(dt)$  (Figure 3.1). As we will show, assumptions iv) and v) are good approximations for large  $N$  and allow us to combine summations. Then, assuming vi)  $\langle x_i^N(t) \rangle$  is a continuous function of time, we rearrange and take the limit  $dt \rightarrow 0$ . For the proliferation terms we replace the cell length with the discrete cell density  $q_i^N = 1/l_i^N$  to obtain

$$\begin{aligned} \frac{dx_i^N}{dt} &= \frac{1}{\eta} [f_i^N(l_i^N) - f_{i-1}^N(l_{i-1}^N)] \\ &\quad - \left( \frac{1}{q_{i-1}^N(t)} \right) \left[ \sum_{j=1}^{i-2} P\left(\frac{1}{q_j^N(t)}\right) + \frac{1}{2} P\left(\frac{1}{q_{i-1}^N(t)}\right) \right]. \end{aligned} \quad (3.3.11)$$

Equation (3.3.11) is only valid for the time interval  $[t, t + dt]$  under the assumptions iv) and v) above.

Thus far, we have extended the discrete model with mechanical relaxation to include the effects of cell proliferation. However, the statistically averaged model still retains information about discrete cell entities. We thus average over space to define a continuum cell density. Following Murphy et al. (2019) (Chapter 2, [157]), we introduce the microscopic density of cells,

$$\hat{q}(x, t) = \sum_{i=1}^N \delta(x - x_i^N(t)), \quad (3.3.12)$$

where  $\delta$  is the Dirac delta function (Evans and Morris 2008, Lighthill 1958). We define a local spatial average over a length scale  $\delta x$ , denoted  $\langle \cdot \rangle_{\delta x}$ , such that  $a_i \ll \delta x \ll L$ , which is sufficiently large to capture local heterogeneities for cellular properties that are constant during cell motion, including  $k$  and  $a$ , but sufficiently small to define continuous properties across  $L$ . The continuous cell density function,  $q(x, t)$ , is thus defined as

$$q(x, t) = \langle \hat{q}(x, t) \rangle_{\delta x} = \frac{1}{2\delta x} \int_{x-\delta x}^{x+\delta x} \hat{q}(y, t) dy. \quad (3.3.13)$$

Differentiating Equation (3.3.13) with respect to time gives

$$\frac{\partial q(x, t)}{\partial t} = -\frac{\partial}{\partial x} \left\langle \sum_{i=1}^N \delta(x - x_i^N(t)) \frac{dx_i^N}{dt} \right\rangle_{\delta x}, \quad (3.3.14)$$

where we use properties of the Dirac delta distribution (Evans and Morris 2008, Lighthill 1958,

Supplementary Material Equations S.15-S.17) and interchange the derivative with the spatial average as  $\delta x$  is small. Consistent with assumptions iv)-v) above, the sum over the microscopic densities can be considered to be fixed over  $N$  cells in Equation (3.3.14) within the small time interval  $[t, t + dt]$ .

On the right hand side of Equation (3.3.11), the first two terms involving  $f_i^N$  and  $f_{i-1}^N$  correspond to a mechanical contribution. This contribution is unchanged compared to Murphy et al. (2019) (Chapter 2, [157]) and, when substituted into Equation (3.3.14), it gives rise to the mechanical relaxation term on the right-hand side of continuum model, Equation (3.3.3) (Supplementary Material SM1.4). We now focus only on the contribution due to proliferation determined by substituting the proliferation terms of Equation (3.3.11) into Equation (3.3.14), giving a contribution which we denote  $\partial q(x, t)/\partial t|_P$ ,

$$\frac{\partial q(x, t)}{\partial t}\Big|_P = \frac{\partial}{\partial x} \left\langle \sum_{i=1}^N \delta(x - x_i^N(t)) \frac{1}{q_{i-1}^N(t)} \left[ \sum_{j=1}^{i-2} P\left(\frac{1}{q_j^N(t)}\right) + \frac{1}{2} P\left(\frac{1}{q_{i-1}^N(t)}\right) \right] \right\rangle_{\delta x}. \quad (3.3.15)$$

Now, assuming vii) that the spatial average interval is sufficiently far from the tissue boundary, i.e.  $i \gg 1$ , we make the following approximation:

$$\sum_{j=1}^{i-2} P\left(\frac{1}{q_j^N}\right) + \frac{1}{2} P\left(\frac{1}{q_{i-1}^N}\right) \approx \sum_{j=1}^{i-1} P\left(\frac{1}{q_j^N}\right). \quad (3.3.16)$$

To switch the dependence on the cell index to cell position, we multiply each term indexed by  $j$  in the sum on the right hand side of Equation (3.3.16) by  $1 = l_j q_j$ . Then, relating the discrete cell density to the continuous density through  $q_j^N = q(x_j^N(t), t)$ , gives

$$\sum_{j=1}^{i-1} q(x_j^N(t), t) P\left(\frac{1}{q(x_j^N(t), t)}\right) l_j. \quad (3.3.17)$$

We discretise the spatial domain  $x_1 \leq x \leq x_{i-1}$  with a uniform mesh with nodes  $y_s, s = 1, 2, \dots, S$ , where  $y_1 = x_1$ ,  $y_S = x_{i-1}$ , and  $y_s - y_{s-1} = \Delta y \ll l_j$ . Then, evaluating the continuous density at each node position,  $y_s$ , we interpret Equation (3.3.17) as the following Riemann sum

$$\sum_{s=1}^S q(y_s, t) P\left(\frac{1}{q(y_s, t)}\right) \Delta y = \int_0^{x_i^N} q(y, t) P\left(\frac{1}{q(y, t)}\right) dy, \quad (3.3.18)$$

where the integral on the right hand side is obtained by taking the limit  $\Delta y \rightarrow 0$ . Substituting

Equation (3.3.18) into Equation (3.3.15) gives

$$\frac{\partial q(x,t)}{\partial t} \Big|_P = \frac{\partial}{\partial x} \left\langle \sum_{i=1}^N \delta(x - x_i^N(t)) \left( \frac{1}{q_{i-1}^N} \right) \left[ \int_0^{x_i^N} q(y,t) P\left(\frac{1}{q(y,t)}\right) dy \right] \right\rangle_{\delta x}. \quad (3.3.19)$$

Calculating the spatial average, which only includes contributions from within the spatial average interval due to the Dirac delta functions, gives

$$\frac{\partial q(x,t)}{\partial t} \Big|_P = \frac{\partial}{\partial x} \left( \left( \frac{n}{2\delta x} \right) \frac{1}{n} \sum_{r=1}^n \left( \frac{1}{q_{r-1}^N} \right) \left[ \int_0^{x_r^N} q(y,t) P\left(\frac{1}{q(y,t)}\right) dy \right] \right), \quad (3.3.20)$$

where the index  $r$  labels the  $n$  cell boundaries contained within the spatial average interval  $(x - \delta x, x + \delta x)$ . Since  $a_i \ll \delta x \ll L$  and  $n \gg 1$  we have  $q_r^N = q(x_r^N(t), t) \approx q(x, t)$  for all  $r$ , which is now independent of  $r$ . Similarly,  $x_r^N \approx x$  for all  $r$ , where  $x$  is the centre of the spatial average interval. This gives

$$\frac{\partial q(x,t)}{\partial t} \Big|_P = \frac{\partial}{\partial x} \left( \left( \frac{n}{2\delta x} \right) \frac{1}{q(x,t)} \int_0^x q(y,t) P\left(\frac{1}{q(y,t)}\right) dy \right). \quad (3.3.21)$$

As  $n/(2\delta x) = q(x, t)$  in this spatial average interval, Equation (3.3.21) simplifies to

$$\frac{\partial q(x,t)}{\partial t} \Big|_P = q(x,t) P\left(\frac{1}{q(x,t)}\right). \quad (3.3.22)$$

At this point, we see that all explicit references to the total number of cells,  $N(t)$ , vanish. This allows the validity of the derivation, initially restricted to the time interval  $[t, t + dt]$ , to be extended to arbitrary times. As  $N(t) = \int_0^L q(x,t) dx$ , the change in the total cell number with time due to proliferation is accounted for through the source term written in Equation (3.3.22). We also stated assumption vii) that held true when sufficiently far from the tissue boundary but we find that this works at the boundary also (Sections 3.4.1, 3.4.2). Equation (3.3.22) shows proliferation arises as a single source term consistent with usual continuum-based formulations of proliferation whereas in Equation (85) of Baker et al. (2019) [19] proliferation arises as this term with an additional contribution.

## Death

The derivation of the cell death sink term follows similarly to that of the cell proliferation source term. We again consider an infinitesimally small time interval  $[t, t + dt]$ , so that at most one cell death event can occur in  $[t, t + dt]$ , and condition on cell death events to understand all possible events that occur and influence cell boundary  $i$  at  $t + dt$ . This gives

$$\begin{aligned} x_i^N(t + dt) = & \left[ x_i^N(t) + \frac{dt}{\eta} \{ f_i^N(l_i^N) - f_{i-1}^N(l_{i-1}^N) \} \right] \times \mathbb{1}\{\text{no death}\} \\ & + \left[ x_i^{N+1}(t) + \frac{dt}{\eta} \{ f_i^{N+1}(l_i^{N+1}) - f_{i-1}^{N+1}(l_{i-1}^{N+1}) \} \right] \\ & \times \mathbb{1}\{\text{death right of cell } i\} \\ & + \left[ x_{i+1}^{N+1}(t) + \frac{dt}{\eta} \{ f_{i+1}^{N+1}(l_{i+1}^{N+1}) - f_i^{N+1}(l_i^{N+1}) \} \right] \\ & \times \mathbb{1}\{\text{death left of cell } i\} \\ & + \left[ \frac{x_i^{N+1}(t) + x_{i+1}^{N+1}(t)}{2} + \frac{dt}{2\eta} \{ f_{i+1}^{N+1}(l_{i+1}^{N+1}) - f_{i-1}^{N+1}(l_{i-1}^{N+1}) \} \right] \\ & \times \mathbb{1}\{\text{death of cell } i\}. \end{aligned} \quad (3.3.23)$$

The cell death probabilities for Equation (3.3.23) for a system of  $N$  cells are given by

$$\mathbb{P}(\text{no death in } [t, t + dt]) = 1 - dt \sum_{j=1}^N D(l_j^N), \quad (3.3.24a)$$

$$\mathbb{P}(\text{death to the right of cell } i \text{ in } [t, t + dt]) = dt \sum_{j=i}^N D(l_j^N), \quad (3.3.24b)$$

$$\mathbb{P}(\text{death to the left of cell } i \text{ in } [t, t + dt]) = dt \sum_{j=1}^{i-2} D(l_j^N), \quad (3.3.24c)$$

$$\mathbb{P}(\text{death of cell } i \text{ in } [t, t + dt]) = dt D(l_i^N). \quad (3.3.24d)$$

Proceeding similarly to the proliferation derivation, we obtain

$$\begin{aligned} \frac{x_i^N(t + dt) - x_i^N(t)}{dt} = & \frac{1}{\eta} [f_i^N(l_i^N) - f_{i-1}^N(l_{i-1}^N)] \\ & - x_i^N(t) \sum_{j=1}^N D(l_j^N) + x_i^{N+1}(t) \sum_{j=i+1}^{N+1} D(l_j^{N+1}) \\ & + x_{i+1}^{N+1}(t) \sum_{j=1}^{i-1} D(l_j^{N+1}) + \left( \frac{x_i^{N+1}(t) + x_{i+1}^{N+1}(t)}{2} \right) D(l_i^{N+1}) + \mathcal{O}(dt). \end{aligned} \quad (3.3.25)$$

Then, following the same approach as the proliferation derivation, we arrive at the sink term in Equation (3.3.3) for cell death,  $-q(x, t)D(1/q(x, t))$ .

### Cell properties

Each cell is prescribed with intrinsic mechanical, proliferation, and death properties which are taken to be constant for each cell throughout the simulation. For mechanical cell properties, which include cell stiffness and resting cell length, we have the relationships  $\chi(x_i^N(t), t) = \chi_i$  for  $\chi = k, a$ . Similar relationships can be defined for the proliferation and death cell properties,  $\beta, \gamma, l_d$ . Differentiating these equations with respect to time we obtain Equations (3.3.6) (Chapter 2, [157]).

## 3.4 Numerical results

In this section, we first explore the evolution of a homogeneous tissue with different proliferation and death mechanisms and then explore mechanical cell competition for a heterogeneous tissue. To conclude we demonstrate the importance of the discrete to continuum approach through a series of problems where we compare the averaged discrete data with solutions of the corresponding continuum equations.

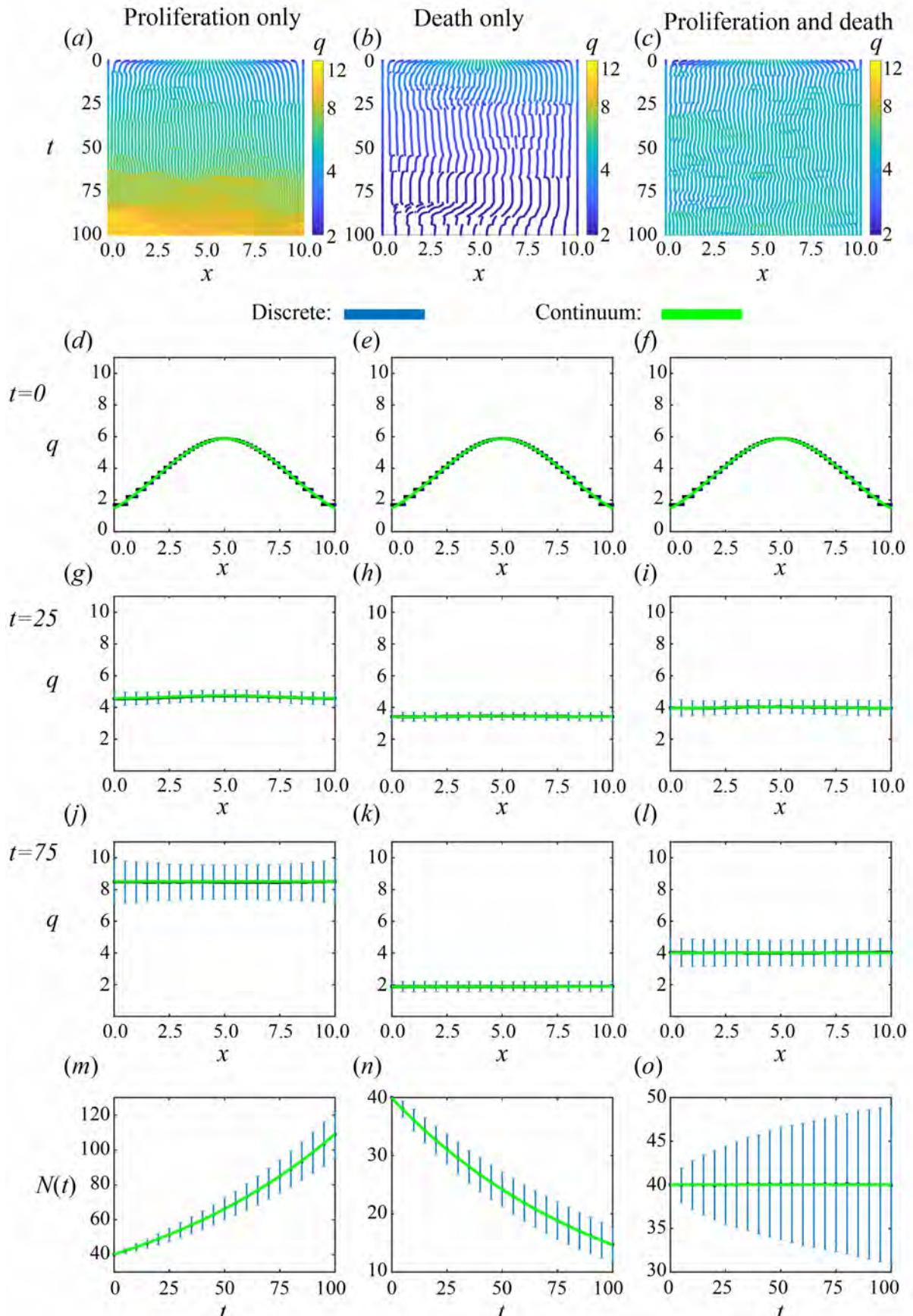
### 3.4.1 A homogeneous tissue

The simplest case to consider first is a homogeneous tissue composed of a population of identical cells. We explore three different proliferation and death mechanisms: constant, linear, and logistic (Table 3.1, Figure 3.2). For each mechanism we explore proliferation only, death only, and proliferation with death. We choose cell proliferation and death parameters (Figure 3.2) so that we can compare results fairly. We first choose the parameters for the proliferation mechanism so that, in the absence of cell death, a tissue with  $N(0) = 40$  evolves to have, on average,  $N(400) = 100$  (Figures 3.3(m), 3.4(m), S4(m)). Parameters for the death mechanisms were then subsequently chosen so that when proliferation and death are both included the total cell number stabilises at 40 cells for  $t > 0$ , which allows the standard deviations of  $N(t)$  to be compared fairly (Figures 3.3(n), 3.4(n), S4(n)).

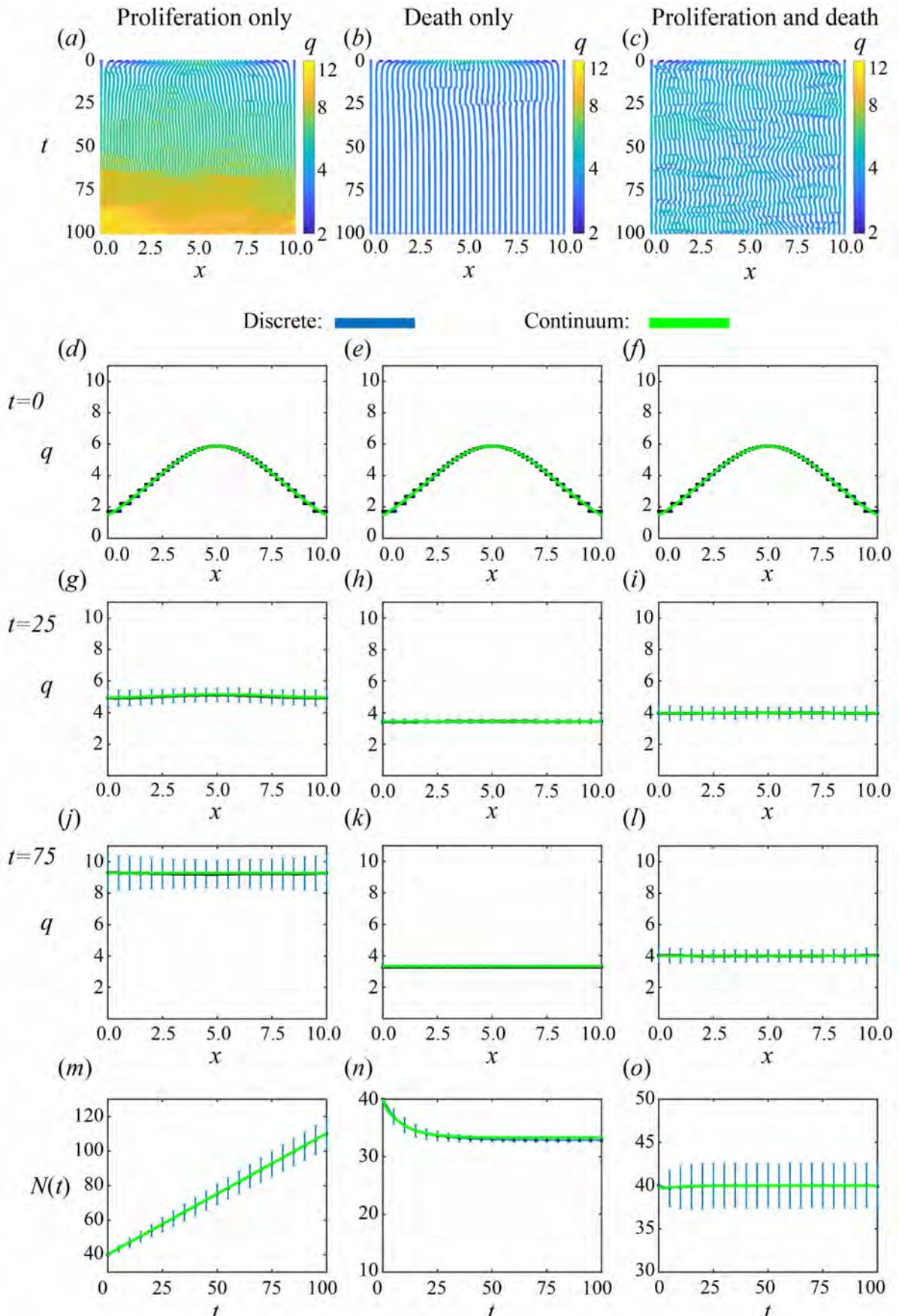
In all simulations we set  $L = 10$  and  $\eta = 1$ , use a Gaussian initial density centred at  $x = L/2$  with variance three and scaled to have  $N(0) = 40$ . We set  $k = 10$ , so that mechanical relaxation is fast in comparison to the proliferation and death [19]. For individual realisations this results in uniform densities except for short-time transient behaviour following a cell proliferation or death event (Figures 3.3a-c, 3.4a-c, S4a-c). Since epithelial cells in a tissue are in tension [228], we set  $a = 0$  for simplicity. Setting  $a > 0$  gives qualitatively similar results as long as cells remain in tension throughout the simulation.

For individual discrete realisations, cell proliferation causes a localised force imbalance followed by fast mechanical relaxation towards mechanical equilibrium and an overall increase in density (Figures 3.3a, 3.4a, S5a). Similarly, cell death results in a decrease in density followed by fast mechanical relaxation and an overall decrease in density (Figures 3.3b, 3.4b, S5b). With proliferation and death, cell boundaries are repeatedly introduced and removed, and the overall density remains, on average, constant (Figures 3.3c, 3.4c, S5c).

We observe excellent agreement when we compare the mean of many identically prepared discrete realisations and the corresponding solutions of the continuum model for both density



**Figure 3.3:** Homogeneous population with **constant** proliferation and death mechanisms. Proliferation only, death only, and proliferation with death shown in the left, middle and right columns, respectively. (a)-(c) Single realisations of cell boundary characteristics for  $0 \leq t \leq 100$ . (d)-(f), (g)-(i), (j)-(l) Density snapshots at times  $t = 0, 25, 75$ , respectively. (m)-(o) Total cell number. The average and standard deviation (blue error bars) of 2000 discrete simulations are compared to solution of continuum model (green).



**Figure 3.4:** Homogeneous population with **linear** proliferation and death mechanisms. Proliferation only, death only, and proliferation with death shown in the left, middle and right columns, respectively. (a)-(c) Single realisations of cell boundary characteristics for  $0 \leq t \leq 100$ . (d)-(f), (g)-(i), (j)-(l) Density snapshots at times  $t = 0, 25, 75$ , respectively. (m)-(o) Total cell number. The average and standard deviation (blue error bars) of 2000 discrete simulations are compared to solution of continuum model (green).

snapshots and total cell number (Figures 3.3d-o, 3.4d-o, S5d-o).

We note that the continuum model does not always provide a good match with an individual realization of the discrete model. For example, for constant proliferation and constant death with equal rates, every discrete realization will eventually become extinct (Supplementary Material SM3.1) as proliferation and death are independent of mechanical relaxation. This is expected as the total cell number is a linear birth-death process (Ross 1996) where the net proliferation rate is always equal to zero (Figure S4). As a consequence, the standard deviation of the total cell number increases with time (Figures 3.3o). When cell proliferation and death are cell-length-dependent there is closer agreement between the continuum model and single realisations. The net proliferation rate adjusts, due to changes in the number of cells and their lengths, to stabilise the population at its equilibrium value (Figure S4). Therefore extinction is extremely unlikely and the standard deviation of averaged discrete realisations is smaller (Figures 3.4o, S5o).

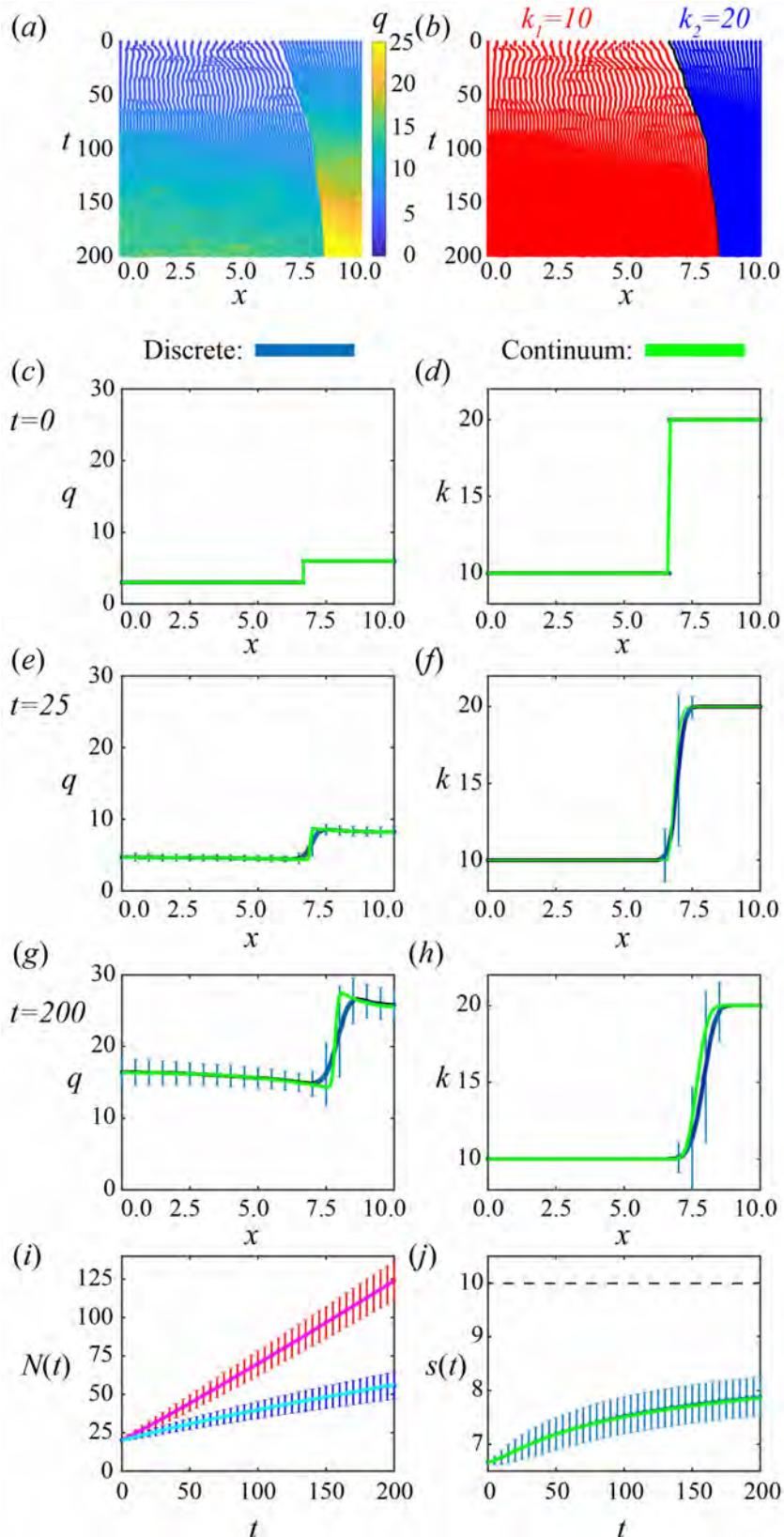
### 3.4.2 Mechanical cell competition

How tissues compete with each other for space is of great interest with many open biological questions being pursued in the experimental cell biology literature [29, 129, 232]. For example, in cancer invasion in an epithelial tissue a key question is whether cancer cells will eventually replace the entire healthy tissue or can the cancer cells coexist with the healthy cells? We consider this question by simulating a heterogeneous tissue composed of two populations, cancer cells adjacent to healthy cells (Figure 3.1c). Biologically, it is a hallmark of cancer cells that they are more proliferative and resistant to death than healthy cells [90]. In existing models the standard procedure would be to include these hallmarks as modelling assumptions and not consider the role of mechanical relaxation. However, we will now show this assumption is not necessary. We find that mechanical differences are sufficient for these hallmarks to arise and for cancer cells to outcompete healthy cells. We prescribe cancer and healthy cells the same proliferation and death mechanisms and parameters. We ask a further key question, how does mechanical relaxation alone compare to mechanical relaxation with proliferation, and to mechanical relaxation with proliferation and death?

In all scenarios, the left tissue (tissue 1) is coloured red to represent cancer cells and the right tissue (tissue 2) is coloured blue to represent healthy cells (Figure 3.1c). Each tissue starts with 20 cells. We assume cancer cells have lower stiffness than healthy cells (Lekka 2016) so we set cells in tissue 1 and 2 with cell stiffnesses  $K_1 = 10$  and  $K_2 = 20$ , respectively. Again, for simplicity and to represent that cells in an epithelial tissue are understood to be in tension [228], we set  $a = 0$ .

With only mechanical relaxation the interface position,  $s(t)$ , relaxes to the long-time interface position,  $\mathcal{S} = \lim_{t \rightarrow \infty} s(t) = 6.66$  (Chapter 2, [157]). In this scenario, the cancer and healthy cells coexist. However, the assumption of mechanical relaxation alone is only realistic over a short timescale where proliferation and death are negligible. When we include proliferation and death below, we use this long-time solution as the initial condition. As the mechanical relaxation rate is faster than the proliferation and death rates, using this initial condition only neglects initial short-time transient behaviour and does not significantly impact the long-time solution.

For mechanical relaxation with proliferation (Figure 3.5), we prescribe the linear proliferation mechanism for both the cancer and healthy cells with the same parameters. As cancer cells have lower cell stiffness than healthy cells, the cancer cells are always longer than the healthy cells (Supplementary Material SM4.1) except for the short transients after proliferation events where the cells have yet to mechanically relax. Initially, the cancer cells, with length 1/3,

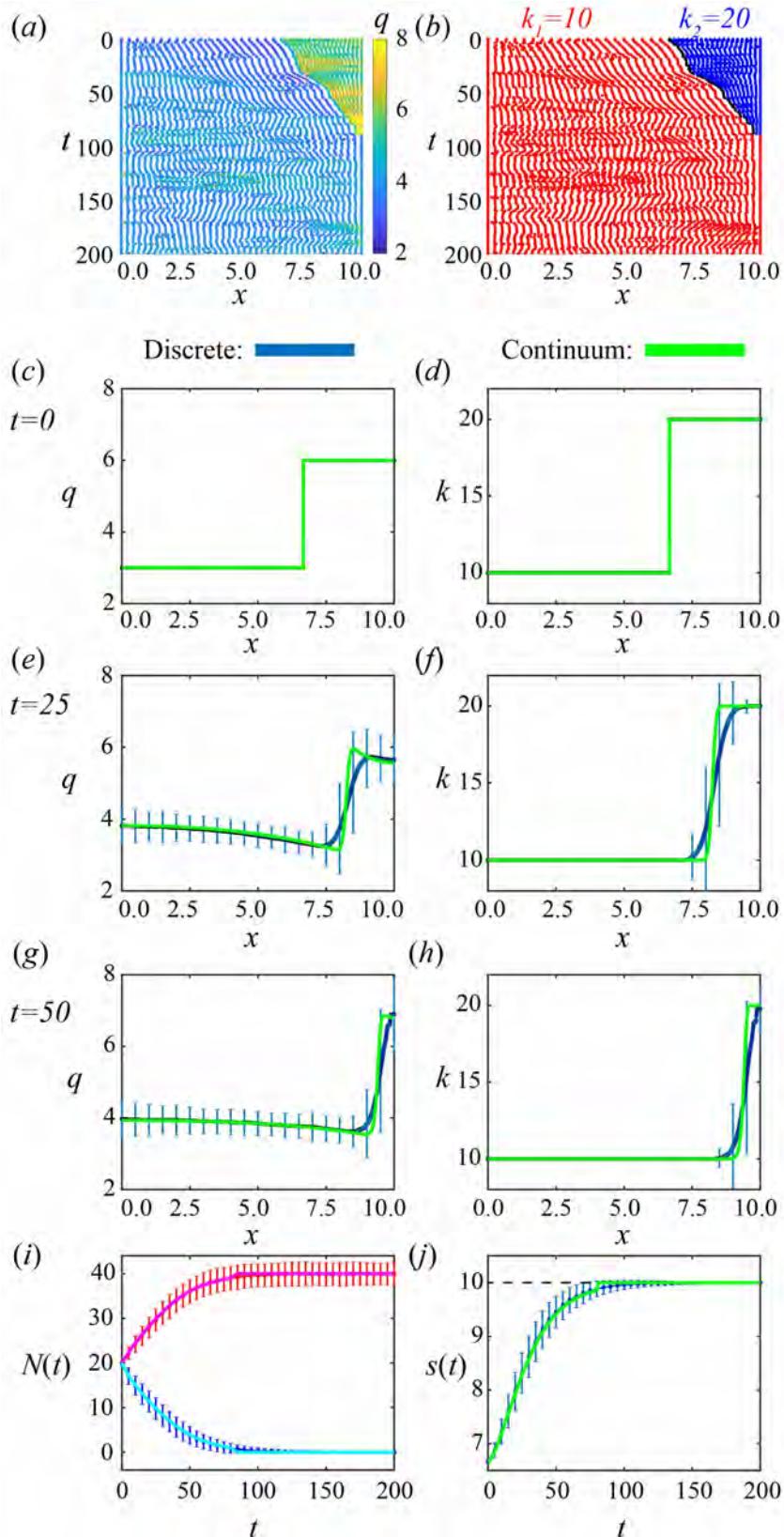


**Figure 3.5:** Results for cancer invasion with adjacent populations using **linear** proliferation and death mechanisms with **proliferation only**. (a),(b) A single realization of cell boundary characteristics for  $0 \leq t \leq 200$ . Colouring in (a),(b) represents cell density and cell stiffness, respectively. (c)-(d), (e)-(f), (g)-(h) Density and cell stiffness snapshots, left and right, respectively, at times  $t = 0, 25, 200$ , respectively. (i) Total cell number,  $N(t) > 0$ , for cancer (red/magenta) and healthy cells (blue/cyan) for the discrete/continuum solutions. (j) Interface position,  $s(t)$ , where the dotted line shows the edge of the domain. The average and standard deviation (blue error bar) of 2000 discrete simulations are compared to the solution of the continuum model (green).

are double the length of healthy cells. Referring to Figure 3.2b we see that the difference in cell lengths corresponds to cancer cells being more likely to proliferate than the healthy cells. Therefore, the cancer cells proliferate more than the healthy cells not because they were set to have advantageous intrinsic proliferation or death properties through a modelling assumption, but simply due to the coupling of mechanical relaxation with the length-dependent proliferation mechanism. With each proliferation event all cells become smaller, with the healthy cells remaining smaller than the cancer cells. Here we have coexistence but, as there is always a non-zero probability of proliferation and no cell death, all cells will eventually become unrealistically small and this happens first for healthy cells. In the absence of cell death, changing the proliferation mechanism will still result in coexistence.

For mechanical relaxation with proliferation and death (Figure 3A.13) a cell is more likely to die when it is smaller (Figure 3.2b). As we have observed for mechanical relaxation with proliferation, the healthy cells are smaller first, due to their higher relative stiffness, and therefore are more likely to die first. Once all of the healthy cells have died we have a homogeneous population of cancerous cells (Section 3.4.1). Importantly, we find that the cancerous cells, despite having identical proliferation and death mechanisms, are the *winner* cells of mechanical cell competition; they outcompete the healthy cells and take over the domain purely as a result of having lower cell stiffness. These results are robust to changes to the initial ratio of healthy cells to cancer cells (Figures S8, S9) and to the ratio of stiffness between healthy and cancer cells, provided cancer cells have lower stiffness than healthy cells (Figure S10).

Similar results regarding cancer invasion are found when considering the logistic mechanisms with both proliferation and death (Supplementary Material SM4.3). In contrast, for the constant proliferation and death mechanisms, where the proliferation and death mechanisms are both independent of the cell length and therefore independent of mechanical relaxation, to observe cancer cells invading the full domain we would have to prescribe the cancer cells to be more proliferative and resistant to death than the healthy cells.



**Figure 3.6:** Results for cancer invasion with adjacent populations using **linear** proliferation and death mechanisms with **proliferation and death**. First row shows a single realization of cell boundary characteristics for  $0 \leq t \leq 200$ . Colouring in (a),(b) represent cell density and cell stiffness, respectively. (c)-(d), (e)-(f), (g)-(h) Density and cell stiffness snapshots, left and right, respectively, at times  $t = 0, 25, 50$ , respectively. (i) Total cell number,  $N(t) > 0$ , for cancer (red/magenta) and healthy cells (blue/cyan) for the discrete/continuum solutions. (j) Interface position,  $s(t)$ , where the dotted line shows the edge of the domain. The average and standard deviation (blue error bar) of 2000 discrete simulations are compared to the solution of the continuum model (green).

### 3.4.3 Importance of the discrete to continuum approach

The discrete to continuum approach is important as it provides a principled means to determine how cell-level properties scale to the macroscale. Further, the approach provides conditions for whether or not the continuum model is beneficial, as we now explore.

In previous sections we choose proliferation and death mechanisms with parameters which lead to a good match between the appropriately averaged data from repeated discrete realisations and the solution of the corresponding continuum model. In Section 3.4.1, we demonstrate that individual realisations of the discrete model can go extinct while the continuum model does not. This provides a first indication that the continuum model does not always capture all relevant information from the underlying discrete model. We now demonstrate that if the approximations outlined in the derivation of the continuum model in Section 3.3.2 are not satisfied then the continuum approximation is not always satisfactory, and in such cases the discrete model should be used.

As an illustrative example we consider a proliferation mechanism which varies rapidly with cell length. For simplicity we consider the following piecewise cell-length-dependent proliferation mechanism

$$P(l_i) = \begin{cases} 0, & 0 \leq l_i < l_p, \\ 0.01, & l_i \geq l_p, \end{cases} \quad (3.4.1)$$

where we set the proliferation threshold to be  $l_p = 0.2$ . As before,  $N(0) = 40$  but we now choose a constant initial density condition so  $l_i = 0.25$  for each cell. Therefore, in the discrete model, each cell is initially able to proliferate. When the first cell proliferates it divides into two equally sized daughter cells with lengths  $l_i = 0.125$ . With fast mechanical relaxation, i.e. sufficiently large  $k$ , all 41 cells relax to equal size,  $l_i = 0.244$ , before the next proliferation event. This repeats until  $l_i < l_p$  for each cell  $i$  when proliferation stops (Figure 3.7d,f,h). This results in a tissue with 50 cells (Figure 3.7i,j), which is consistent with the continuum model where the density increases at the same rate everywhere in the tissue until reaching  $N = 50$  (Figure 3.7j, S6b,d,f). As the initial density condition is uniform the continuum solution holds true for any  $k$ . However, the behaviour of the discrete model for very slow mechanical relaxation, i.e. sufficiently small  $k$ , is very different. Proliferation occurs faster than mechanical relaxation so each of the initial 40 cells can proliferate, resulting in 80 cells (Figure 3.7c,e,g,i, S6a,c,e). It is clear that the continuum model does not accurately describe this problem and so we conclude that the discrete model should be used in this case. Increasing  $k$  results in an

improved agreement between the discrete and continuum models (Figure S7).

This example is important. The mismatch between the continuum and discrete results for this case remains even if we consider similar problems with larger numbers of cells, so simply increasing  $N(0)$  does not alleviate the issue. We do observe that increasing the mechanical relaxation rate, by increasing  $k$ , does provide a better match. However, in this piecewise proliferation mechanism example we require very high values of  $k$ , for example  $k = 1000$ , for a good match. Results in previous sections with excellent agreement are generated using  $k = 10$ . Revisiting the mechanical cell competition example and reducing to  $k = 0.0001$  still provides a reasonable match (Figure S13). This is because the rates involved in the proliferation and death mechanisms are smoother and slowly varying with respect to cell-length.

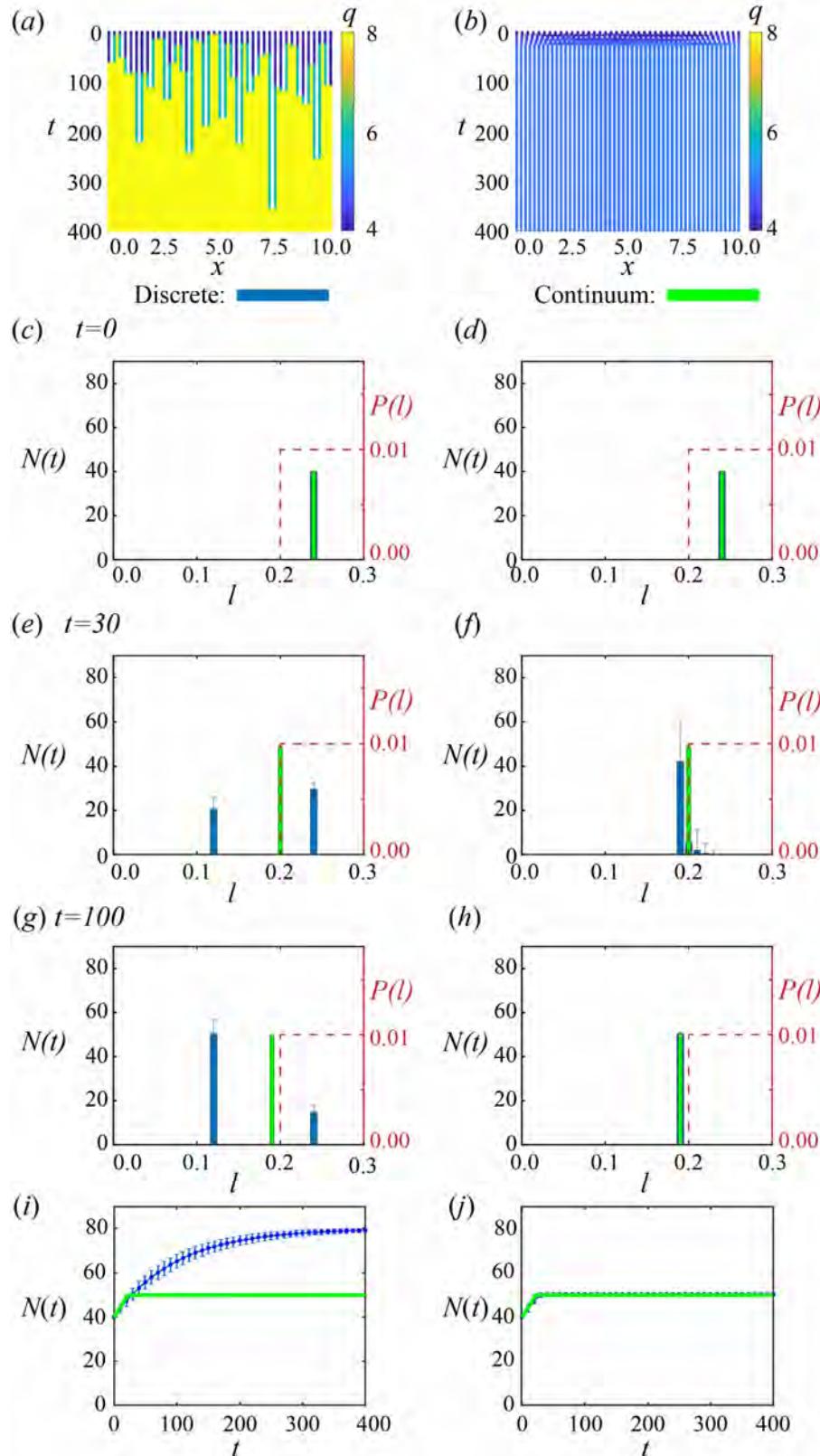
The results in Figure 3.7 may be surprising from the perspective of continuum mechanics. A common approach in continuum mechanics [14, 78, 154] is to start with conservation of mass and linear momentum and invoke constitutive laws. To derive our model using this approach one could start with the conservation of mass equation and then heuristically add source and sink terms to represent proliferation and cell death to give

$$\frac{\partial q(x, t)}{\partial t} + \frac{\partial}{\partial x} (q(x, t)u(x, t)) = q(x, t)P \left( \frac{1}{q(x, t)} \right) - q(x, t)D \left( \frac{1}{q(x, t)} \right). \quad (3.4.2)$$

The continuous analogue of the discrete conservation of momentum Equation (3.3.1) could be written by expanding the discrete cell-cell interaction force law with respect to cell-length in a Taylor series expansion to obtain

$$\eta u(x, t) = \frac{1}{q(x, t)} \frac{\partial f(x, t)}{\partial x}. \quad (3.4.3)$$

Equations (3.4.2) and (3.4.3) agree with Equations (3.3.3)-(3.3.6) derived earlier using a systematic coarse-graining approach. However, in the common continuum mechanics approach we would not have any opportunity to compare solutions of these continuum models with any underlying discrete description. This simple approach does not give any explicit indication of the underlying approximations inherent in the continuum model nor does it inform us when the continuum model may be a poor representation of the biology (Figure 3.7). Especially in biological contexts where cell numbers are large but local fluctuations can play an important role, we prefer to adopt the approach of starting with a biologically motivated discrete model and carefully derive the associated continuum limit, since this approach explicitly highlights the underlying assumptions inherent in the continuum model and provides us with a way of testing the accuracy of such assumptions.



**Figure 3.7:** Homogeneous population with rapidly varying proliferation mechanisms. With slow mechanical relaxation,  $k = 0.0001$ , and faster mechanical relaxation,  $k = 1000$ , shown in left and right columns, respectively. (a)-(b) Single realisations of cell boundary characteristics for  $0 \leq t \leq 100$ . (c)-(h) Cell length distributions against proliferation mechanism for times  $t = 0, 50, 100$  where one discrete realisation (blue) is compared against continuum model (green). (i)-(j) Total cell number where the average and standard deviation (blue error bars) of 2000 discrete simulations are compared to solution of continuum model (green).

### 3.5 Conclusion

In this work, we present a new one-dimensional cell-based model of heterogeneous epithelial tissue mechanics that includes cell proliferation and death. The main focus is to determine the corresponding continuum model which is a novel coupled system of nonlinear partial differential equations. The cell density equation is a parabolic partial differential equation while the cell property equations are hyperbolic partial differential equations. In deriving the continuum model, the discrete mechanisms and assumptions that underpin the continuum model have been made explicit by presenting the details of the coarse-graining derivation. Assumptions that relate to mean-field approximations and statistical independence of quantities are normally implicitly assumed in continuum models. By specifying the details of the derivation, and all assumptions required, our work provides insight into situations when these assumptions hold, as well as giving insight into when these assumptions fail to hold, such as when the number of cells,  $N(t)$ , is not sufficiently large, when cell properties vary rapidly in space, when mechanical relaxation is slow relative to rate of proliferation, or with proliferation and death mechanisms which vary rapidly with respect to cell-length. Under these conditions we recommend that the discrete description is more appropriate than the approximate continuum description. Further, we stress the limitations of developing continuum models by simply adding source and sink terms to an existing model without considering the underlying discrete model in complex biological systems.

By coupling mechanics with proliferation and death we are able to explore biological scenarios that could not be described in previous modelling frameworks. Specifically we can focus on mechanical cell competition driven by variations in cell stiffness and resting cell length. By choosing mechanical relaxation rates sufficiently fast relative to proliferation rates we observe good agreement between the average of many identically prepared stochastic realisations of the discrete model and the corresponding solutions of the continuum model. The quality of agreement holds even when our simulations only consider 40 cells and  $L = 10$  which is an extremely small number in comparison to the number of cells in an epithelial tissue. Further, assuming one cell is approximately 10  $\mu\text{m}$  in length [68, 73] then setting  $L = 10$  corresponds to a tissue length of approximately 0.4mm which is a biologically relevant length scale for applications such as wound healing. A continuum model is beneficial as we now have a tissue-level understanding of the mechanisms encoded in the discrete model and the time to solve the continuum model is independent of  $N(t)$ . The discrete model remains beneficial and can provide additional information. For example, the average of many discrete realisations can match the continuum model but every discrete realization could go extinct which is not observed in the

continuum model.

We explore mechanical cell competition applied to cancer invasion by considering cancer cells adjacent to healthy cells which compete for space. Interestingly, when we only allow cancer cells and healthy cells to differ in their cell stiffnesses, as a result of mechanical coupling, we observe that the cancer cells have more opportunities to proliferate and are less likely to die than healthy cells. We can then identify the cancer cells, as a result of the property of lower cell stiffness, as being the *winner* cells which invade the full domain. Cell stiffness and cell size may therefore be important factors to include when interpreting proliferation and death rates in experimental data. This analysis would not be possible using other existing models.

In all simulations we set  $a = 0$  to model cells being in tension [228]. Setting  $a > 0$  gives qualitatively similar results for homogeneous and heterogeneous populations as long as cells remain in tension throughout the simulation. This modelling framework is well-suited to be extended to cases where cells may also become compressed, for example in a tumour spheroid [53]. The model is well-suited to also study other observations of melanoma tumour spheroids such as subpopulations with differing proliferation rates located in different regions of the tumour, cells switching between these subpopulations, and the role of oxygen and nutrient concentrations [86, 239].

Many interesting extensions to this work are possible. Mathematically, the extent to which the continuum-limit holds with a free boundary is not yet clear (explored in Chapters 4 and 5). A free boundary also allows us to consider tissue growth [198] and shrinkage in mechanically less constrained environments, such as in developmental biology. Further, explicitly incorporating additional biological mechanisms that regulate cell size [101, 256, 257] and the evolution of intrinsic cell properties [89] would be both mathematically interesting and biologically relevant. In addition, while some features of cell ageing are implicit in the model, for example initially after a proliferation event daughter cells are less likely to proliferate than the parent cell with cell-length-dependent proliferation mechanisms, it would be of interest to explicitly incorporate the cell cycle and associated cell ageing processes such as growth in the resting cell length [147]. The theoretical foundations presented here for building a discrete model and constructing the continuum limit of that discrete model could be used to describe these additional mechanisms in future analyses (see Chapters 4 and 5 where we include free boundaries).

## **Chapter 3A**

# **Supplementary Material for Chapter 3**

<b>Contents</b>	<b>Page No.</b>
SM1 Model formulation	99
SM1.1 Discrete model with $m > 1$ springs per cell	99
SM1.2 Death at boundaries	102
SM1.3 Derivation of proliferation with $m > 1$ springs per cell	103
SM1.4 Mechanical relaxation	111
SM2 Numerical methods	112
SM2.1 Discrete model	112
SM2.2 Continuum model	114
SM3 Homogeneous population	119
SM3.1 Extinction for constant proliferation and death	119
SM3.2 Reduced variance with cell-length-dependent mechanisms	121
SM3.3 Logistic proliferation and death	123
SM3.4 Piecewise proliferation: varying mechanical relaxation rate	125
SM4 Mechanical cell competition	128
SM4.1 Two populations: cell size at mechanical equilibrium	128
SM4.2 Robustness to initial conditions	129
SM4.3 Logistic proliferation and death mechanisms	133
SM4.4 Varying mechanical relaxation rate	136

### 3A.1 Model formulation

In Section 2.1 of the main manuscript, we present the discrete model with one spring per cell and the derivation of the corresponding continuum model. Here, we present the discrete model with  $m > 1$  springs per cell, see Supplementary Material Section 3A.1.1. This is important to define the field functions in the continuum model, in particular the mechanical relaxation term, for sufficiently small  $N$  [157] where the size of a cell is no longer small in comparison to the size of the domain. Proliferation still occurs at a cellular level rather than at a spring level with  $m > 1$  therefore we still require many cells, i.e.  $a_i \ll \delta x \ll L$ . We also present: the special cases we need to consider at the boundaries for cell death, see Supplementary Section 3A.1.2; the derivation of continuum model for proliferation with  $m > 1$  springs per cell, see Supplementary Material Section 3A.1.3; and in Supplementary Material Section 3A.1.4 highlight key points for the derivation of the mechanical relaxation term [157].

#### 3A.1.1 Discrete model with $m > 1$ springs per cell

The model described in the main manuscript is presented by considering each cell to be represented by a single mechanical spring,  $m = 1$ , and tracking the evolution of cell boundaries. We now replace each cell with  $m > 1$  identical springs (Figure 3A.1a) [157]. We have spring boundaries at the cell boundaries as before but now we also have spring boundaries internal to the cell. We now track the evolution of all spring boundaries. In a system of  $N$  cells, cell  $i$  has spring boundaries  $x_{i,\nu}^N, \nu = 1, \dots, m$ , where  $x_i^N = x_{i,1}^N$ . The spring length is defined as  $l_{i,\nu}^N = x_{i,\nu}^N - x_{i,\nu-1}^N > 0$  and is related to cell length through  $l_i^N \sim ml_{i,\nu}^N$  as  $m \rightarrow \infty$ , and with equality for all  $m$  as  $t \rightarrow \infty$ . The mobility coefficient for a cell,  $\eta$ , and mechanical cell properties,  $k_i$  and  $a_i$ , are related to mobility coefficient for a spring boundary,  $\eta_\nu$ , spring stiffness,  $k_{i,\nu}^N$  and resting spring length,  $a_{i,\nu}^N$ , through the following scalings

$$\eta_\nu = \frac{\eta}{m}, \quad k_{i,\nu}^N = mk_i^N, \quad a_{i,\nu}^N = \frac{a_i^N}{m}, \quad (3A.1.1)$$

The spring boundaries,  $x_{i,\nu}^N$ , evolve according to

$$\begin{aligned} \eta_\nu \frac{dx_{i,\nu}^N(t)}{dt} &= f_{i,\nu}^N - f_{i,\nu-1}^N, \quad i = 2, \dots, N, \quad \nu = 1, 2, \dots, m, \\ f_{i,\nu}^N &= k_{i,\nu}^N (l_{i,\nu}^N - a_{i,\nu}^N). \end{aligned} \quad (3A.1.2)$$

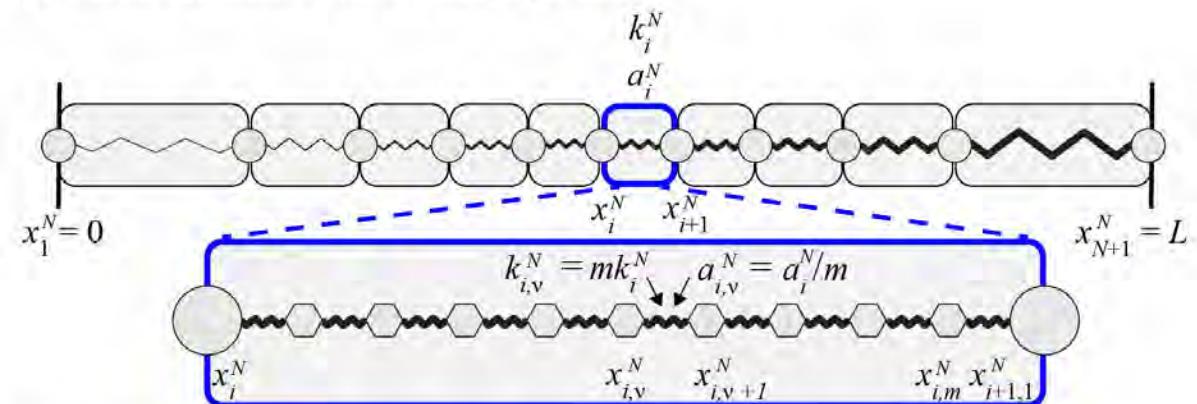
We consider proliferation to be a property of a cell rather than a property of springs within a cell. Specifically, when spring  $\nu$  in cell  $i$  is chosen to proliferate we consider that the whole

cell proliferates and introduce an additional  $m$  springs (Figure 3A.1b). Accordingly, we will introduce scaled spring proliferation rates. As with  $m = 1$ , we introduce the new cell boundary at the midpoint of original cell before proliferation. However, now we introduce  $m$  additional springs. To do so we equally space  $2m$  springs within the original cell. Similarly, for cell death we now instantly coalesce the cell boundaries and all internal spring boundaries to the centre of the dying cell. We have spring proliferation and death laws,  $P_\nu$  and  $D_\nu$ , respectively,

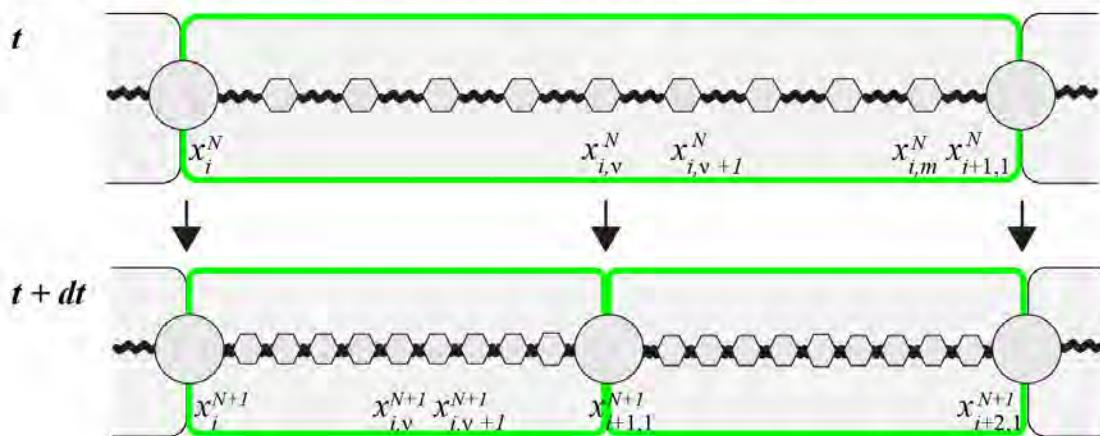
$$P_\nu(l_{i,\nu}^N) = \frac{P(l_i^N)}{m}, \quad D_\nu(l_{i,\nu}^N) = \frac{D(l_i^N)}{m}. \quad (3A.1.3)$$

The scalings are chosen such that the cell boundary velocities and proliferation/death rates are maintained and are independent of  $m$  [157].

(a) Model schematic with  $m$  springs per cell



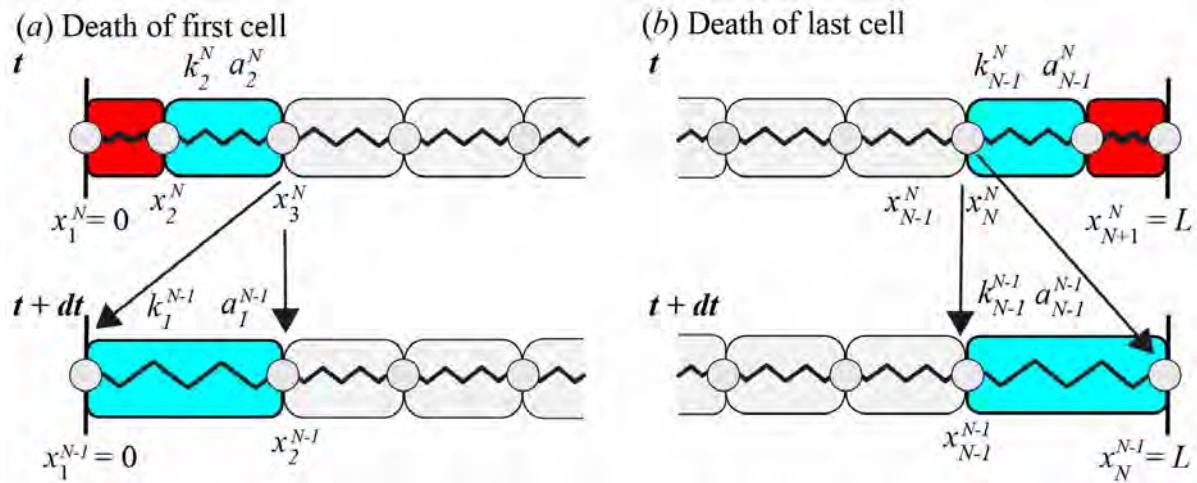
(b) Proliferation of cell  $i$  in model with  $m$  springs per cell



**Figure 3A.1:** (a) Schematic for the discrete model with  $m$  springs per cell. Heterogeneous population with  $N$  cells in a fixed domain of length  $L$ . Cell  $i$ , with a blue border, is prescribed with cell stiffness  $k_i^N$  and resting cell length  $a_i^N$ . The length of cell  $i$  is  $l_i^N = x_{i+1}^N - x_i^N$ . Spring  $\nu$  in cell  $i$  has spring boundaries  $x_{i,\nu}^N, x_{i,\nu+1}^N$ . Each spring is prescribed with a spring stiffness  $k_{i,\nu}^N = m k_i^N$  and a resting spring length  $a_{i,\nu}^N = a_i^N/m$ . Each spring has spring length  $l_{i,\nu}^N = x_{i,\nu+1}^N - x_{i,\nu}^N$ . The cell and spring boundaries are shown as discs and hexagons, respectively. (b) Proliferation of cell  $i$ , with a green border, in a model with  $m$  springs per cell. The original cell divides into two cells. An additional  $m$  equally spaced springs are introduced. This schematic is presented for even  $m$ .

### 3A.1.2 Death at boundaries

Cell death at a boundary is a special case of the discrete model which we present here. For death of the first cell whose left boundary is at  $x = 0$ , the first cell is removed and the left boundary of the second cell is set to  $x = 0$  (Figure 3A.2a). Similarly, for the death of the last cell whose right boundary is at  $x = L$ , the final cell is removed and the right boundary of cell  $N - 1$  is set to  $x = L$  (Figure 3A.2b).



**Figure 3A.2:** Model schematic for special cases of cell death of (a) first and (b) last cell.

### 3A.1.3 Derivation of proliferation with $m > 1$ springs per cell

In Section 2.2 of the main manuscript, we present the derivation with one spring per cell,  $m = 1$ . Here, for completeness, we rewrite the derivation with  $m$  springs per cell. The advantage of  $m$  springs per cell is that we can more appropriately define the continuous field functions, in particular the mechanical relaxation term for low cell numbers. New text relevant for  $m$  springs per cell is shown in purple. Setting  $m = 1$  in this new derivation recovers the results in the main manuscript.

Starting from the discrete model described in Section 2.1 but now for  $m$  springs per cell as described in Supplementary Material Section 1.1, we now derive the proliferation source term in Equation (3). As cell proliferation is included stochastically, we consider an infinitesimal time interval  $[t, t + dt]$  and condition on the possible proliferation events that could occur and influence the position of spring boundary  $\nu$  in cell  $i$  in a system of  $N$  cells. We note that proliferation is still considered a cell event rather than a spring event. Specifically, we say that a cell proliferates once any spring in the cell has been chosen to proliferate; this is captured through the scalings in Equation (3A.1.3).

Choosing  $dt$  sufficiently small so that at most one proliferation event can occur in  $[t, t + dt]$  there are five possibilities: i) either there is no proliferation, in which case the spring boundary position  $x_{i,\nu}^N$  only changes by mechanical relaxation; ii) there is proliferation to the right of cell  $i$ ; iii) there is proliferation to the left of cell  $i - 1$ ; iv) cell  $i - 1$  proliferates; and v) cell  $i$  proliferates. This leads to the following infinitesimal evolution law for the position of spring boundary  $\nu$  in

cell  $i$   $x_{i,\nu}^N$ , accounting for cell relabelling when a new cell is added:

$$\begin{aligned}
 x_{i,\nu}^N(t+dt) = & \left[ x_{i,\nu}^N(t) + \frac{dt}{\eta} \left\{ f_{i,\nu}^N(l_{i,\nu}^N) - f_{i,\nu-1}^N(l_{i,\nu-1}^N) \right\} \right] \times \mathbb{1}\{\text{no proliferation}\} \\
 & + \left[ x_{i,\nu}^{N-1}(t) + \frac{dt}{\eta} \left\{ f_{i,\nu}^{N-1}(l_{i,\nu}^{N-1}) - f_{i,\nu-1}^{N-1}(l_{i,\nu-1}^{N-1}) \right\} \right] \\
 & \quad \times \mathbb{1}\{\text{proliferation right of cell } i\} \\
 & + \left[ x_{i-1,\nu}^{N-1}(t) + \frac{dt}{\eta} \left\{ f_{i-1,\nu}^{N-1}(l_{i-1,\nu}^{N-1}) - f_{i-1,\nu-1}^{N-1}(l_{i-1,\nu-1}^{N-1}) \right\} \right] \\
 & \quad \times \mathbb{1}\{\text{proliferation left of cell } i-1\} \\
 & + \left[ x_{i-1,(\frac{m}{2}+\frac{\nu+1}{2})}^{N-1}(t) + \frac{dt}{\eta} \left\{ f_{i-1,(\frac{m}{2}+\frac{\nu+1}{2})}^{N-1}(l_{i-1,(\frac{m}{2}+\frac{\nu+1}{2})}^{N-1}) - f_{i-1,(\frac{m}{2}+\frac{\nu-1}{2})}^N(l_{i-1,(\frac{m}{2}+\frac{\nu-1}{2})}^N) \right\} \right] \\
 & \quad \times \mathbb{1}\{\text{proliferation of cell } i-1\} \\
 & + \left[ x_{i,(\frac{\nu+1}{2})}^{N-1}(t) + \frac{dt}{\eta} \left\{ f_{i+1,(\frac{\nu+1}{2})}^{N-1}(l_{i+1,(\frac{\nu+1}{2})}^{N-1}) - f_{i-1,(\frac{\nu+1}{2})}^N(l_{i-1,(\frac{\nu+1}{2})}^N) \right\} \right] \\
 & \quad \times \mathbb{1}\{\text{proliferation of cell } i\}.
 \end{aligned} \tag{3A.1.4}$$

The derivation is shown for even  $m$  and odd  $\nu$  for simplicity (Figure 3A.1b). With other choices of  $m$  and  $\nu$  we obtain slightly different terms for the proliferation of cell  $i-1$  and  $i$ . However, these choices are not important to the following derivation. Each term in square brackets is the resulting force from neighbouring cells due to mechanical relaxation, given by Equations (3A.1.2), for each potential event. In addition, we include Boolean variables expressed as indicator functions,  $\mathbb{1}\{\cdot\}$ , defined as

$$\mathbb{1}\{\text{event}\} = \begin{cases} 1, & \text{if event occurs in } [t, t+dt], \\ 0, & \text{otherwise,} \end{cases} \tag{3A.1.5}$$

whose expectations in the context of Equation (3A.1.4) can be interpreted as proliferation probabilities. For a system of  $N$  cells with  $m$  springs per cell, where  $dt$  is sufficiently small, these

proliferation probabilities are given by

$$\mathbb{P}(\text{no proliferation in } [t, t + dt]) = 1 - dt \sum_{j=1}^N \sum_{\nu=1}^m P_\nu(l_{j,\nu}^N), \quad (3A.1.6a)$$

$$\mathbb{P}(\text{proliferation to the right of cell } i \text{ in } [t, t + dt]) = dt \sum_{j=i+1}^N \sum_{\nu=1}^m P_\nu(l_{j,\nu}^N), \quad (3A.1.6b)$$

$$\mathbb{P}(\text{proliferation to the left of cell } i-1 \text{ in } [t, t + dt]) = dt \sum_{j=1}^{i-2} \sum_{\nu=1}^m P_\nu(l_{j,\nu}^N), \quad (3A.1.6c)$$

$$\mathbb{P}(\text{proliferation of cell } i-1 \text{ in } [t, t + dt]) = dt \sum_{\nu=1}^m P_\nu(l_{i-1,\nu}^N), \quad (3A.1.6d)$$

$$\mathbb{P}(\text{proliferation of cell } i \text{ in } [t, t + dt]) = dt \sum_{\nu=1}^m P_\nu(l_{i,\nu}^N), \quad (3A.1.6e)$$

where using the proliferation rate scaling, from Equation (3A.1.3), Equations (3A.1.6) can be written in terms of the cell proliferation rates,

$$\mathbb{P}(\text{no proliferation in } [t, t + dt]) = 1 - dt \sum_{j=1}^N P(l_j^N), \quad (3A.1.7a)$$

$$\mathbb{P}(\text{proliferation to the right of cell } i \text{ in } [t, t + dt]) = dt \sum_{j=i+1}^N P(l_j^N), \quad (3A.1.7b)$$

$$\mathbb{P}(\text{proliferation to the left of cell } i-1 \text{ in } [t, t + dt]) = dt \sum_{j=1}^{i-2} P(l_j^N), \quad (3A.1.7c)$$

$$\mathbb{P}(\text{proliferation of cell } i-1 \text{ in } [t, t + dt]) = dt P(l_{i-1}^N), \quad (3A.1.7d)$$

$$\mathbb{P}(\text{proliferation of cell } i \text{ in } [t, t + dt]) = dt P(l_i^N). \quad (3A.1.7e)$$

Taking a statistical expectation, denoted  $\langle \cdot \rangle$ , of Equation (3A.1.4),  $\langle x_{i,\nu}^N(t) \rangle$  now represents the expected position of **spring boundary**  $\nu$  in **cell**  $i$  at time  $t$  in a system of  $N$  cells. We use the proliferation probabilities with the following simplifying assumptions: i)  $\langle x_{i,\nu}^N(t) \mathbb{1}\{\text{event}\} \rangle = \langle x_{i,\nu}^N(t) \rangle \langle \mathbb{1}\{\text{event}\} \rangle$ , namely independence of the position of the cell boundary in space and proliferation propensity, and a mean-field approximation as proliferation propensities depend on **spring length**; ii)  $\langle x_{i,\nu}^N(t) f^N(l_{j,\nu}^N) \rangle = \langle x_{i,\nu}^N(t) \rangle \langle f^N(l_{j,\nu}^N) \rangle$ , namely independence of force and the propensity to proliferate, and a mean-field approximation as force depends on **spring length**; iii) a statistical mean-field approximation for force,  $\langle f^N(l_{j,\nu}^N) \rangle = f^N(\langle l_{j,\nu}^N \rangle)$ , and proliferation terms,  $\langle P(l_j^N) \rangle = P(\langle l_j^N \rangle)$ . For simplicity, we now drop the  $\langle \cdot \rangle$  notation.

Then,

$$\begin{aligned} \frac{x_{i,\nu}^N(t + dt) - x_{i,\nu}^N(t)}{dt} &= \frac{1}{\eta} [f_{i,\nu}^N(l_{i,\nu}^N) - f_{i,\nu-1}^N(l_{i,\nu-1}^N)] \\ &- x_{i,\nu}^N(t) \sum_{j=1}^N P(l_j^N) + x_{i,\nu}^{N-1}(t) \sum_{j=i+1}^{N-1} P(l_j^{N-1}) + x_{i,(\frac{\nu+1}{2})}^{N-1}(t) P(l_i^{N-1}) \\ &+ x_{i-1,\nu}^{N-1}(t) \sum_{j=1}^{i-2} P(l_j^{N-1}) + x_{i-1,(\frac{m}{2}+\frac{\nu+1}{2})}^{N-1}(t) P(l_{i-1}^{N-1}) + \mathcal{O}(dt). \end{aligned} \quad (3A.1.8)$$

We also have

$$x_{i,(\frac{\nu+1}{2})}^{N-1}(t) = x_{i,\nu}^{N-1}(t) + \left[ x_{i,(\frac{\nu+1}{2})}^{N-1}(t) - x_{i,\nu}^{N-1}(t) \right], \quad (3A.1.9)$$

and we make assumption vii) that we are sufficiently far from the tissue boundary such that the term in square brackets of Equation (3A.1.9) is negligible in comparison to the first term in the right hand side. This is consistent with Equation (16) in the derivation with  $m = 1$ . Similarly for the term  $x_{i-1,(\frac{m}{2}+\frac{\nu+1}{2})}^{N-1}(t)$ . This gives

$$\begin{aligned} \frac{x_{i,\nu}^N(t + dt) - x_{i,\nu}^N(t)}{dt} &= \frac{1}{\eta} [f_{i,\nu}^N(l_{i,\nu}^N) - f_{i,\nu-1}^N(l_{i,\nu-1}^N)] \\ &- x_{i,\nu}^N(t) \sum_{j=1}^N P(l_j^N) + x_{i,\nu}^{N-1}(t) \sum_{j=i+1}^{N-1} P(l_j^{N-1}) + x_{i,\nu}^{N-1}(t) P(l_i^{N-1}) \\ &+ x_{i-1,\nu}^{N-1}(t) \sum_{j=1}^{i-2} P(l_j^{N-1}) + x_{i-1,\nu}^{N-1}(t) P(l_{i-1}^{N-1}) + \mathcal{O}(dt) \end{aligned} \quad (3A.1.10)$$

We also assume: iv) the total propensity to proliferate is not significantly changed due to a single proliferation event,  $\sum_{j=1}^{N-1} P(l_j^{N-1}) dt = \sum_{j=1}^N P(l_j^N) dt + \mathcal{O}(dt^2, \frac{1}{N})$ ; v) a single proliferation event does not significantly alter the position of a cell boundary,  $x_{i,\nu}^{N-1}(t) = x_{i,\nu}^N(t) + \mathcal{O}(dt)$ . As we will show, assumptions iv) and v) are good approximations for large  $N$  and allow us to combine summations to give

$$\begin{aligned} \frac{x_{i,\nu}^N(t + dt) - x_{i,\nu}^N(t)}{dt} &= \frac{1}{\eta} [f_{i,\nu}^N(l_{i,\nu}^N) - f_{i,\nu-1}^N(l_{i,\nu-1}^N)] \\ &- x_{i,\nu}^N(t) \left[ \sum_{j=1}^{i-1} P(l_j^N) \right] + x_{i-1,\nu}^N(t) \left[ \sum_{j=1}^{i-1} P(l_j^N) \right] + \mathcal{O}(dt, N^{-1}). \end{aligned} \quad (3A.1.11)$$

Then, assuming vi)  $\langle x_{i,\nu}^N(t) \rangle$  is a continuous function of time, we rearrange and take the limit  $dt \rightarrow 0$ . For the proliferation terms we replace the cell length with the discrete cell density

$q_i^N = 1/l_i^N$  to obtain

$$\frac{dx_{i,\nu}^N}{dt} = f_{i,\nu}^N(l_{i,\nu}^N) - f_{i,\nu-1}^N(l_{i,\nu-1}^N) - \left( \frac{1}{q_i^N(t)} \right) \left[ \sum_{j=1}^{i-1} P \left( \frac{1}{q_j^N(t)} \right) \right]. \quad (3A.1.12)$$

Equation (3A.1.12) is only valid for the time interval  $[t, t + dt]$  under the assumptions iv) and v) above.

Thus far, we have extended the discrete model with mechanical relaxation to include the effects of cell proliferation and death. However, the statistically averaged model still retains information about discrete cell entities. We thus average over space to define a continuum cell density. Following Murphy et al. (2019) [157], we introduce the microscopic density of cells,

$$\hat{q}(x, t) = \frac{1}{m} \sum_{i=1}^N \sum_{\nu=1}^m \delta(x - x_{i,\nu}^N(t)), \quad (3A.1.13)$$

where  $\delta$  is the Dirac delta function (Evans and Morrise 2008, Lighthill 1958). We define a local spatial average over a length scale  $\delta_x$ , denoted  $\langle \cdot \rangle_{\delta_x}$ , such that  $a_{i,\nu} \ll a_i \ll \delta_x \ll L$ , which is sufficiently large to capture local heterogeneities for cellular properties that are constant during cell motion, including  $k$  and  $a$ , but sufficiently small to define continuous properties across  $L$ . The continuous cell density function,  $q(x, t)$ , is thus defined as

$$q(x, t) = \langle \hat{q}(x, t) \rangle_{\delta_x} = \frac{1}{2\delta_x} \int_{x-\delta_x}^{x+\delta_x} \hat{q}(y, t) dy. \quad (3A.1.14)$$

We proceed by differentiating Equation (3A.1.14) with respect to time to obtain

$$\frac{\partial q(x, t)}{\partial t} = \frac{\partial}{\partial t} \langle \hat{q}(x, t) \rangle_{\delta_x} = \left\langle \frac{\partial \hat{q}(x, t)}{\partial t} \right\rangle_{\delta_x} \quad (3A.1.15)$$

From Evans and Morrise (2008), the term inside the spatial average on the right hand side of

Equation (3A.1.15) can be written as

$$\begin{aligned}
 \frac{\partial \hat{q}(x, t)}{\partial t} &= \frac{\partial}{\partial t} \left( \frac{1}{m} \sum_{i=1}^N \sum_{\nu=1}^m \delta(x - x_{i,\nu}^N(t)) \right), \\
 &= \frac{1}{m} \sum_{i=1}^N \sum_{\nu=1}^m \frac{dx_{i,\nu}}{dt} \frac{\partial}{\partial x_{i,\nu}} \delta(x - x_{i,\nu}^N(t)), \\
 &= -\frac{1}{m} \sum_{i=1}^N \sum_{\nu=1}^m \frac{dx_{i,\nu}}{dt} \frac{\partial}{\partial x} \delta(x - x_{i,\nu}^N(t)), \\
 &= -\frac{\partial}{\partial x} \left( \frac{1}{m} \sum_{i=1}^N \sum_{\nu=1}^m \frac{dx_{i,\nu}}{dt} \delta(x - x_{i,\nu}^N(t)) \right),
 \end{aligned} \tag{3A.1.16}$$

where we obtain the third line of Equation (3A.1.16) by making use of properties associated with the Direct delta function (Evans and Morris 2008, Lighthill 1958). Then substituting Equation (3A.1.16) into Equation (3A.1.15) gives

$$\frac{\partial q(x, t)}{\partial t} = -\frac{\partial}{\partial x} \left\langle \frac{1}{m} \sum_{i=1}^N \sum_{\nu=1}^m \frac{dx_{i,\nu}}{dt} \delta(x - x_{i,\nu}^N(t)) \right\rangle_{\delta x}, \tag{3A.1.17}$$

where on the right hand side, as  $\delta x$  is small, we have interchanged the spatial average and the derivative with respect to  $x$ . We note that the final step using the assumption  $\delta x$  is small can be shown more formally by explicitly treating the Dirac delta function as a generalised function and using test functions (Lighthill 1958). Consistent with assumptions 4)-5) above, the sum over the microscopic densities can be considered to be fixed over  $N$  cells in Equation (3A.1.17) within the small time interval  $[t, t + dt]$ .

On the right hand side of Equation (3A.1.12), the first two terms involving  $f$  correspond to a mechanical contribution. This contribution is unchanged compared to Murphy et al. (2019) [157] and, when substituted into Equation (3A.1.17), it gives rise to the mechanical relaxation term on the right hand side of the continuum equation (3) (Supplementary Material 3A.1.4). We now focus only on the contribution determined by substituting the proliferation terms of Equation (3A.1.12) into Equation (3A.1.17), giving a contribution which we denote  $\partial q(x, t)/\partial t|_P$ ,

$$\frac{\partial q(x, t)}{\partial t} \Big|_P = \frac{\partial}{\partial x} \left\langle \frac{1}{m} \sum_{i=1}^N \sum_{\nu=1}^m \delta(x - x_{i,\nu}^N(t)) \frac{1}{q_{i-1}^N(t)} \left[ \sum_{j=1}^{i-1} P \left( \frac{1}{q_j^N(t)} \right) \right] \right\rangle_{\delta x}. \tag{3A.1.18}$$

At this point in the derivation for  $m = 1$  we make the assumption that we are sufficiently far from the tissue boundary. We name this as assumption vii). For  $m > 1$  we have already made assumption vii) and this is not required again here. To switch the dependence on the cell index

to cell position, we focus on the sum in square brackets on the right hand side of Equation (3A.1.18). We multiply each term  $j$  by  $1 = l_j q_j$ . Then relating the discrete cell density to the continuous density through  $q_j^N = q(x_j^N(t), t)$  gives

$$\sum_{j=1}^{i-1} q(x_j^N(t), t) P \left( \frac{1}{q(x_j^N(t), t)} \right) l_j. \quad (3A.1.19)$$

We discretise the spatial domain  $x_1 \leq x \leq x_{i-1}$  with a uniform mesh with nodes  $y_s, s = 1, 2, \dots, S$ , where  $y_1 = x_1, y_S = x_{i-1}$ , and  $y_s - y_{s-1} = \Delta y \ll l_j$ . Then, evaluating the continuous density at each node position,  $y_s$ , we interpret Equation (3A.1.19) as the following Riemann sum

$$\sum_{s=1}^S q(y_s, t) P \left( \frac{1}{q(y_s, t)} \right) \Delta y = \int_0^{x_i^N} q(y, t) P \left( \frac{1}{q(y, t)} \right) dy. \quad (3A.1.20)$$

where the integral on the right hand side is obtained by taking the limit  $\Delta y \rightarrow 0$ . Substituting Equation (3A.1.20) into Equation (3A.1.18) gives

$$\frac{\partial q(x, t)}{\partial t} \Big|_P = \frac{\partial}{\partial x} \left\langle \frac{1}{m} \sum_{i=1}^N \sum_{\nu=1}^m \delta(x - x_{i,\nu}(t)) \left( \frac{1}{q_{i-1}^N} \right) \left[ \int_0^{x_i^N} q(y, t) P \left( \frac{1}{q(y, t)} \right) dy \right] \right\rangle_{\delta x}. \quad (3A.1.21)$$

Calculating the spatial average, which only includes contributions from within the spatial average interval due to the Dirac delta functions, gives

$$\frac{\partial q(x, t)}{\partial t} \Big|_P = \frac{\partial}{\partial x} \left( \left( \frac{n}{2\delta x} \right) \frac{1}{n} \sum_{r=1}^n \sum_{\nu=1}^m \left( \frac{1}{q_{r-1}^N} \right) \left[ \int_0^{x_r^N} q(y, t) P \left( \frac{1}{q(y, t)} \right) dy \right] \right), \quad (3A.1.22)$$

where the index  $r$  labels the  $n$  cell boundaries contained within the spatial average interval  $(x - \delta x, x + \delta x)$ . Equation (3A.1.22) is now independent of  $m$ , which is to be expected as proliferation is considered a cell event rather than a spring event. Simplifying gives

$$\frac{\partial q(x, t)}{\partial t} \Big|_P = \frac{\partial}{\partial x} \left( \left( \frac{n}{2\delta x} \right) \frac{1}{n} \sum_{r=1}^n \left( \frac{1}{q_{r-1}^N} \right) \left[ \int_0^{x_r^N} q(y, t) P \left( \frac{1}{q(y, t)} \right) dy \right] \right). \quad (3A.1.23)$$

Since  $a_i \ll \delta x \ll L$  and  $n \gg 1$  we have  $q_r^N = q(x_r^N(t), t) = q(x, t)$  for all  $r$ , which is independent of  $r$ . Similarly,  $x_r \approx x$  for all  $r$ , where  $x$  is the centre of the spatial average interval. This gives

$$\frac{\partial q(x, t)}{\partial t} \Big|_P = \frac{\partial}{\partial x} \left( \left( \frac{n}{2\delta x} \right) \frac{1}{q(x, t)} \int_0^x q(y, t) P \left( \frac{1}{q(y, t)} \right) dy \right). \quad (3A.1.24)$$

As  $n/(2\delta x) = q(x, t)$  in this spatial average interval, Equation (3A.1.24) simplifies to

$$\frac{\partial q(x, t)}{\partial t} \Big|_P = q(x, t) P \left( \frac{1}{q(x, t)} \right). \quad (3A.1.25)$$

At this point, we see that all explicit references to the total number of cells,  $N(t)$ , vanish. This allows the validity of the derivation, initially restricted to the time interval  $[t, t + dt]$ , to be extended to arbitrary times. As  $N(t) = \int_0^L q(x, t) dx$ , the change in the total cell number with time due to proliferation is accounted for through the source term written in Equation (3A.1.25). We also stated assumption vii) that held true when sufficiently far from the tissue boundary but we find that this works at the boundary also (Section 2).

### 3A.1.4 Mechanical relaxation

We outline the key steps to derive the mechanical terms, with one spring per cell, and refer the reader to Section 2 of Murphy et al. (2019) [157] for full details. We introduce field functions for the force,  $f(x, t)$ , cell stiffness,  $k(x, t)$ , and resting cell length  $a(x, t)$ , which relate to the discrete model through

$$\begin{aligned} f(x_i^N(t), t) &= f_i^N, \\ k(x_i^N(t), t) &= k_i^N, \\ a(x_i^N(t), t) &= a_i^N. \end{aligned} \tag{3A.1.26}$$

Substituting Equation (11) into Equation (14), we focus on the force term on the right side and consider no proliferation or death. We expand the cell-cell interaction force using the small cell length parameter  $l_i^N$ , small as the number of cell boundaries inside the spatial average interval is large, i.e.  $n \gg 1$  in  $(x - \delta x, x + \delta x)$ . We then simplify to leading order, integrate over the spatial average interval, and perform spatial mean-field approximations using  $n \gg 1$  in the spatial average interval. We arrive at the force term,  $-(1/\eta) \partial^2 f / \partial x^2$ , written in Equation (3), where the continuous cell-cell interaction force,  $f$ , is given by Equation (4). This force term has an important physical interpretation where the cell density flux,  $j(x, t)$ , is equal to the gradient of the cell-cell interaction force,

$$j(x, t) = \frac{1}{\eta} \frac{\partial f(x, t)}{\partial x}. \tag{3A.1.27}$$

Further we find the cell velocity,  $u(x, t)$ , is related to the cell density and gradient of the cell-cell interaction force through Equation (5).

## 3A.2 Numerical methods

Here we present the numerical methods used to solve the discrete and continuum models. Key algorithms used to generate results are available on GitHub

<https://github.com/ryanmurphy42/Murphy2020a.git>.

### 3A.2.1 Discrete model

We numerically solve the discrete model with a constant time step algorithm, using a forward Euler approximation to integrate the discrete equations (1), and rejection sampling to determine when proliferation and death events occur [74]. This method is valid for all proliferation and death law mechanisms which we consider. However, we note that to improve computational efficiency, in the case of the constant proliferation and death law mechanism, we could use Gillespie's algorithm [76]. This is possible as the propensities of cells to proliferate or die are constant within the calculated time to the next reaction interval. This is more difficult for the linear and logistic proliferation and death mechanisms where, due to mechanical coupling, the propensity of a cell to proliferate or die can vary appreciably within the calculated time to the next reaction interval per the Gillespie method. In such a case the Extrande method may be considered [240].

#### Euler's method

To simulate a single discrete realization, we initialise the model with  $N$  cells. We prescribe each cell  $i$  with the mechanical cell properties including cell stiffness  $k_i^N$  and resting cell length  $a_i^N$ . We prescribe proliferation and death mechanisms to each cell  $i$  and any associated proliferation or death cell properties. We define the initial cell positions and then for each time step of size  $\Delta t = 0.0001$  we update the cell positions using a simple forward Euler method to integrate Equations (1) numerically. At the end of each time step we determine whether a proliferation or death event occurs and if so which cell has proliferated or died. To do so we use rejection sampling [74] where we generate three independent random numbers from a uniform distribution,  $r_1, r_2, r_3 \sim U[0, 1]$ . Then a cell event, which could be either a cell proliferation or cell death event, occurs when

$$r_1 < \sum_{i=1}^{N(t)} P(l_i^N) \Delta t + \sum_{i=1}^{N(t)} D(l_i^N) \Delta t,$$

i.e. with probability  $\sum_{i=1}^{N(t)} [P(l_i^N) + D(l_i^N)] \Delta t$ . Given that a cell event occurs, a proliferation event occurs if

$$r_2 < \frac{\sum_{i=1}^{N(t)} P(l_i^N)}{\sum_{i=1}^{N(t)} P(l_i^N) + \sum_{i=1}^{N(t)} D(l_i^N)}.$$

Otherwise we have a cell death event. To determine which cell is proliferating, similarly for dying, we find the index  $j$  which satisfies,

$$\frac{\sum_{i=1}^j P(l_i^N)}{\sum_{i=1}^{N(t)} P(l_i^N)} < r_3 \leq \frac{\sum_{i=1}^{j+1} P(l_i^N)}{\sum_{i=1}^{N(t)} P(l_i^N)}. \quad (3A.2.1)$$

We then update the node positions, cell properties, and indices according to the model description in Section 2.1. We repeat for each time step until we have reached the final time. This approach requires that at most one cell event can occur within each time step which is satisfied for the parameters used in this work. For other parameters where this assumption may not be valid the size of the time step could be reduced.

### 3A.2.2 Continuum model

We now outline the numerical method we use to solve the continuum model. First, for completeness, we rewrite the governing equations for cell density,  $q(x, t)$ ,

$$\frac{\partial q(x, t)}{\partial t} = -\frac{1}{\eta} \frac{\partial^2 f(x, t)}{\partial x^2} + q(x, t) P\left(\frac{1}{q(x, t)}\right) - q(x, t) D\left(\frac{1}{q(x, t)}\right), \quad (3A.2.2)$$

where the cell-cell interaction force,  $f(x, t)$ , is given by

$$f(x, t) = k(x, t) \left( \frac{1}{q(x, t)} - a(x, t) \right), \quad (3A.2.3)$$

and  $P(1/q(x, t))$ ,  $D(1/q(x, t))$  are the proliferation and death mechanisms, respectively. The cell properties are governed by

$$\frac{\partial \chi(x, t)}{\partial t} + u(x, t) \frac{\partial \chi(x, t)}{\partial x} = 0, \quad \chi = k, a, \beta, \gamma, l_d, \quad (3A.2.4)$$

where the cell velocity  $u(x, t)$  is given by,

$$u(x, t) = \frac{1}{\eta q(x, t)} \frac{\partial f(x, t)}{\partial x}. \quad (3A.2.5)$$

In Equation (3A.2.2) for the cell density we have a second-order spatial derivative, and in the cell property equations, Equations (3A.2.4), we have first-order spatial derivatives. Both equations have first-order time derivatives. To begin we discretise the domain of fixed length  $L$  with a uniform mesh with spatial step  $\Delta x$ . We discretise time with a uniform mesh with time step  $\Delta t$ . Second-order spatial derivative terms are approximated by standard central differences. First-order spatial derivatives are approximated by standard upwind differences. Temporal derivatives are approximated by a Crank-Nicolson approximation. We then have a system of nonlinear algebraic equations for the cell density, cell stiffness, resting cell length, and any other cell properties specified by the proliferation and/or death laws. Each resulting system of nonlinear algebraic equations we solve sequentially within the same Newton-Raphson iteration [44] until a convergence tolerance,  $\epsilon$ , is satisfied. In each iteration each resulting system of linearised tridiagonal algebraic equations is solved using the Thomas algorithm [254]. There are three key choices with this method  $\Delta x$ ,  $\Delta t$ , and  $\epsilon$ . Any implementation of the numerical method should ensure that  $\Delta x$ ,  $\Delta t$ , and  $\epsilon$  are sufficiently small that the solution is grid-independent. In the results we present we take  $\Delta x = 0.01$ ,  $\Delta t = 0.00001$ , and  $\epsilon = 0.001$ .

For convenience, we now explain the Newton-Raphson method in more detail and explicitly

present how we discretise the cell density equation, its boundary condition, discretise cell property equations, their boundary conditions, and update the interface position when there are two adjacent populations.

### Newton-Raphson method

The following notation is used. We use the subscript  $j = 1, 2, \dots, J$  to represent spatial nodes. We use the superscript  $n = 1, 2, \dots, T$  to represent temporal nodes. We use the superscript  $r = 0, 1, \dots, R_n$  to represent the Newton-Raphson iterate within time step  $n$ , where iterate  $r = R_n$  is the final iterate which meets the convergence tolerance  $\epsilon$ . For convenience, we will drop the  $R_n$  notation for the final iterate of time step  $n$ , for example, we write  $q_j^n = q_j^{n,R_n}$  for the cell density at spatial node  $j$ .

We now solve for each variable at time step  $n+1$ . The initial iterate, corresponding to  $r = 0$ , is given by  $q_j^{n+1,0} = q_j^n$  and  $\chi_j^{n+1,0} = \chi_j^n$  for  $\chi = k, a, \beta, \gamma, l_d$ . The following is for the  $r^{\text{th}}$  iterate.

We first solve for the **cell density**. We substitute Equation (3A.2.3) into Equation (3A.2.2). For internal spatial nodes,  $j = 2, 3, \dots, J - 1$ , we have, after rearranging so that all terms are on the right hand side, and for convenience multiplying all terms by  $\Delta t$ , a system of algebraic equations

$$\begin{aligned} 0 = & -q_j^{n+1,r} + q_j^n \\ & - \frac{\Delta t}{2\eta(\Delta x)^2} \left[ k_{j-1}^n \left( \frac{1}{q_{j-1}^n} - a_{j-1}^n \right) - 2k_j^n \left( \frac{1}{q_j^n} - a_j^n \right) + k_{j+1}^n \left( \frac{1}{q_{j+1}^n} - a_{j+1}^n \right) \right] \\ & - \frac{\Delta t}{2\eta(\Delta x)^2} \left[ k_{j-1}^{n+1,r-1} \left( \frac{1}{q_{j-1}^{n+1,r}} - a_{j-1}^{n+1,r-1} \right) - 2k_j^{n+1,r-1} \left( \frac{1}{q_j^{n+1,r}} - a_j^{n+1,r-1} \right) \right. \\ & \quad \left. + k_{j+1}^{n+1,r-1} \left( \frac{1}{q_{j+1}^{n+1,r}} - a_{j+1}^{n+1,r-1} \right) \right] \\ & + \frac{\Delta t}{2} \left[ q_j^n P \left( \frac{1}{q_j^n} \right) - q_j^n D \left( \frac{1}{q_j^n} \right) \right] \\ & + \frac{\Delta t}{2} \left[ q_j^{n+1,r} P \left( \frac{1}{q_j^{n+1,r}} \right) - q_j^{n+1,r} D \left( \frac{1}{q_j^{n+1,r}} \right) \right]. \end{aligned} \tag{3A.2.6}$$

For boundary nodes,  $j = 1, J$ , we have fixed boundaries which correspond to zero velocity boundary conditions

$$\frac{\partial f(x, t)}{\partial x} = 0, \quad x = 0, L. \tag{3A.2.7}$$

We apply a forward difference approximation to Equation (3A.2.7) at  $x = 0$ , corresponding to

spatial node  $j = 1$ ,

$$0 = \frac{1}{\Delta x} \left[ k_2^{n+1,r-1} \left( \frac{1}{q_2^{n+1,r}} - a_2^{n+1,r-1} \right) - k_1^{n+1,r-1} \left( \frac{1}{q_1^{n+1,r}} - a_1^{n+1,r-1} \right) \right], \quad (3A.2.8)$$

and a backward difference approximation to Equation (3A.2.7) at  $x = L$ , corresponding to spatial node  $j = J$ ,

$$0 = \frac{1}{\Delta x} \left[ k_J^{n+1,r-1} \left( \frac{1}{q_J^{n+1,r}} - a_J^{n+1,r-1} \right) - k_{J-1}^{n+1,r-1} \left( \frac{1}{q_{J-1}^{n+1,r}} - a_{J-1}^{n+1,r-1} \right) \right]. \quad (3A.2.9)$$

Equations (3A.2.6), (3A.2.8) and (3A.2.9) can be combined to form a tridiagonal matrix  $\mathbf{F}(q^{n+1,r})$ , where  $q^{n+1,r}$  is the system of algebraic equations of the  $r^{\text{th}}$  iterate of the Newton-Raphson method at time step  $n + 1$ . We then calculate the Jacobian,  $\mathbf{J}(q^{n+1,r})$ , of  $\mathbf{F}(q^{n+1,r})$ . We form a linear system for the  $r^{\text{th}}$  iterate of the Newton-Raphson method within the time step  $n + 1$  as

$$\mathbf{J}(q^{n+1,r}) \delta q^{n+1,r} = -\mathbf{F}(q^{n+1,r}), \quad (3A.2.10)$$

where  $\delta q^{n+1,r}$  is the Newton-Raphson correction. As  $\mathbf{F}(q^{n+1,r})$  is tridiagonal we solve this using the Thomas algorithm [254], to determine the next iterate,

$$q^{n+1,r+1} = q^{n+1,r} + \delta q^{n+1,r} \quad (3A.2.11)$$

Next we sequentially solve for each **cell property**. We calculate the velocity at each node,  $v_j^{n+1,r}$ ,

$$v_j^{n+1,r} = \frac{1}{\eta} \frac{1}{q_j^{n+1,r}} \left[ k_{j+1}^{n+1,r-1} \left( \frac{1}{q_{j+1}^{n,r}} - a_{j+1}^{n+1,r-1} \right) - k_{j-1}^{n+1,r-1} \left( \frac{1}{q_{j-1}^{n+1,r}} - a_{j-1}^{n+1,r-1} \right) \right]. \quad (3A.2.12)$$

Then we substitute Equation (3A.2.5) into Equation (3A.2.4) and consider  $\chi = k$ . If  $v_j^{n+1,r} > 0$ , then we apply a backward difference approximation to the first-order spatial derivatives and a Crank-Nicolson approximation for the time derivative

$$\begin{aligned} 0 = & k_j^{n+1,r} - k_j^n \\ & - \frac{\Delta t}{2\eta(\Delta x)^2} \left\{ \left[ k_j^n \left( \frac{1}{q_j^n} - a_j^n \right) - k_{j-1}^n \left( \frac{1}{q_{j-1}^n} - a_{j-1}^n \right) \right] [k_j^n - k_{j-1}^n] \right. \\ & \left. + \left[ k_j^{n+1,r} \left( \frac{1}{q_j^{n+1,r}} - a_j^{n+1,r-1} \right) - k_{j-1}^{n+1,r} \left( \frac{1}{q_{j-1}^{n+1,r}} - a_{j-1}^{n+1,r-1} \right) \right] [k_j^{n+1,r} - k_{j-1}^{n+1,r}] \right\}. \end{aligned} \quad (3A.2.13)$$

Similarly, if  $v_j^{n+1,r} < 0$ , then we apply a forward difference approximation.

For only mechanical relaxation or for mechanical relaxation with proliferation, we have fixed boundary conditions at  $x = 0, L$ , at spatial nodes  $j = 1, J$ , respectively,

$$\begin{aligned} k_1^{n+1,r} &= k_1^n, \\ k_J^{n+1,r} &= k_J^n. \end{aligned} \quad (3A.2.14)$$

These boundary conditions also apply for cell death in a homogeneous population. However, when we have cell death with two adjacent populations, for example in Section 3.2, we need to modify a boundary condition when one population becomes extinct. To do so, we calculate the total cell number for each tissue within each time step. When the total number of cells in a tissue decreases below one we remove this tissue by setting the interface position equal to the relevant domain boundary, and make cell properties homogeneous across the domain by setting them equal to the values in the remaining tissue. Then, similarly to Equations (3A.2.10) and (3A.2.11), we solve for  $k^{n+1,r}$ .

In the examples presented in this work, the resting cell length,  $a$ , and proliferation and death properties,  $\beta, \gamma$  and  $l_d$ , are homogeneous across the population. Therefore, they do not need to be simulated and we set  $\delta a^{n+1,r} = \delta \beta^{n+1,r} = \delta \gamma^{n+1,r} = \delta l_d^{n+1,r} = 0$ . However, if any of  $a, \beta, \gamma$  or  $l_d$  are heterogeneous, we extend the above by discretising the relevant cell property equations and boundary conditions similarly to Equations (3A.2.13) and (3A.2.14).

When we have two populations we also update the interface position,  $s(t)$ . To do so, we find the closest spatial node to  $s(t)$  and calculate the velocity at this node. Suppose that the closest spatial node is at node  $j$  then we set  $\delta s^{n+1,r} = v_j^{n+1,r} \Delta t$  and update the interface position through

$$s^{n+1,r} = s^n + \delta s^{n+1,r}. \quad (3A.2.15)$$

We iterate until  $\|\delta q^{n+1,r}, \delta k^{n+1,r}, \delta a^{n+1,r}, \delta \beta^{n+1,r}, \delta \gamma^{n+1,r}, \delta l_d^{n+1,r}, \delta s^{n+1,r}\|_\infty < \epsilon$ .

## Initial conditions

Equations (3)-(6) have the following initial conditions, for  $0 < x < L$ ,

$$q(x, 0) = q_0(x), \quad \chi(x, 0) = \chi_0(x) \text{ for } \chi = k, a, \beta, \gamma, l_d. \quad (3A.2.16)$$

If initial conditions are provided only for the discrete model they can be converted to continuum model initial conditions. This has been discussed in previous work, see Supplementary Section 1 of Murphy et al. (2019) [157].

### 3A.3 Homogeneous population

We now present additional results for the homogeneous population: an exact calculation for extinction with constant proliferation and death mechanisms, and results for logistic proliferation and death mechanisms.

#### 3A.3.1 Extinction for constant proliferation and death

In Section 3.1 of the main manuscript, we consider a homogeneous population with constant proliferation and constant death mechanisms with equal rates. In this case, the average of the discrete realizations shows good agreement with the corresponding solution of the continuum model. However, each individual realization exhibits very different behaviour to the continuum model. Due to the total cell number following a linear birth-death process independent of mechanical relaxation, each individual realization will eventually become extinct and the averaged total cell number displays increasing variance with time (Ross 1996). We now present the exact expressions for the extinction probability, mean, and standard deviation of the linear birth-death process. Corresponding results from the discrete realizations are then shown to match these expressions.

For the constant proliferation and death law combination with equal proliferation and death rates,  $\beta$ , we follow the work of Morgan (1977) who applies conditioning arguments to find the following probability generating function

$$G(z; t) = \left( \frac{z + \beta t (1 - z)}{1 + \beta t (1 - z)} \right)^{N(0)}, \quad (3A.3.1)$$

where  $N(0)$  is the initial cell population and  $z$  is a dummy variable defined for  $|z| \leq 1$ . The extinction probability,  $\mathbb{P}(E)$ , is

$$\mathbb{P}(E) = \mathbb{P}(N(t) = 0) = G(0; t) = \left( \frac{\beta t}{1 + \beta t} \right)^{N(0)}. \quad (3A.3.2)$$

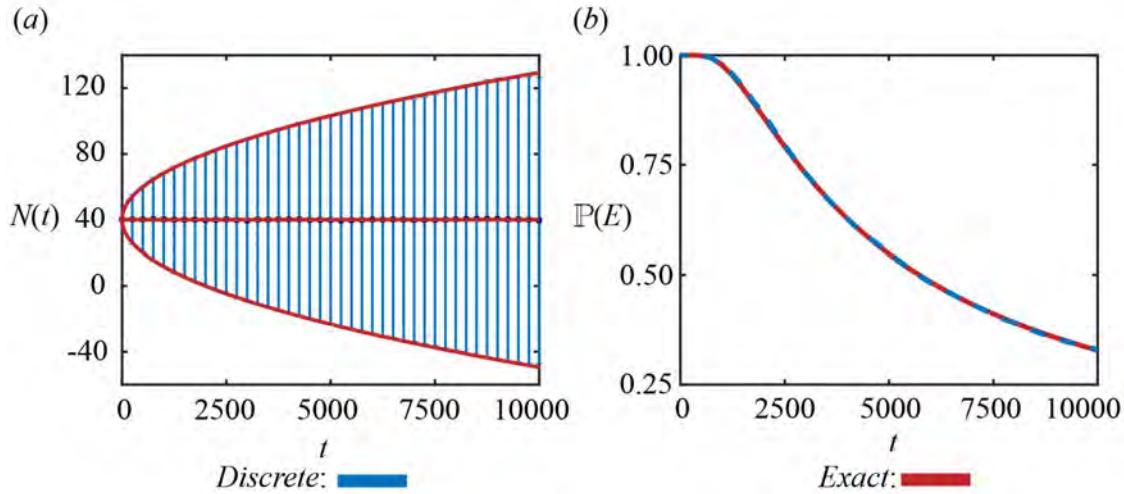
We observe that  $\mathbb{P}(N(t) = 0) \rightarrow 1$  as  $t \rightarrow \infty$ , i.e. every individual realization will eventually become extinct. The mean is the same as the initial cell number,

$$\mathbb{E}(N(t)) = \frac{\partial G(z; t)}{\partial z} \Big|_{z=1} = N(0). \quad (3A.3.3)$$

The standard deviation,  $\sigma(t)$ , is

$$\sigma(t) = \sqrt{2N(0)\beta t}. \quad (3A.3.4)$$

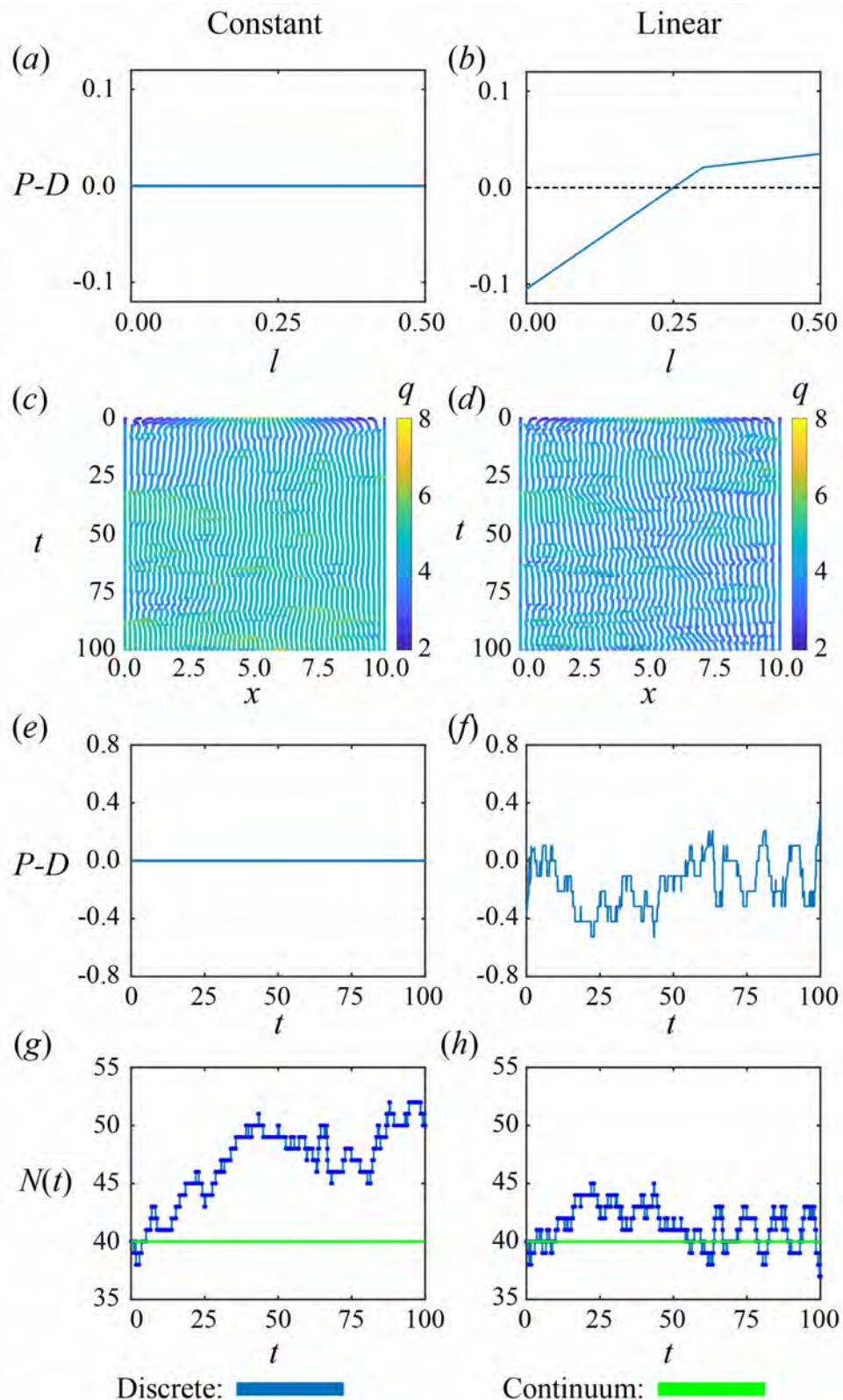
We now compare these exact formulas with our discrete simulation, with  $N(0) = 40$  and find very good agreement (Figure 3A.3).



**Figure 3A.3:** Comparison of 5000 discrete realizations with the exact expressions for the (a) mean and variance of total cell number and the (b) probability of extinction,  $\mathbb{P}(E)$ . Results shown for a homogeneous population with  $k = 10$ ,  $a = 0$ , constant proliferation and death laws with equal rates,  $\beta = 0.01$ , and  $0 < t < 10000$ .

### 3A.3.2 Reduced variance with cell-length-dependent mechanisms

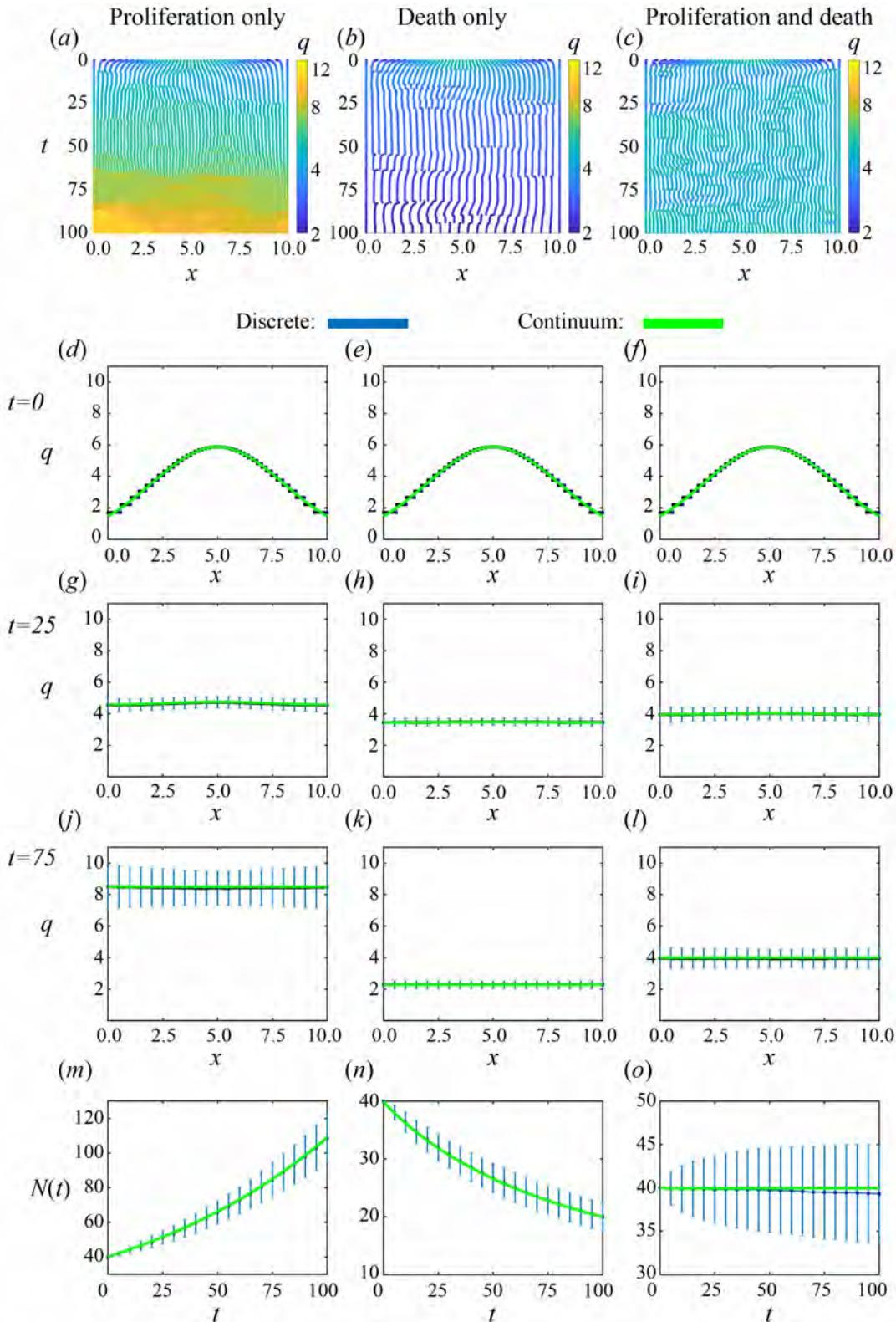
In Figures 3 and 4 in the main manuscript, we observe reduced variance in population with cell-length dependent proliferation and death mechanisms in comparison to cell-length independent proliferation and death mechanisms. Here, we explore the difference between constant and linear proliferation death mechanisms by presenting a single realisation for each (Figure 3A.4). In the constant proliferation and death mechanism case the net proliferation rate, which is the sum of probabilities of each cell to proliferate minus the sum of probabilities of each cell to die, is always zero (Figure 3A.4e). However, with the linear proliferation and death mechanism the net proliferation rate adjusts, due to changes in number of cells and their cell lengths, to stabilise the population at its equilibrium value (Figure 3A.4f,h).



**Figure 3A.4:** Homogeneous population with constant and linear cell-length dependent proliferation and death mechanisms. (a)-(b) Net proliferation rate for a single cell,  $P - D$ , dependent cell length,  $l$ . (c)-(d) Single realizations of cell boundary characteristics for  $0 \leq t \leq 100$ . (e)-(f) Net proliferation rate for the single realizations. (g)-(h) Total cell number for single realization (blue) compared to the continuum solution (green).

### 3A.3.3 Homogeneous population: logistic proliferation and death

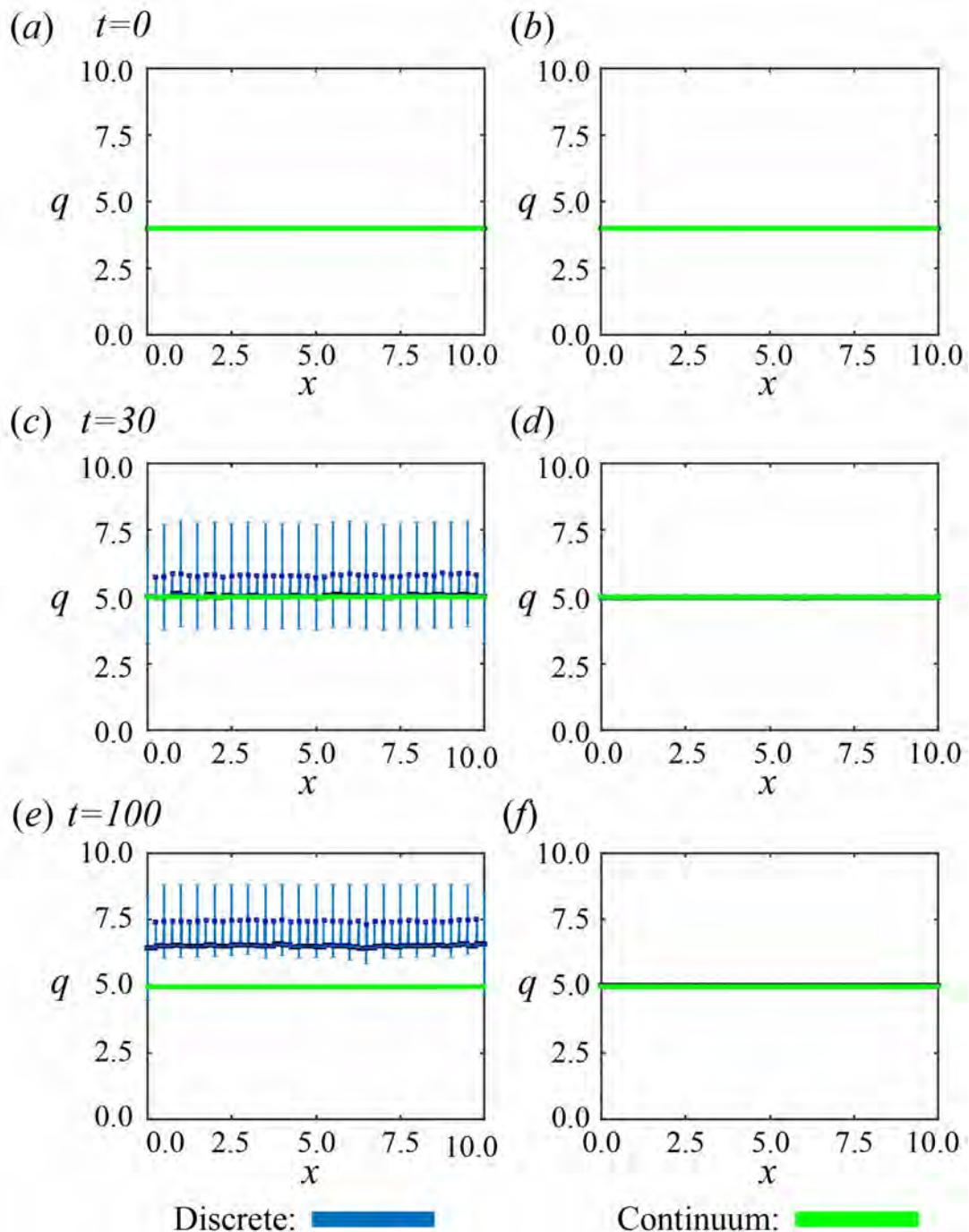
In Figures 3 and 4 in the main manuscript we present results for a homogeneous population with constant and linear cell-length-dependent proliferation mechanisms, respectively. We now present results for homogeneous population for the logistic proliferation and death mechanism (Figure 3A.5).



**Figure 3A.5:** Homogeneous population with **logistic** proliferation and death mechanisms. Proliferation only, death only, and proliferation with death shown in left, middle and right columns, respectively. (a)-(c) Single realizations of cell boundary characteristics for  $0 \leq t \leq 100$ . (d)-(f), (g)-(i), (j)-(l) Density snapshots at times  $t = 0, 25, 75$ , respectively. (m)-(o) Total cell number. The average and standard deviation (blue error bars) of 2000 discrete simulations are compared to solution of continuum model (green).

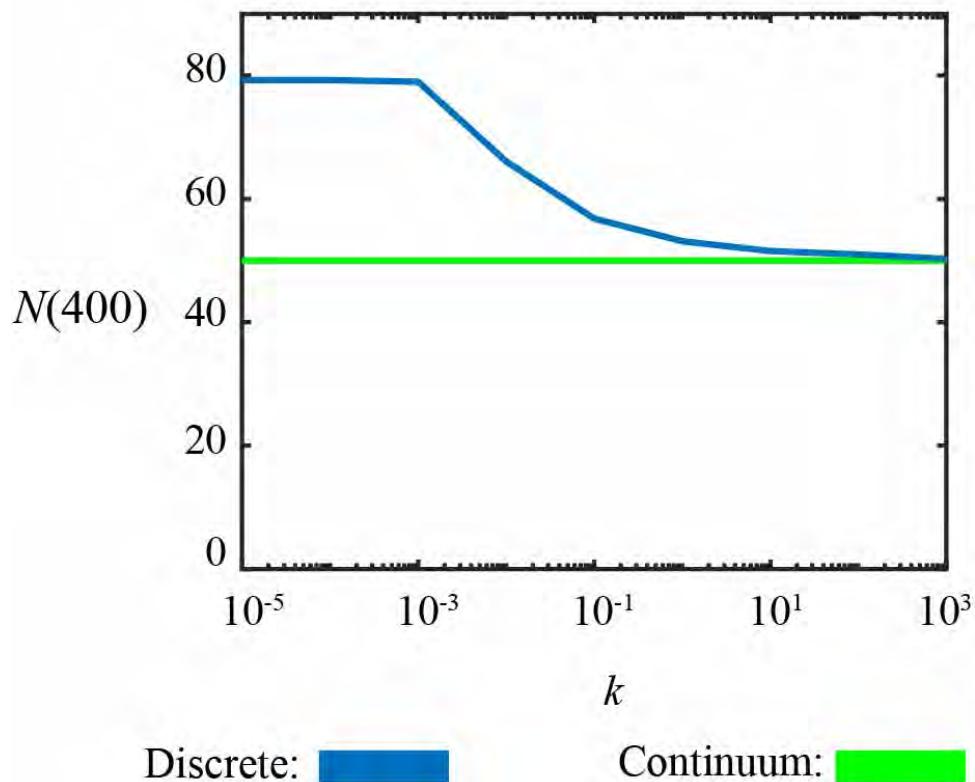
### 3A.3.4 Piecewise proliferation: varying mechanical relaxation rate

In Figure 7 in the main manuscript we show that the solution of the continuum model can differ significantly from the solution to the discrete model with slow mechanical relaxation,  $k = 0.0001$ , and matches extremely well when  $k = 1000$ . We now present the density snapshots corresponding to Figure 7. For  $k = 0.0001$  the density snapshots have jumps at locations of the initial cell boundaries and density in the discrete model is higher than the continuum model (Figure 3A.6a,c,e). For  $k = 1000$  the discrete model and continuum model are consistent (Figure 3A.6b,d,f).



**Figure 3A.6:** Homogeneous population with piecewise proliferation mechanism. With slow mechanical relaxation,  $k = 0.0001$ , and faster mechanical relaxation,  $k = 1000$ , shown in left and right columns, respectively. (a)-(f) Density snapshots at times  $t = 0, 30, 100$  where the average and standard deviation (blue error bars) of 2000 discrete simulations are compared to solution of continuum model (green).

In Figure 7 in the main manuscript we show that at later times the solution of the continuum model can differ significantly from the solution to the discrete model with slow mechanical relaxation,  $k = 0.0001$ , and matches extremely well when  $k = 1000$ . We now show  $N(400)$  for intermediate values of  $k$  in Figure 3A.7.



**Figure 3A.7:** Homogeneous population with piecewise proliferation mechanism.  $N(400)$  with varying mechanical relaxation rates through cell stiffness  $k$ . Increasing the cell stiffness  $k$  improves matching between discrete and continuum model. Cell stiffness plotted on a logarithmic axis.

### 3A.4 Mechanical cell competition

We now present additional results for a heterogeneous population: an exact calculation for cell size at mechanical equilibrium, and results for logistic proliferation and death.

#### 3A.4.1 Two populations: cell size at mechanical equilibrium

Here we show that, for cells in extension, cells with lower cell stiffness are larger than cells with higher cell stiffness at mechanical equilibrium. As mechanical relaxation is chosen to be fast relative to the proliferation and death rates the system will be at mechanical equilibrium, except for the short transition following a proliferation or death event.

In Section 3.2 of the main manuscript, we are interested in two adjacent populations which we denote tissue 1 and tissue 2. We assume there are  $N_i$  cells in tissue  $i$  and cells in tissue  $i$  have cell stiffness  $K_i$  and resting cell length  $A_i$ . From previous work [157] the interface position at mechanical equilibrium,  $S = \lim_{t \rightarrow \infty} s(t)$ , is

$$S = \frac{\frac{K_1 A_1}{K_2} + \frac{L}{N_2} - A_2}{\frac{K_1}{K_2 N_1} + \frac{1}{N_2}}. \quad (3A.4.1)$$

Assuming  $A_i = 0$  this simplifies to

$$S = \frac{L}{\frac{K_1 N_2}{K_2 N_1} + 1}. \quad (3A.4.2)$$

Letting  $l_i$  be the length of a cell in tissue  $i$  then

$$\begin{aligned} l_1 &= \frac{S}{N_1}, \\ l_2 &= \frac{L - S}{N_2}. \end{aligned} \quad (3A.4.3)$$

Substituting Equation (3A.4.2) into Equation (3A.4.3) gives

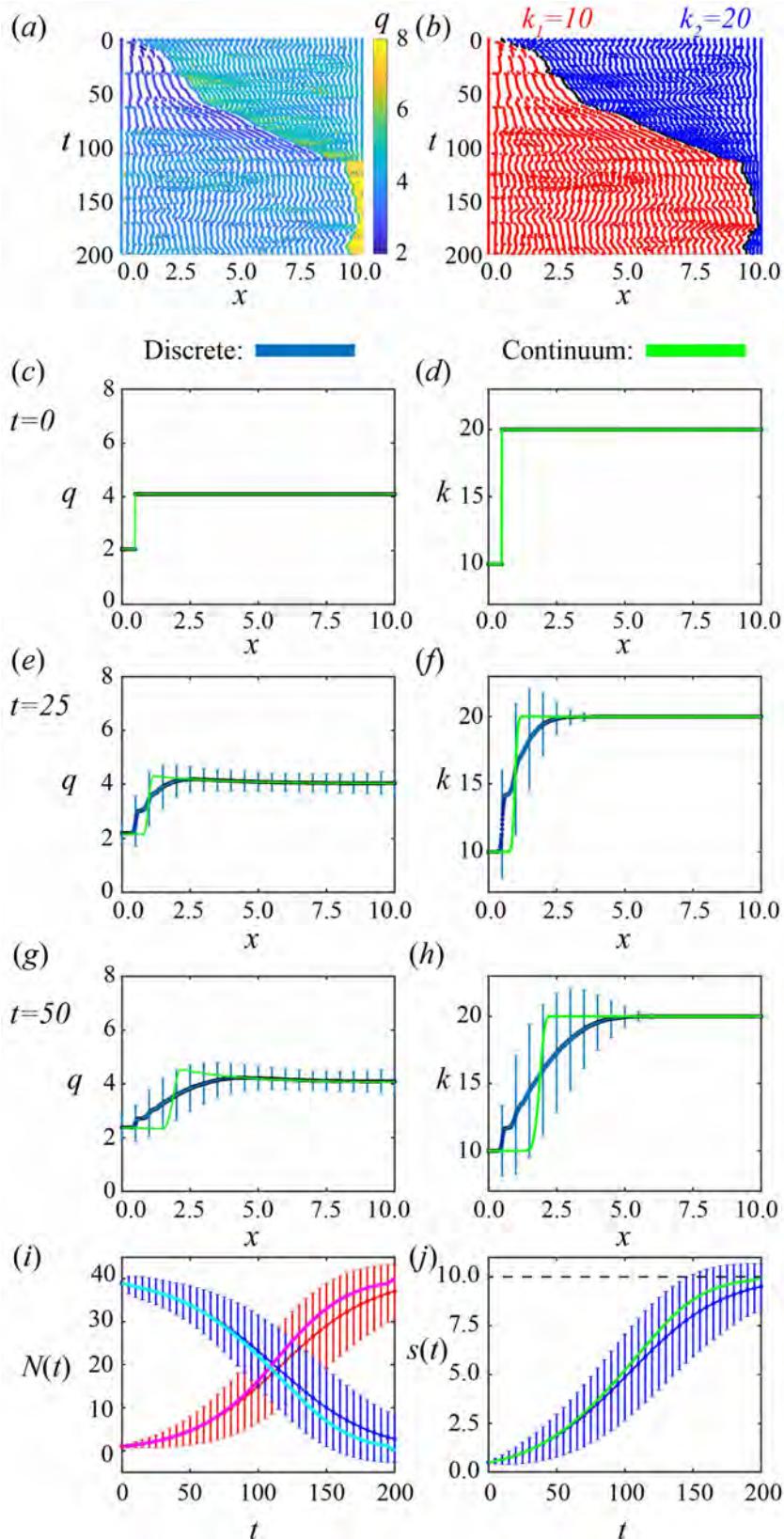
$$\begin{aligned} l_1 &= \frac{L}{\frac{K_1 N_2}{K_2} + N_1}, \\ l_2 &= \frac{L}{N_2 + \frac{K_2 N_1}{K_1}}. \end{aligned} \quad (3A.4.4)$$

It can then be shown that if  $K_1 < K_2$  then  $l_1 > l_2$ . This corresponds to cells of lower stiffness being larger at mechanical equilibrium and this is independent of the number of cells in tissue 1 and 2.

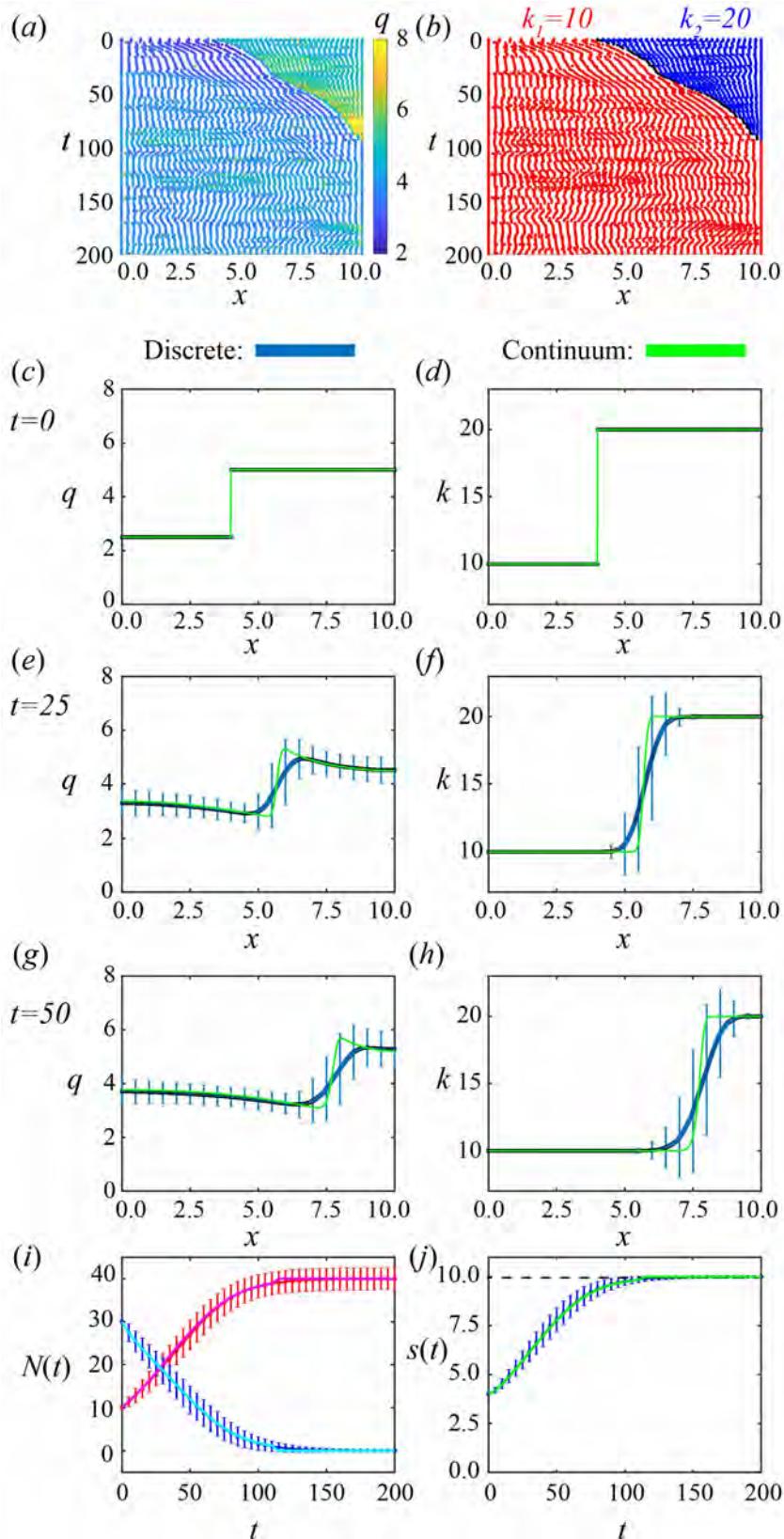
### 3A.4.2 Mechanical cell competition: robustness to initial conditions

In this section, we explore the robustness of results to changes in the initial conditions. We consider the linear combination as in Section 3.2 and Figure 6. In Figure 6 we started the simulation with 20 cancer cells, each with cell stiffness  $k_1 = 10$ , and 20 healthy cells, each with cell stiffness  $k_2 = 20$ . We first keep  $k_1 = 10$  and  $k_2 = 20$  and start the simulations with one cancer cell and 39 healthy cells in Figure 3A.8, and with 10 cancer cells and 30 healthy cells in Figure 3A.9. We then set  $k_1 = 10$  and  $k_2 = 11$  and start with 20 cancer cells and 20 healthy cells in Figure 3A.10. In all results we observe that same behaviour that the difference in cell stiffness is sufficient to allow cancer to invade and takeover the tissue, provided cancer cells have lower stiffness than healthy cells.

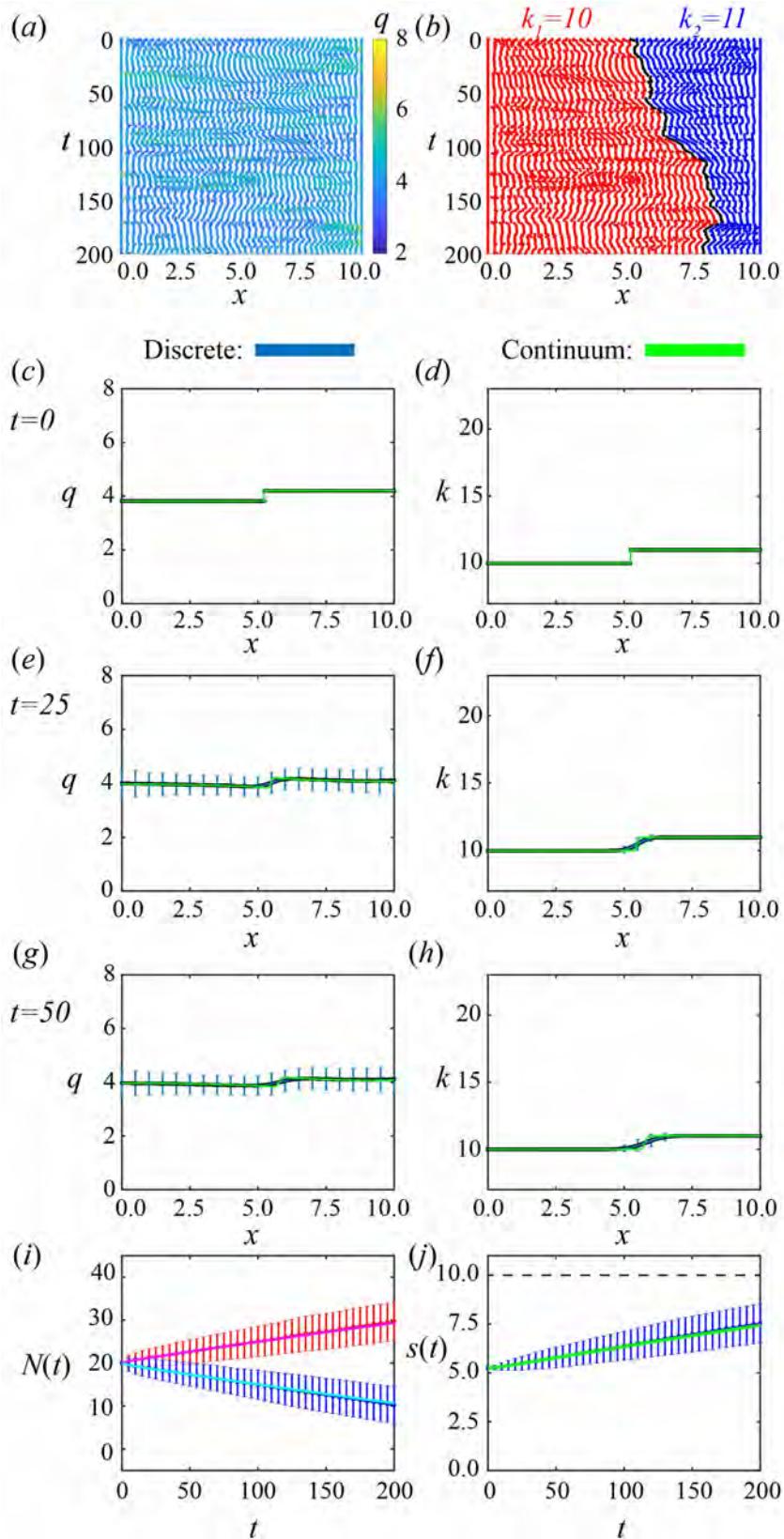
We also note that when starting the simulation with one cancer cell, as in Figure 3A.8, there is a slight mismatch between the discrete model and the continuum model. This is because in the discrete model the cancer cells will initially go extinct with a higher probability than if the model is started with a higher initial number of cancer cells as in Figures 3A.9 and 6.



**Figure 3A.8:** Results for cancer invasion with adjacent populations using **linear** proliferation and death mechanisms with **proliferation and death**. Starting with 1 cancer cell and 39 healthy cells. First row shows a single realization of cell boundary characteristics for  $0 \leq t \leq 200$ . Colouring in (a),(b) represent cell density and cell stiffness, respectively. (c)-(d), (e)-(f), (g)-(h) Density and cell stiffness snapshots, left and right, respectively, at times  $t = 0, 25, 50$ , respectively. (i) Total cell number,  $N(t) > 0$ , for cancer (red/magenta) and healthy cells (blue/cyan) for the discrete/continuum solutions. (j) Interface position,  $s(t)$ , where the dotted line shows the edge of the domain. The average and standard deviation (blue error bar) of 2000 discrete simulations are compared to the solution of the continuum model (green).



**Figure 3A.9:** Results for cancer invasion with adjacent populations using **linear** proliferation and death mechanisms with **proliferation and death**. Starting with 1 cancer cell and 39 healthy cells. First row shows a single realization of cell boundary characteristics for  $0 \leq t \leq 200$ . Colouring in (a),(b) represent cell density and cell stiffness, respectively. (c)-(d), (e)-(f), (g)-(h) Density and cell stiffness snapshots, left and right, respectively, at times  $t = 0, 25, 50$ , respectively. (i) Total cell number,  $N(t) > 0$ , for cancer (red/magenta) and healthy cells (blue/cyan) for the discrete/continuum solutions. (j) Interface position,  $s(t)$ , where the dotted line shows the edge of the domain. The average and standard deviation (blue error bar) of 2000 discrete simulations are compared to the solution of the continuum model (green).



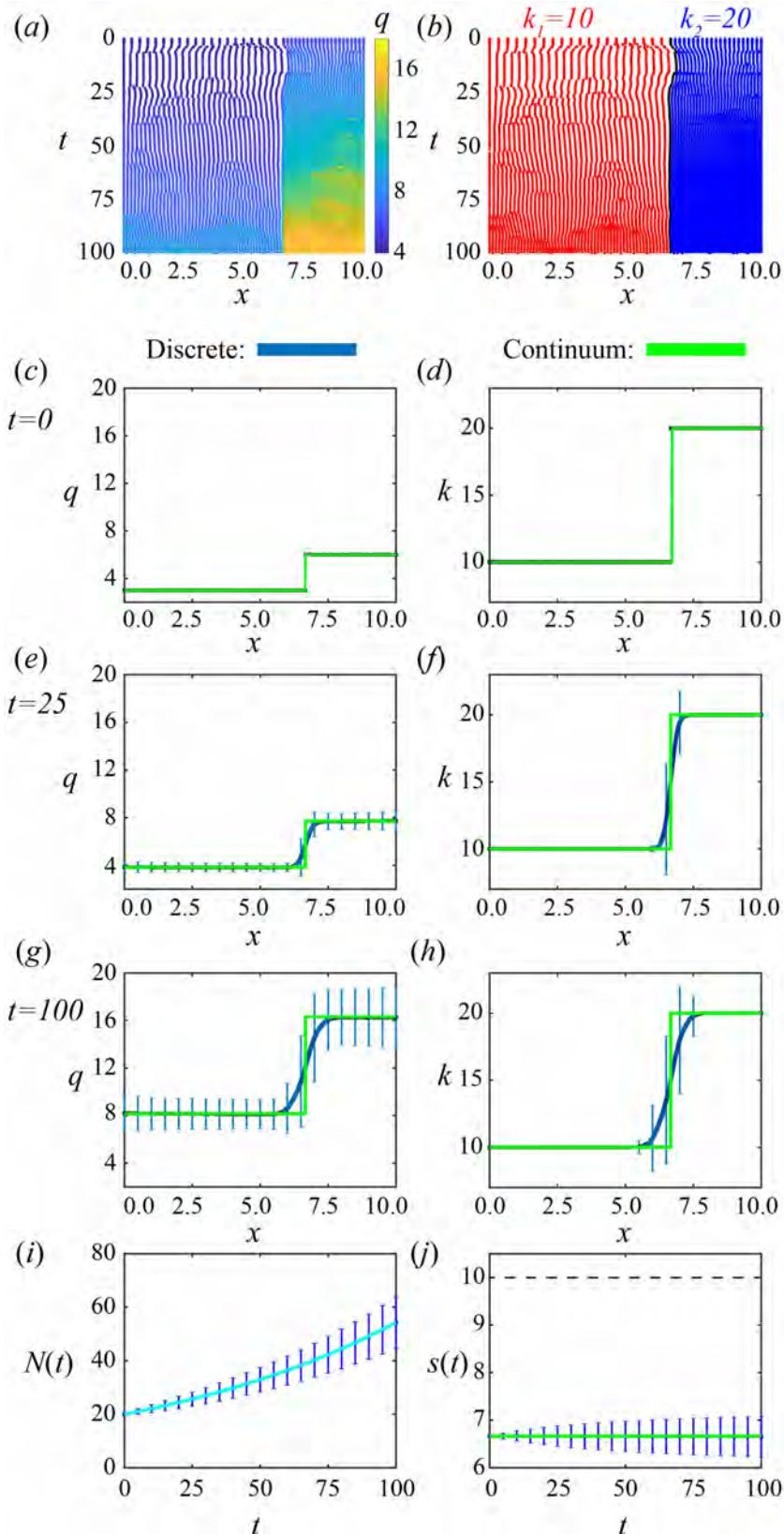
**Figure 3A.10:** Results for cancer invasion with adjacent populations using **linear** proliferation and death mechanisms with **proliferation and death**. Starting with 20 cancer cell and 20 healthy cells and set  $k_1 = 10$  and  $k_2 = 11$ . First row shows a single realization of cell boundary characteristics for  $0 \leq t \leq 200$ . Colouring in (a),(b) represent cell density and cell stiffness, respectively. (c)-(d), (e)-(f), (g)-(h) Density and cell stiffness snapshots, left and right, respectively, at times  $t = 0, 25, 50$ , respectively. (i) Total cell number,  $N(t) > 0$ , for cancer (red/magenta) and healthy cells (blue/cyan) for the discrete/continuum solutions. (j) Interface position,  $s(t)$ , where the dotted line shows the edge of the domain. The average and standard deviation (blue error bar) of 2000 discrete simulations are compared to the solution of the continuum model (green).

### 3A.4.3 Mechanical cell competition: logistic combination

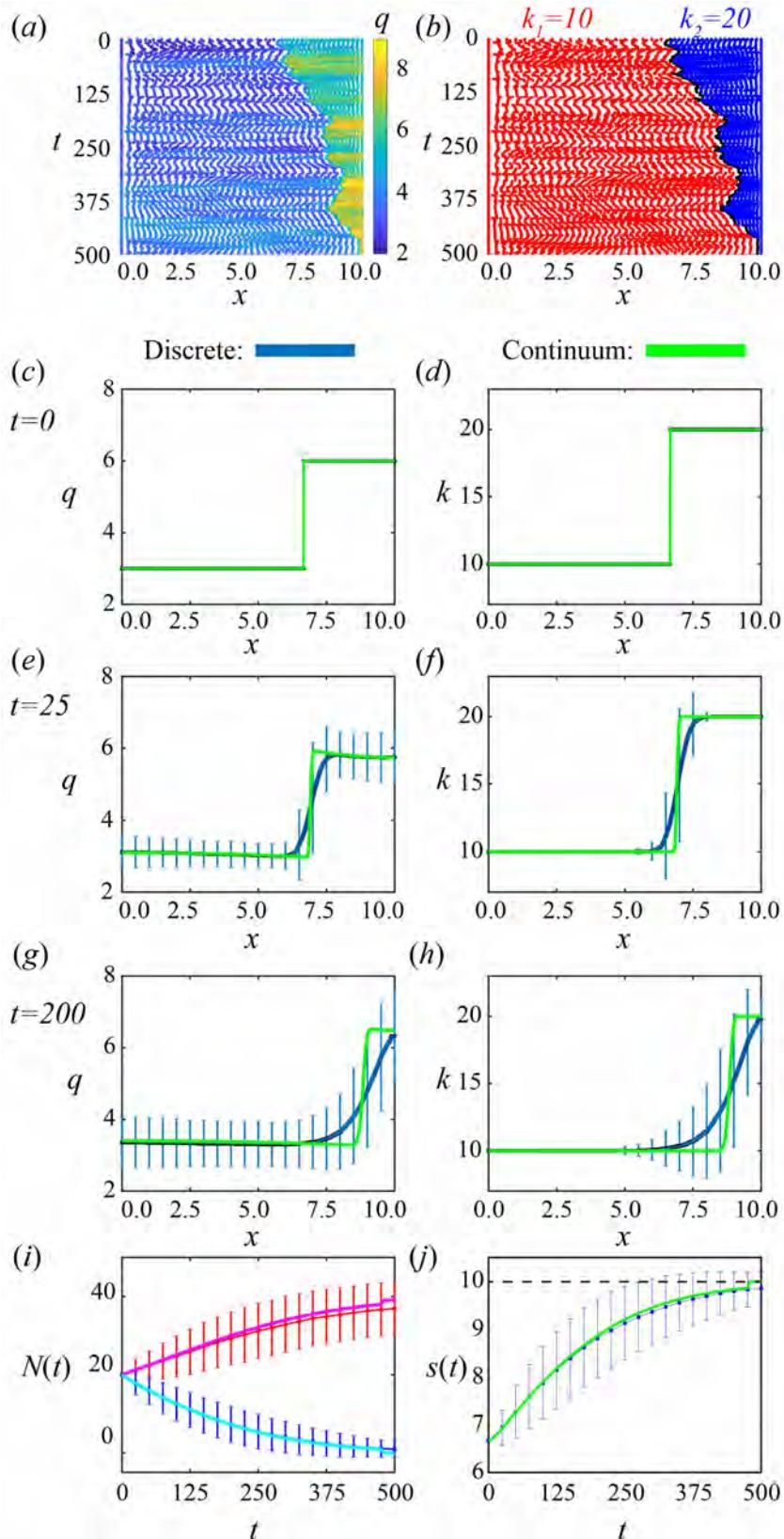
In this section, we repeat the scenario presented in Section 3.2 but now with the logistic proliferation and death mechanisms rather than the linear proliferation and death mechanisms. We observe similar qualitative results.

With mechanical relaxation and proliferation only, we observe that the interface position is on average the same as the initial condition which was chosen as mechanical equilibrium (Figure 3A.11). This is expected as proliferation is independent of cell length and therefore mechanical relaxation. Therefore we also expect that the total number of cells in tissue 1 and tissue 2 are, on average, equal (Figure 3A.11i). From Equation (3A.4.2) we expect the initial condition of mechanical equilibrium to be maintained. Similarly, in the continuum model both tissues have the same proliferation rate therefore the interface position,  $s(t)$ , is maintained (Figure 3A.11j).

For mechanical relaxation with proliferation and death, we now observe that the cancer cells extend and the healthy cells are smaller. The smaller healthy cells then eventually die and the cancer cells take over the domain. This is because the death mechanism is now dependent on cell length and therefore dependent on mechanical relaxation (Figure 3A.12). Note that the total cell number for each population does not decrease below zero, the error bars represent the standard deviation about the mean.



**Figure 3A.11:** Results for cancer invasion with adjacent populations using **logistic** proliferation and death mechanisms with **proliferation only**. (a),(b) A single realization of cell boundary characteristics for  $0 \leq t \leq 100$ . Colouring in (a),(b) represent cell density and cell stiffness, respectively. (c)-(d), (e)-(f), (g)-(h) Density and cell stiffness snapshots, left and right, respectively, at times  $t = 0, 25, 100$ , respectively. (i) Total cell number,  $N(t) > 0$ , for cancer (red/magenta) and healthy cells (blue/cyan) for the discrete/continuum solutions. (j) Interface position,  $s(t)$ , where the dotted line shows the edge of the domain. The average and standard deviation (blue error bars) of 2000 discrete simulations are compared to the solution of the continuum model (green).

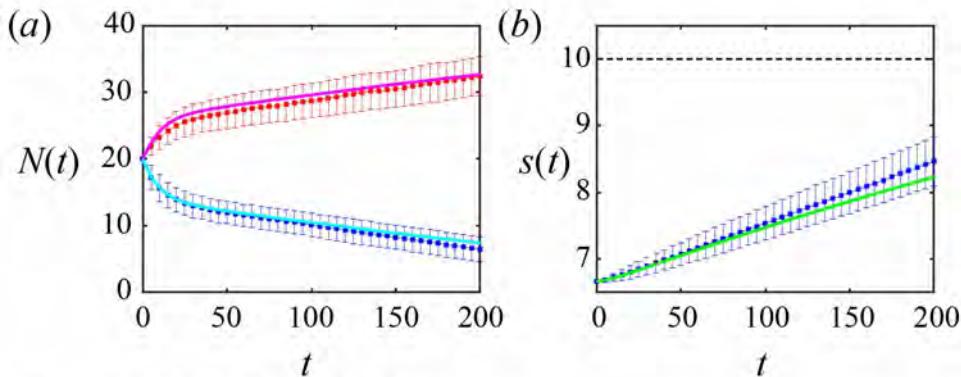


**Figure 3A.12:** Results for cancer invasion with adjacent populations using **logistic** proliferation and death mechanisms with **proliferation and death**. First row shows a single realization of cell boundary characteristics for  $0 \leq t \leq 500$ . Colouring in (a),(b) represent cell density and cell stiffness, respectively. (c)-(d), (e)-(f), (g)-(h) Density and cell stiffness snapshots, left and right respectively, at times  $t = 0, 25, 200$ , respectively. (i) Total cell number,  $N(t) > 0$ , for cancer (red/magenta) and healthy cells (blue/cyan) for the discrete/continuum solutions. (j) Interface position,  $s(t)$ , where the dotted line shows the edge of the domain. The average and standard deviation (blue error bars) of 2000 discrete simulations are compared to the solution of the continuum model (green).

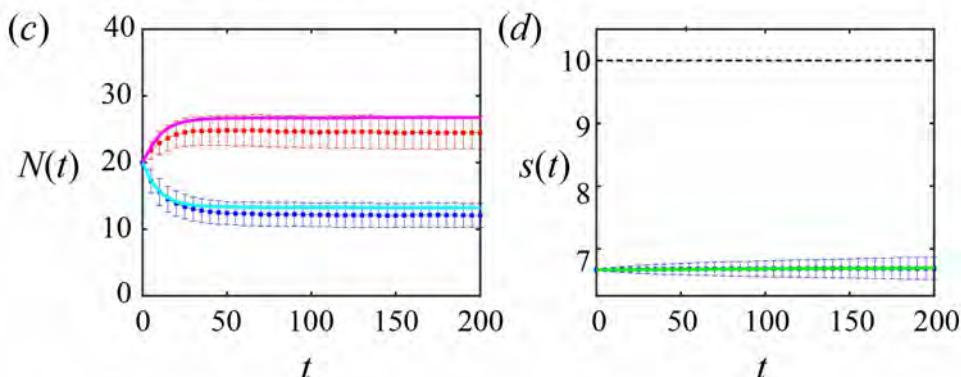
### 3A.4.4 Varying mechanical relaxation rate

In Figure 6 in the main manuscript, mechanical cell competition is considered with cell stiffnesses  $K_1 = 10, K_2 = 20$ , and linear proliferation and death mechanisms. Here, in Figure 3A.13, we present the same problem with reduced mechanical relaxation rates and observe that the continuum model is a reasonably good approximation even for cell stiffnesses  $K_1 = 0.0001, K_2 = 0.0002$ .

$$K_1 = 0.1, K_2 = 0.2$$



$$K_1 = 0.0001, K_2 = 0.0002$$



**Figure 3A.13:** Results for cancer invasion with adjacent populations using **linear** proliferation and death mechanisms with **proliferation and death** with slower mechanical relaxation,  $k = 0.1$  in (a,b) and  $k = 0.0001$  in (c,d). (a,c) Total cell number,  $N(t) > 0$ , for cancer (red/magenta) and healthy cells (blue/cyan) for the discrete/continuum solutions. (b,d) Interface position,  $s(t)$ , where the dotted line shows the edge of the domain. The average and standard deviation (blue error bar) of 2000 discrete simulations are compared to the solution of the continuum model (green).

## **Chapter 4**

# **Travelling waves in a free boundary mechanobiological model of an epithelial tissue**

## 4.0 Preamble

An article published in *Applied Mathematics Letters*

**Murphy RJ**, Buenzli PR, Baker RE, Simpson MJ (2021). Travelling waves in a free boundary mechanobiological model of an epithelial tissue. *Applied Mathematics Letters*. 111: 106636. doi: 10.1016/j.aml.2020.106636. arxiv preprint

This chapter includes Publication 3, addresses objective 3 and research question 3. This chapter extends the work of Chapters 2 and 3, by considering a free boundary rather than a fixed boundary and analysing the continuum model derived in earlier chapters. The key results include extending the continuum model to a free boundary and examining travelling wave behaviour. Using this we show travelling wave solutions that may invade or retreat depending on whether the carrying capacity density corresponds to cells being in compression or extension. Further, travelling wave solutions have well-defined fronts and are not associated with heteroclinic orbits in the phase plane. Furthermore, as this chapter focuses solely on the continuum model, derived in earlier chapters, and due to the formatting of the journal where this article has been published, the layout of this chapter is slightly different to other chapters in Part 1 of this thesis. This chapter includes the following sections: Abstract; Introduction; Mathematical Model; and Travelling waves, comprising the results, discussion, and conclusion. Supplementary material, including additional results, associated with publication 3, is presented in Chapter 4A.

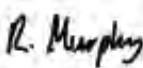
### Statement of Contribution of Co-Authors for Thesis by Published Paper

The authors listed below have certified that:

1. they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
2. they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
3. there are no other authors of the publication according to these criteria;
4. potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit, and
5. they agree to the use of the publication in the student's thesis and its publication on the QUT's ePrints site consistent with any limitations set by publisher requirements.

In this case of this chapter:

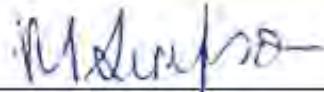
Murphy RJ, Buenzli PR, Baker RE, Simpson MJ (2021). Travelling waves in a free boundary mechanobiological model of an epithelial tissue. *Applied Mathematics Letters*. 111: 106636. doi: 10.1016/j.aml.2020.106636.

Contributor	Statement of Contribution
Ryan J. Murphy (Candidate)  02/09/2021	Conceived and designed the study, derived the continuum model, developed the codes for numerical simulation of the continuum models, performed numerical simulations, generated results, interpreted results, drafted the manuscript, and revised the manuscript during the peer-review process. Acted as corresponding author.
Pascal R. Buenzli	Supervised the research, and provided comments and gave final approval for publication. Supervised the research.
Ruth E. Baker	Provided comments and gave final approval for publication.
Matthew J. Simpson	Supervised the research, and provided comments and gave final approval for publication. Supervised the research.

#### Principal Supervisor Confirmation

I have sighted email or other correspondence from all Co-authors confirming their certifying authorship.

Matthew Simpson



2/9/2021

Name

Signature

Date

## 4.1 Abstract

We consider a free boundary model of epithelial cell migration with logistic growth and non-linear diffusion induced by mechanical interactions. Using numerical simulations, phase plane and perturbation analysis, we find and analyse travelling wave solutions with negative, zero, and positive wavespeeds. Unlike classical travelling wave solutions of reaction-diffusion equations, the travelling wave solutions that we explore have a well-defined front and are not associated with a heteroclinic orbit in the phase plane. We find leading order expressions for both the wavespeed and the density at the free boundary. Interestingly, whether the travelling wave solution invades or retreats depends only on whether the carrying capacity density corresponds to cells being in compression or extension.

## 4.2 Introduction

Nonlinear reaction-diffusion equations describing the dynamics of a single species often support travelling wave solutions [161, 201]. The classical example is the Fisher-KPP equation which has linear diffusion and a logistic reaction term [64, 141]. Travelling wave solutions of the Fisher-KPP equation are associated with heteroclinic orbits in the phase plane and correspond to invasion with a positive minimum non-dimensional wavespeed  $c \geq 2$  [96, 161]. Since the cell density  $q(x, t) \rightarrow 0$  as  $x \rightarrow \infty$ , these solutions do not have compact support and do not allow us to identify a well-defined front often observed in cell invasion experiments and ecological invasion [118, 140, 141, 212].

One way of overcoming the lack of a well-defined front is to incorporate degenerate nonlinear diffusion, as in the Porous-Fisher equation [194, 195, 200, 247]. An alternative approach to obtain travelling wave solutions with a well-defined front is to re-formulate the Fisher-KPP and Porous-Fisher models as moving boundary problems with a Stefan condition at the moving boundary [57, 58, 62, 63]. Interestingly the Fisher-KPP, Porous-Fisher, and Fisher-Stefan models always lead to invading travelling waves where previously vacant regions are eventually colonised. None of these single-species models lead to retreating travelling waves where colonised regions eventually become uncolonised. Similar invading behaviour has been observed in discrete space and velocity jump processes and their continuum approximations [139, 143]. Retreating and invading waves have previously been observed for multi-species models [116].

In this work we consider a single-species model which leads to travelling wave solutions with a well-defined front that can either invade or retreat. Our free boundary model, which we derived previously (Chapter 3, [19, 158, 164]), is motivated from a discrete model of a one-dimensional chain of epithelial cells. In this model cells are treated as mechanical springs that can be stretched or compressed and relax to a natural resting length. Cells are also able to proliferate logically up to a maximum carrying capacity density (Chapter 3, [19, 158, 164]). We find travelling wave solutions that are very different to the classical travelling waves of the Fisher-KPP, Porous-Fisher, or Fisher-Stefan models. We find travelling wave solutions for  $-\infty < c < \infty$  which depend on the two dimensionless parameters. In the phase plane these travelling waves are not associated with heteroclinic orbits. Instead, they are associated with an orbit that leaves a saddle equilibrium node until the trajectory passes through a special point in the phase plane determined by the free boundary conditions. We find and validate analytical expressions for both the wavespeed and the density at the free boundary. Interestingly, the distinction between whether the population retreats ( $c < 0$ ) or invades ( $c > 0$ ) depends only on

whether the carrying capacity density corresponds to cells being in compression or extension.

### 4.3 Mathematical model

We consider a one-dimensional chain of cells forming an epithelial sheet of total length  $L(t)$ . Each cell can be thought to act like a mechanical spring which mechanically relaxes towards its resting cell length,  $a$ , according to Hooke's law. Each cell can proliferate or die logically. Our previous work (Chapter 3, [19, 158]) shows this results in a moving boundary problem with nonlinear diffusivity, a logistic reaction term, and no-flux mechanical relaxation boundary conditions. After nondimensionalisation, the cell density,  $q(x, t) > 0$ , which depends on position  $x$  and time  $t$ , is governed by (Chapters 2 and 3, [19, 157, 158])

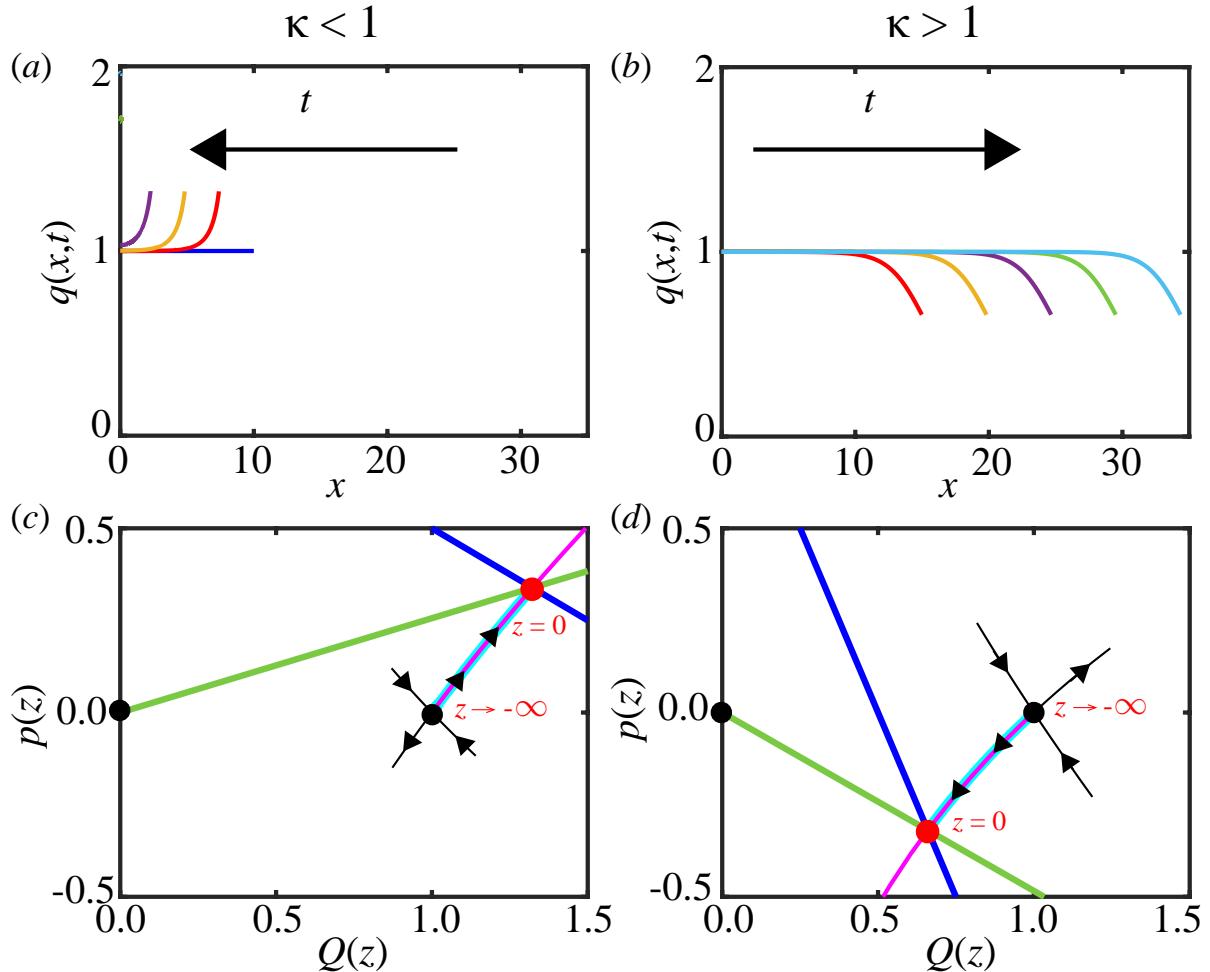
$$\frac{\partial q(x, t)}{\partial t} = \frac{\partial}{\partial x} \left( \frac{1}{q(x, t)^2} \frac{\partial q(x, t)}{\partial x} \right) + q(x, t) (1 - q(x, t)), \quad 0 < x < L(t), \quad (4.3.1)$$

$$\frac{\partial q(x, t)}{\partial x} = 0, \quad x = 0, \quad (4.3.2)$$

$$\frac{\partial q(x, t)}{\partial x} = \frac{q(x, t)^3}{\phi} \left( \frac{1}{q(x, t)} - \kappa \right), \quad x = L(t), \quad (4.3.3)$$

$$\frac{dL(t)}{dt} = -\frac{1}{q(x, t)^3} \frac{\partial q(x, t)}{\partial x}, \quad x = L(t), \quad (4.3.4)$$

with two dimensionless parameters  $\kappa$  and  $\phi$  occurring only in the free boundary condition at  $x = L(t)$  in Eq. (4.3.3). The first,  $\kappa = Ka$ , is the product of the carrying capacity density,  $K$ , and the resting cell length,  $a$ , and determines whether the carrying capacity density corresponds to cells being in compression ( $\kappa < 1$ ), at the resting length ( $\kappa = 1$ ), or in extension ( $\kappa > 1$ ). The second,  $\phi = \sqrt{\beta\eta/(4k)}$ , is the ratio of the proliferation rate,  $\beta$ , and mechanical relaxation rate, that depends on the cell stiffness  $k$  and mobility coefficient  $\eta$ . Eq. (4.3.4) governs the evolution of the free boundary due to mechanical relaxation and mass conservation but can be thought of as a nonlinear analogue of a Stefan condition [57, 58, 63]. Eqs. (4.3.1)–(4.3.4) can be solved numerically by using a boundary fixing transformation [120], discretising the subsequent equations on a uniform mesh using a central difference approximation. The resulting system of ordinary differential equations are solved using an implicit Euler approximation, leading to a system of nonlinear algebraic equations that are solved using Newton-Raphson iteration. Key code and algorithms are available on GitHub.



**Figure 4.1:** Travelling waves depend on  $\kappa$ :  $c < 0$  for  $\kappa = 0.5 < 1$ , and  $c > 0$  for  $\kappa = 2 > 1$ . (a)-(b) Density snapshots for varying  $\kappa$  at  $t = 0$  (blue), 10 (red), 20 (yellow), 30 (purple), 40 (green), 50 (cyan). (c)-(d)  $(Q, p)$  phase planes for varying  $\kappa$ . The travelling wave solution corresponds to a trajectory governed by Eqs. (4.4.2) (magenta) between the saddle node at  $(Q^*, p^*) = (1, 0)$  from Eq. (4.4.3) (black circle) and terminating at the intersection of Eq. (4.4.4) (blue) and (4.4.5) (green) given by Eq. (4.4.6) (red circle). Continuum solution from Eqs. (4.3.1)–(4.3.4) (cyan line). The degenerate node  $(Q^*, p^*) = (0, 0)$  is also shown (black circle). All results for  $\phi = 1$ .

## 4.4 Travelling waves

In Figure 4.1 we present numerical solutions of Eqs. (4.3.1)–(4.3.4) for varying  $\kappa$  and initial density condition  $q(x, 0) = 1$  for  $0 < x < L(0) = 10$ , which remains uniform and stationary, with  $c = 0$  for  $t > 0$ , when  $\kappa = 1$ . The numerical results in Figure 2(a) suggest the emergence of travelling wave solutions with  $c < 0$  when  $\kappa < 1$  (Figure 4.1(a)), with  $c = 0$  when  $\kappa = 1$  (not shown), and  $c > 0$  when  $\kappa > 1$  (Figure 4.1(b)). The travelling waves form after initial transient behaviour. For  $\kappa > 1$  the invading travelling waves in the numerical simulations continue as  $t \rightarrow \infty$ . For  $\kappa < 1$  we observe retreating travelling wave-like behaviour with  $c < 0$  for some intermediate time before  $L(t)$  approaches  $x = 0$  and boundary effects play a role (not shown). The cell density at the free boundary is  $Q_L > 0$ .

After a travelling wave has formed  $L(t) \sim ct$  (Figure S6), where  $c$  is the constant speed of propagation, and we introduce travelling wave coordinates  $z = x - ct$ . Letting  $Q(z) = q(x, t)$  then Eq. (4.3.1) becomes

$$\frac{d}{dz} \left( \frac{1}{Q(z)^2} \frac{dQ(z)}{dz} \right) + c \frac{dQ(z)}{dz} + Q(z)(1 - Q(z)) = 0, \quad -\infty < z < 0. \quad (4.4.1)$$

where we choose  $z = 0$  to correspond to the free boundary at  $x = L(t)$ .

To analyse Eq. (4.4.1) in the two dimensional phase plane we let  $p(z) = (1/Q(z)^2) dQ(z)/dz$  [132] to give

$$\frac{dQ}{dz} = pQ^2, \quad \frac{dp}{dz} = Q[-cpQ - (1 - Q)]. \quad (4.4.2)$$

The dynamical system given by Eqs. (4.4.2) has two equilibrium points. The first at  $(Q^*, p^*) = (0, 0)$  is a degenerate node. The second at  $(Q^*, p^*) = (1, 0)$  is a saddle node for  $c \neq 0$  and a degenerate node when  $c = 0$ . Interestingly, in contrast to the Fisher-KPP equation [161], here linear stability analysis provides no restrictions on  $c$ .

We return to the boundary conditions from Eqs. (4.3.2)–(4.3.4), and after transforming to travelling wave coordinates and writing in terms of  $p$ , we obtain

$$(Q, p) = (1, 0), \quad z \rightarrow -\infty, \quad (4.4.3)$$

$$p = \frac{1}{\phi}(1 - \kappa Q), \quad z = 0, \quad (4.4.4)$$

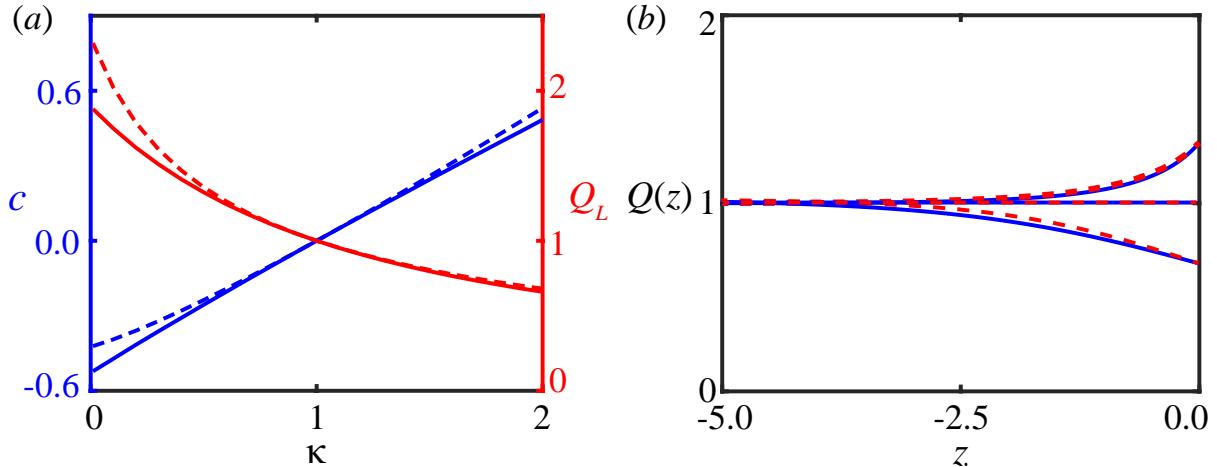
$$p = -cQ, \quad z = 0, \quad (4.4.5)$$

where Eq. (4.4.3) is informed by numerical travelling wave solutions in Figure 4.1.

In Figures 4.1(c),(d) we generate the  $(Q, p)$  phase plane for  $\kappa < 1$  and  $\kappa > 1$ , respectively, using MATLAB functions quiver and ode45 [145]. Trajectories corresponding to travelling wave solutions are initiated on the relevant eigenvector associated with the saddle node. We find that travelling wave solutions correspond to phase plane trajectories that run between  $(Q^*, p^*) = (1, 0)$ , and a special point given by the intersection of Eqs. (4.4.4) and (4.4.5) given by

$$(Q_L, p_L) = \left( \frac{1}{\kappa - c\phi}, \frac{-c}{\kappa - c\phi} \right). \quad (4.4.6)$$

The remainder of the trajectory beyond  $(Q_L, p_L)$ , obtained by solving Eqs. (4.4.2), corresponds to  $z > 0$  and is not associated with the travelling wave solution which is restricted to  $z \leq 0$ . The part of the trajectory with  $z > 0$  tends to infinity rather than to the degenerate equilibrium



**Figure 4.2:** Travelling wave perturbation analysis. (a) Properties of the travelling wave. Wavespeed  $c$  as a function of  $\kappa$  (blue) and density at free boundary  $Q_L$  as a function of  $\kappa$  (red). Solid lines: continuum model given by Eqs. (4.3.1)–(4.3.4). Dashed lines: leading order implicit solution given by Eq. (4.4.8). (b) Travelling wave solutions for  $\kappa = 0.5$  (top),  $\kappa = 1$  (middle),  $\kappa = 2$  (bottom) obtained by continuum model from Eqs. (4.3.1)–(4.3.4) (blue solid) and leading-order perturbation solution from Eq. (4.4.7) (red dashed). All for fixed  $\phi = 1$ .

point at  $(Q^*, p^*) = (0, 0)$ . Therefore, the travelling wave solution is not associated with a heteroclinic orbit. This is very interesting as classical travelling waves solutions are associated with heteroclinic orbits in the phase plane [161]–[62].

To provide insight into the travelling wave solutions in Figure 4.1 we now seek to determine a relationship between  $c$ ,  $\kappa$ , and  $\phi$ . By solving the continuum model we expect  $c \rightarrow 0$  as  $\kappa \rightarrow 1$  (Figure 4.2(a)). Therefore, we seek a perturbation solution  $p(Q) = p_0(Q) + cp_1(Q) + \mathcal{O}(c^2)$  for  $|c| \ll 1$  which we substitute into the equation for  $dp/dQ$  determined from Eqs. (4.4.2) to find

$$p_0(Q) = \pm \sqrt{2 [Q - \log_e(Q) - 1]}, \quad (4.4.7)$$

where the positive root corresponds to  $c < 0$  and the negative root corresponds to  $c > 0$ . The integration constant is chosen such that Eq. (4.4.7) satisfies Eq. (4.4.3). Eq. (4.4.7) corresponds to a small- $c$  approximation of the unstable manifold of the saddle point  $(1, 0)$ . Applying the free boundary condition from Eq. (4.4.5) and using  $Q_L$  from Eq. (4.4.6) gives

$$|c| = (\kappa - c\phi) \sqrt{2 \left[ \frac{1}{\kappa - c\phi} - \log_e \left( \frac{1}{\kappa - c\phi} \right) - 1 \right]}. \quad (4.4.8)$$

Eq. (4.4.8) can be solved implicitly for  $c$  as a function of  $\kappa$  and  $\phi$  and provides good agreement with the long time numerical solutions of Eqs. (4.3.1)–(4.3.4) (Figure 4.2). To find an approximate explicit form for  $c$  and  $Q_L$ , we expand Eq. (4.4.8) about  $\kappa - c\phi = 1$ , and use Eq. (4.4.6) to

give

$$c = \frac{\kappa - 1}{\phi + 1} + \mathcal{O}\left((\kappa - c\phi - 1)^{3/2}\right), \quad Q_L = \frac{1 + \phi}{\kappa + \phi}. \quad (4.4.9)$$

We find these leading order expressions in Eq. (4.4.9) to be accurate close to  $\phi = 1$  (Figures 4.1, 4.2, S1-S5).

In Figure 4.2b we plot the shape of the travelling wave obtained by considering long time numerical solutions of Eqs. (4.3.1)–(4.3.4) and compare this to the leading order perturbation solution. The leading order perturbation solution is obtained by solving Eq. (4.4.7) with the definition of  $p(z)$  in Eqs. (4.4.2) together with Eq. (4.4.6) as the initial condition. We observe excellent agreement for  $|c| \ll 1$  about  $\kappa = 1$ .

In summary, by considering a reaction-diffusion equation arising from a biologically motivated discrete model, we find an interesting result where whether a population invades or retreats corresponds to whether cells at the carrying capacity density are in compression or in extension, respectively. We also obtain exact expressions for the speed of travelling wave solutions of Eqs. (4.3.1)–(4.3.4), together with useful approximations of the shape of the travelling wave solutions when  $|c| \ll 1$ . We do not pursue an existence proof of these travelling wave solutions here, but leave this for future consideration [118].

## **Chapter 4A**

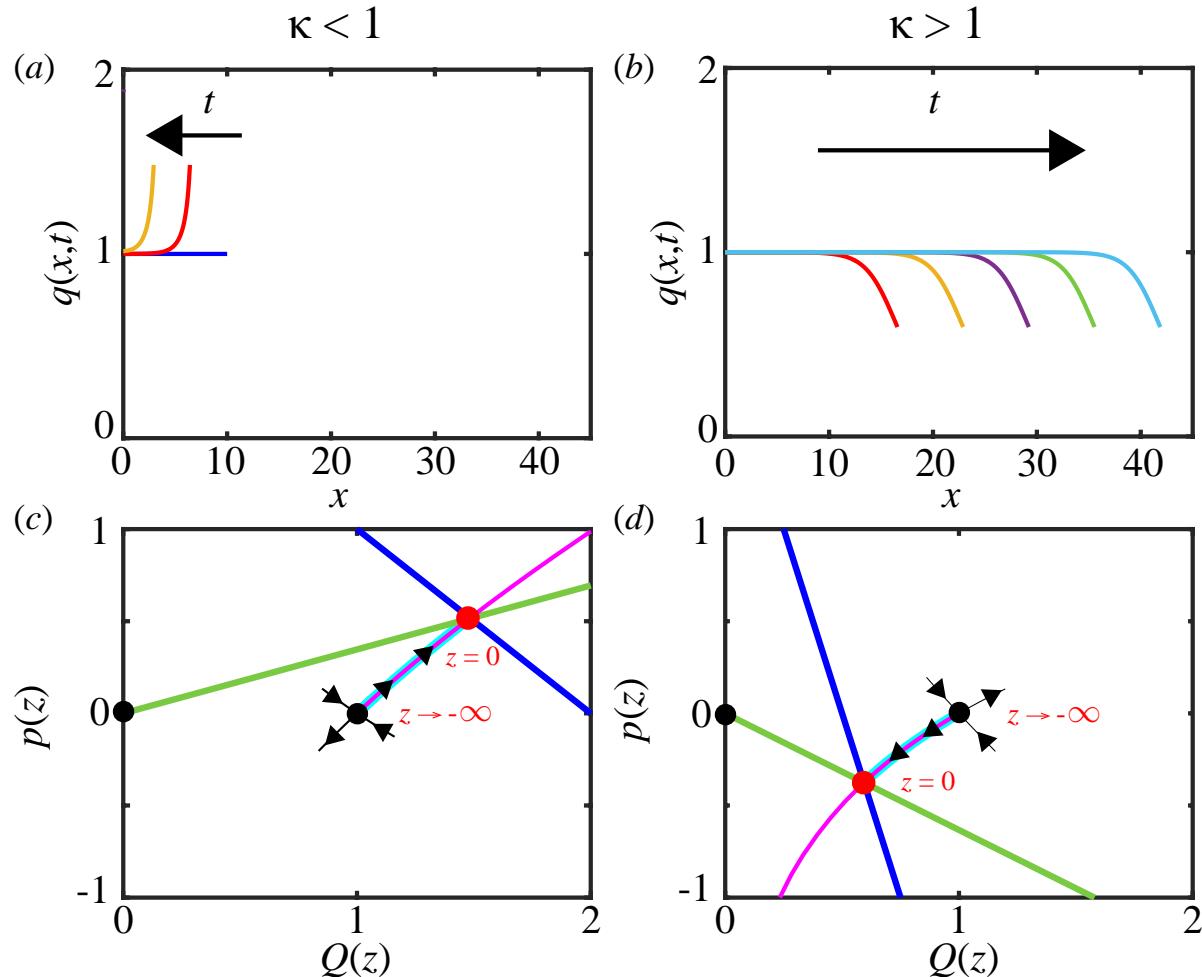
# **Supplementary Material for Chapter 4**

## S1 Additional Figures

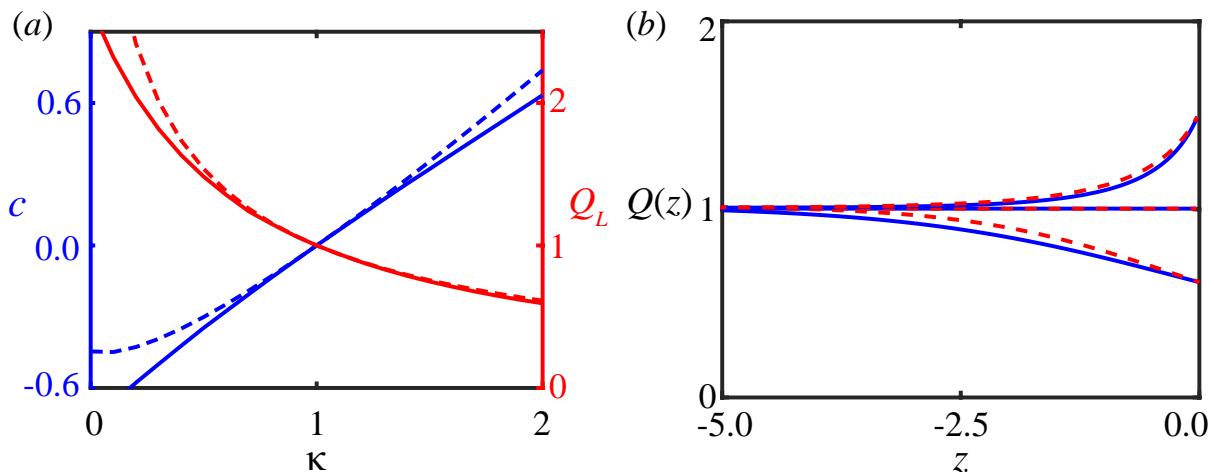
Results for Figure 1 and 2 in the manuscript are presented for  $\phi = 1$ . We now reproduce these figures for  $\phi = 0.5$  and  $\phi = 2$  in Figures S1, S2 and Figures S3, S4, respectively.

In Figure S5 we plot the dependence of the wavespeed,  $c$ , and density at the boundary,  $Q_L$ , against  $\phi$  for  $\kappa = 0.75$  and  $\kappa = 1.25$ .

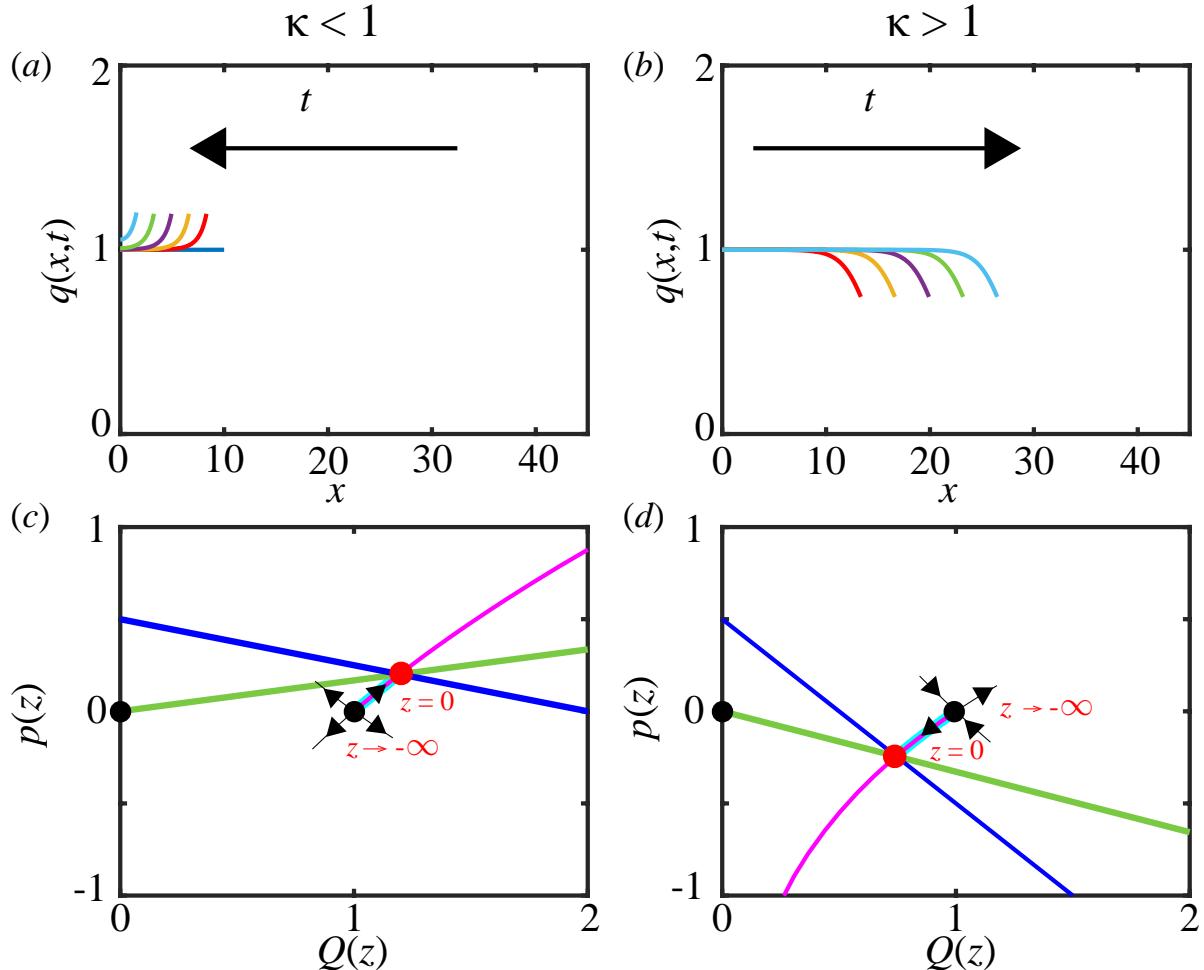
In Figure S6, we support the statement that after the travelling waves have formed  $L \sim ct$  by plotting  $L(t)$  against  $t$  for the results corresponding to Figure 1. Similar excellent agreement is found for other results (not shown).



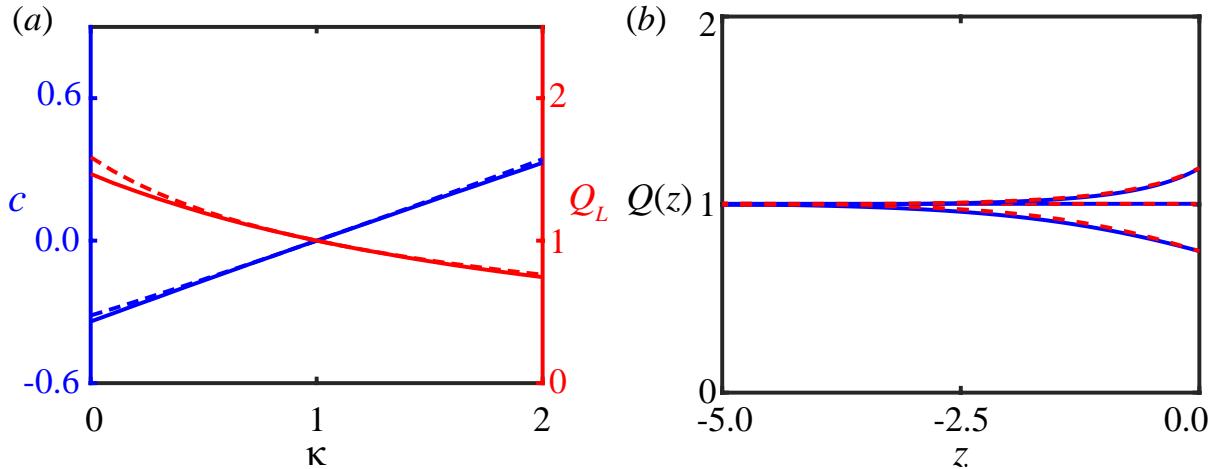
**Figure S1:** Results for  $\phi = 0.5$ . Travelling waves depend on  $\kappa$ :  $c < 0$  for  $\kappa = 0.5 < 1$ , and  $c > 0$  for  $\kappa = 2 > 1$ . (a)-(b) Density snapshots for varying  $\kappa$  at  $t = 0$  (blue), 10 (red), 20 (yellow), 30 (purple), 40 (green), 50 (cyan). (c)-(d)  $(Q, p)$  phase planes for varying  $\kappa$ . The travelling wave solution corresponds to a trajectory governed by Eqs. (6) (magenta) between the saddle node at  $(Q^*, p^*) = (1, 0)$  from Eq. (7) (black circle) and terminating at the intersection of Eq. (8) (blue) and 9) (green) given by Eq. (10) (red circle). Continuum solution from Eqs. (1)–(4) (cyan line). The degenerate node  $(Q^*, p^*) = (0, 0)$  is also shown (black circle).



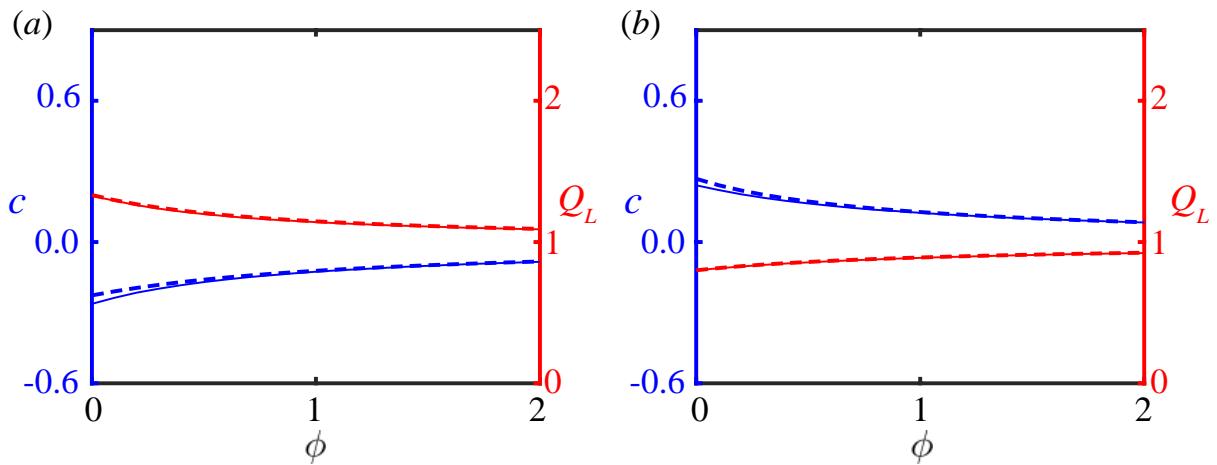
**Figure S2:** Travelling wave perturbation analysis for  $\phi = 0.5$ . (a) Properties of the travelling wave. Wavespeed  $c$  as a function of  $\kappa$  (blue) and density at free boundary  $Q_L$  as a function of  $\kappa$  (red). Solid lines: continuum model given by Eqs. (1)–(4). Dashed lines: leading order implicit solution given by Eq. (12). (b) Travelling wave solutions for  $\kappa = 0.5$  (top),  $\kappa = 1$  (middle),  $\kappa = 2$  (bottom) obtained by continuum model from Eqs. (1)–(4) (blue solid) and leading-order perturbation solution from Eq. (11) (red dashed).



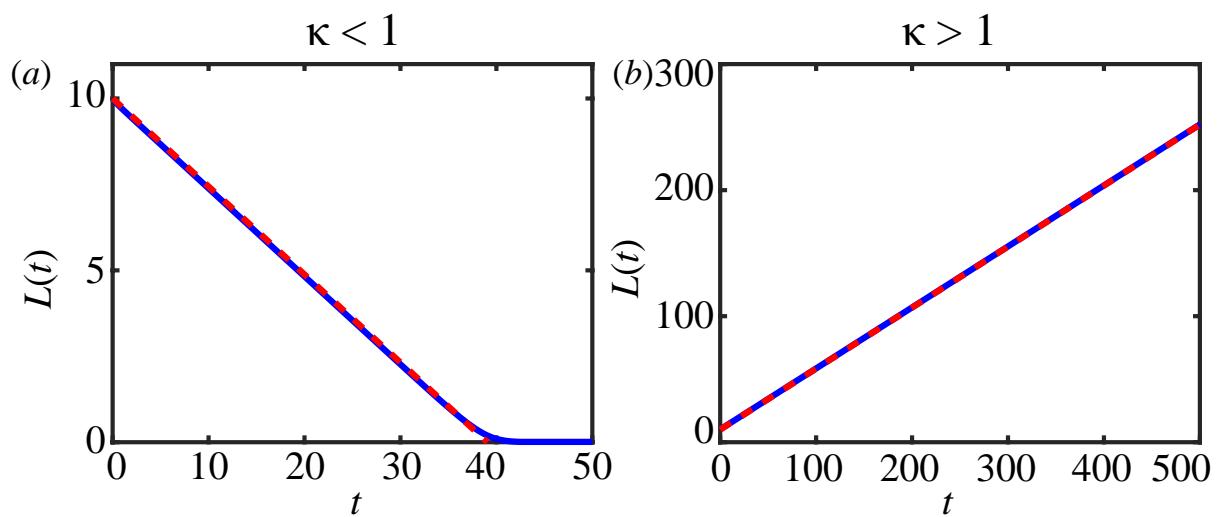
**Figure S3:** Results for  $\phi = 2$ . Travelling waves depend on  $\kappa$ :  $c < 0$  for  $\kappa = 0.5 < 1$ , and  $c > 0$  for  $\kappa = 2 > 1$ . (a)-(b) Density snapshots for varying  $\kappa$  at  $t = 0$  (blue), 10 (red), 20 (yellow), 30 (purple), 40 (green), 50 (cyan). (c)-(d)  $(Q, p)$  phase planes for varying  $\kappa$ . The travelling wave solution corresponds to a trajectory governed by Eqs. (6) (magenta) between the saddle node at  $(Q^*, p^*) = (1, 0)$  from Eq. (7) (black circle) and terminating at the intersection of Eq. (8) (blue) and (9) (green) given by Eq. (10) (red circle). Continuum solution from Eqs. (1)–(4) (cyan line). The degenerate node  $(Q^*, p^*) = (0, 0)$  is also shown (black circle).



**Figure S4:** Travelling wave perturbation analysis for  $\phi = 2$ . (a) Properties of the travelling wave. Wavespeed  $c$  as a function of  $\kappa$  (blue) and density at free boundary  $Q_L$  as a function of  $\kappa$  (red). Solid lines: continuum model given by Eqs. (1)–(4). Dashed lines: leading order implicit solution given by Eq. (12). (b) Travelling wave solutions for  $\kappa = 0.5$  (top),  $\kappa = 1$  (middle),  $\kappa = 2$  (bottom) obtained by continuum model from Eqs. (1)–(4) (blue solid) and leading-order perturbation solution from Eq. (11) (red dashed). All for fixed  $\phi = 2$ .



**Figure S5:** Travelling wave perturbation analysis. Properties of the travelling wave. Wavespeed  $c$  as a function of  $\phi$  (blue) and density at free boundary  $Q_L$  as a function of  $\phi$  (red) for (a)  $\kappa = 0.75$  and (b)  $\kappa = 1.25$ . Solid lines: continuum model given by Eqs. (1)–(4). Dashed lines: leading order implicit solution given by Eq. (12).



**Figure S6:** Evolution of tissue length  $L(t)$ . Comparison of continuum model given by Eqs. (1)–(4) (solid blue lines) with  $L(t) \sim ct$  (red dashed lines). (a)  $\kappa = 0.5 < 1$  and  $L(t) = 10 + ct$  where  $c = 0.484$ . (b)  $\kappa = 2 > 1$  and  $L(t) = 10 + ct$  where  $c = -0.256$ . Solutions correspond to Figure 1 for  $\phi = 1$ .

## **Chapter 5**

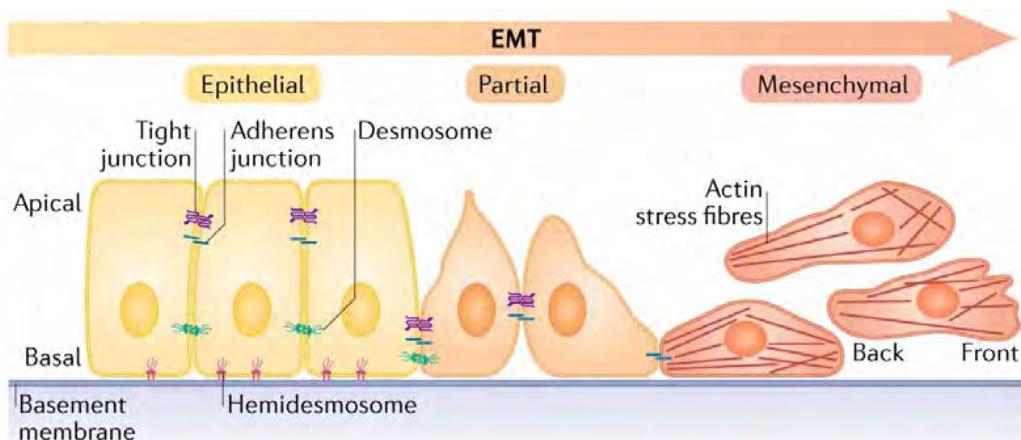
# **The role of mechanical interactions in epithelial-mesenchymal-transitions**

## 5.0 Preamble

An article published in *Physical Biology*

**Murphy RJ, Buenzli PR, Tambyah TA, Thompson EW, Hugo H, Baker RE, Simpson MJ (2021).**  
The role of mechanical interactions in epithelial mesenchymal transitions. *Physical Biology*.  
18:046001. doi:10.1088/1478-3975/abf425. bioRxiv preprint

This chapter includes Publication 4, addresses objective 4 and research question 4. This chapter extends the work of Chapters 2, 3, and 4, by incorporating diffusion of a chemical that influences the rate at which cells detach from the tissue boundary, in a process called epithelial-mesenchymal transition (EMT). As explored in earlier chapters, cells in epithelial tissues are characterised as moving collectively and being closely adherent. However, these epithelial cells can undergo phenotypic and morphological changes to partially or fully transition to mesenchymal cells, typically characterised as cells that are less adherent to other cells and tend to move as individuals, in a process called epithelial-mesenchymal transitions (EMT) (Figure 5.0, [110, 249]). The key results of this chapter are to extend the discrete model from previous chapters to describe the role of mechanical interactions in epithelial-mesenchymal transitions. To incorporate epithelial-mesenchymal transitions we extend the free boundary model in Chapter 4 by allowing cells to detach from the free end of the tissue dependent on a diffusing EMT-inducing chemical. Other key results include a derivation to obtain the corresponding continuum model, and the continuum model. Using this novel nonlinear free boundary problem we explore how mechanochemical coupling influences epithelial-mesenchymal-transitions. Supplementary material, such as numerical methods and additional results, associated with publication 4, is presented in Chapter 5A.



**Figure 5.0:** Schematic for epithelial-mesenchymal-transitions (EMT). Figure 1 from [55] reproduced with permission.

### Statement of Contribution of Co-Authors for Thesis by Published Paper

The authors listed below have certified that:

1. they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
2. they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
3. there are no other authors of the publication according to these criteria;
4. potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit, and
5. they agree to the use of the publication in the student's thesis and its publication on the QUT's ePrints site consistent with any limitations set by publisher requirements.

In this case of this chapter:

Murphy RJ, Buenzli PR, Tambyah TA, Thompson EW, Hugo H, Baker RE, Simpson MJ (2021). The role of mechanical interactions in epithelial-mesenchymal transitions. *Physical Biology*, 18:046001. doi:10.1088/1478-3975/abf425

Contributor	Statement of Contribution
Ryan J. Murphy (Candidate) <i>R. Murphy</i> 02/09/2021	Conceived and designed the study, derived the continuum model, developed the codes for numerical simulation of the discrete and continuum models, performed numerical simulations, generated results, interpreted results, drafted the manuscript, and revised the manuscript during the peer-review process. Acted as corresponding author.
Pascal R. Buenzli	Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research.
Tamara A. Tambyah	Provided comments and gave final approval for publication.
Erik W. Thompson	Provided advice on cancer and EMT. Provided comments and gave final approval for publication.
Honor Hugo	Provided advice on cancer and EMT. Provided comments and gave final approval for publication.
Ruth E. Baker	Provided comments and gave final approval for publication.
Matthew J. Simpson	Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research.

#### Principal Supervisor Confirmation

I have sighted email or other correspondence from all Co-authors confirming their certifying authorship.

Matthew Simpson

2/9/2021

Name	Signature	Date

## 5.1 Abstract

The detachment of cells from the boundary of an epithelial tissue and the subsequent invasion of these cells into surrounding tissues is important for cancer development and wound healing, and is strongly associated with the epithelial-mesenchymal transition (EMT). Chemical signals, such as TGF- $\beta$ , produced by surrounding tissue can be uptaken by cells and induce EMT. In this work, we present a novel cell-based discrete mathematical model of mechanical cellular relaxation, cell proliferation, and cell detachment driven by chemically-dependent EMT in an epithelial tissue. A continuum description of the model is then derived in the form of a novel nonlinear free boundary problem. Using the discrete and continuum models we explore how the coupling of chemical transport and mechanical interactions influences EMT, and postulate how this could be used to help control EMT in pathological situations.

## 5.2 Introduction

Cell detachment driven by epithelial-mesenchymal transitions (EMT) is crucial to many biological processes: embryonic development; later development in adults; wound healing; and cancer development [124, 229, 249]. During EMT, changes in gene expression and post-translational regulation mechanisms lead to increased invasive ability through the loss of epithelial characteristics and the acquisition of mesenchymal characteristics [249]. This transition is characterised by the loosening of cell-cell junctions and breakdown of the basement membrane [249]. EMT can be induced by chemical signals, such as TGF- $\beta$  [41], and is regulated by physical signals such as mechanical stress [77]. While EMT plays an important role in development, where it is a highly controlled and regulated process, EMT can be detrimental when initiated by cancer systems as it accelerates malignant progression and metastasis [178]. Furthermore, as 90% of cancer related deaths are associated to metastatic spread rather than cancer limited to a primary site [42], EMT is an important factor when considering therapy regimes [15, 23, 24, 142]. Previous theoretical models of EMT have largely neglected the role of mechanical interactions. Therefore, in this work we develop a novel mathematical model to explore how mechanical interactions between cells influence EMT and the evolution of a primary tumour. Using our model, we ask when do tumours grow or shrink, and how mechanical interactions and an EMT-inducing chemical could influence the rate of cell detachment. These insights could be used to help understand how to control EMT in pathological situations.

Mathematical models have proven to be a powerful tool to improve our understanding of EMT by providing a conceptual framework in which to integrate and analyse experimental data and make testable predictions, some of which have since been experimentally validated, for example, the existence of the epithelial/mesenchymal hybrid state [41, 110, 111, 137, 226, 229] and waves of temporal cell-cell detachments [12, 187]. Experimental and modelling studies are typically performed either at the single-cell level, by considering regulatory networks [137, 226, 229], or at the population-level, for example where cell populations are modelled with lattice based frameworks and the inclusion of cell-cell communication results in spatial heterogeneity [25]. However, existing models for EMT typically do not account for mechanical relaxation nor cell proliferation, both of which influence cell migration and cell size [25, 77]. These processes are thought to play a key role in cell-cell communication, tissue size, and the rate of cell detachment driven by EMT [25, 26]. Further, existing models typically do not connect descriptions of single-cell processes to population-scale behaviours.

In this work, we develop and explore a novel mathematical model of EMT which includes mechanical cellular relaxation, cell proliferation and cell detachment driven by chemical sig-

nals. We allow individual cells to detach from the tissue at the free boundary where the chemical concentration is highest [25]. This leads to a novel nonlinear free boundary problem. The evolution of epithelial monolayers and tumours have previously been modelled as free boundary problems [140, 198, 201]. However, many previous studies specify a classical one-phase Stefan condition [51] at the free boundary, where the rate of expansion of the free boundary is assumed proportional to the spatial gradient of the density without strong biological justification, or the evolution of the tissue length is specified according to experimental observations [48, 49, 203, 207–209]. Here, the evolution of the tissue boundary arises naturally from the cell-scale processes of cell proliferation, mechanical cellular relaxation (Chapter 4, [19, 159, 225]), and cell detachment.

To implement this model, we start with a cell-based discrete model, where we prescribe individual cell-level properties, and then derive the corresponding tissue-level continuum partial differential equation model. This approach extends previous studies (Chapters 2, 3, and 4, [19, 68, 157–159, 164, 165, 225, 256]) all of which consider mechanical cell movement, but do not consider cell detachment driven by EMT. The continuum model is useful to analyse possible behaviours of the model including tissue shrinkage, tissue homeostasis, and tissue growth depending on the initial number of cells, mechanical cell properties, the rate of proliferation, and chemical diffusivity. Importantly, we provide guidance when the discrete and continuum models are accurate. To simulate the mathematical model numerically, we devise a new method to incorporate chemical diffusion in an evolving population of cells with variable cells lengths.

This work is structured in the following way; we present the cell-based discrete model (Section 5.3.1) and derive the corresponding continuum model (Section 5.3.2). Then using the new discrete and continuum models we explore when tumours grow or shrink, and how mechanical interactions and an EMT-inducing chemical could influence the rate of cell detachment. To do so we consider different mechanisms in the models: cell-length-independent proliferation mechanism and chemically-independent cell detachment (Section 5.4.1); cell-length-dependent proliferation mechanism and chemically-independent cell detachment (Section 5.4.2); chemically-dependent cell detachment driven by an EMT-inducing chemical which diffuses slowly (Section 5.4.3), or which diffuses quickly (Section 5.4.4).

## 5.3 Model description

In this section, we present the new discrete model of mechanical cellular relaxation, cell proliferation, and cell detachment driven by a chemically-dependent EMT process in an epithelial tissue. We then derive the corresponding continuum model. To simplify the model, we suppose that linear diffusion is the key transport mechanism by which the chemical that induces EMT is transported from the free boundary inwards through the cells in the epithelial tissue (Figure 5.1).

### 5.3.1 Discrete model

We consider a one-dimensional chain of cells to represent the cross-section of an epithelial tissue (Figure 5.1). Each cell is assumed to act like a mechanical spring (Chapters 2, 3, and 4, [19, 157–159, 198, 225]). The tissue has a fixed boundary at  $x = 0$  and a free boundary at  $x = L(t) > 0$ . Cells undergo mechanical relaxation which results in changes in cell length and corresponding movements of cell boundaries. Cell  $i$ , for  $i = 1, 2, \dots, N(t)$ , occupies the interval  $x_i(t) < x < x_{i+1}(t)$ , where  $x_i(t)$  and  $x_{i+1}(t)$  are the positions of the left and right boundaries of the cell, respectively, so that cell  $i$  has length  $l_i(t) = x_{i+1}(t) - x_i(t)$ . Each cell is prescribed with cell stiffness  $k > 0$  and resting cell length  $a > 0$ . We assume that the motion of each cell boundary is subject to mechanical interactions and occurs in a viscous medium, resulting in a drag force with drag coefficient  $\eta > 0$  (Chapters 2, 3, and 4, [28, 67, 157, 198]). Then, as cells move in dissipative environments, the motion is assumed to be overdamped (Chapters 2, 3, and 4, [67, 157]) and the cell boundary evolution equations are

$$\frac{dx_1}{dt} = 0, \quad (5.3.1)$$

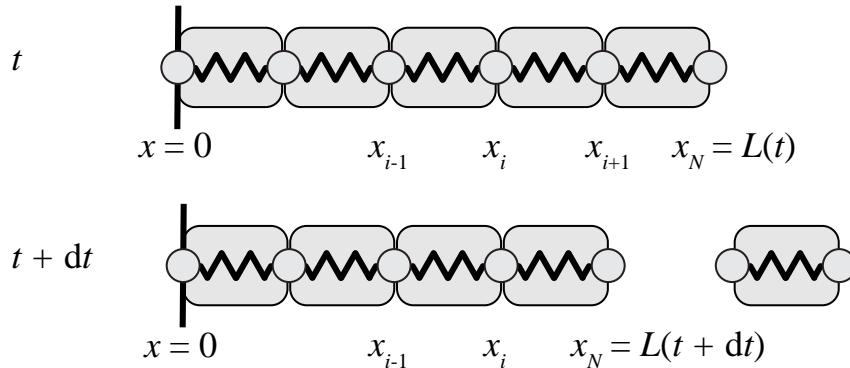
$$\eta \frac{dx_i}{dt} = f_{i+1}(t) - f_i(t) \quad i = 1, 2, \dots, N(t), \quad (5.3.2)$$

$$\eta \frac{dx_{N+1}}{dt} = -f_N(t), \quad (5.3.3)$$

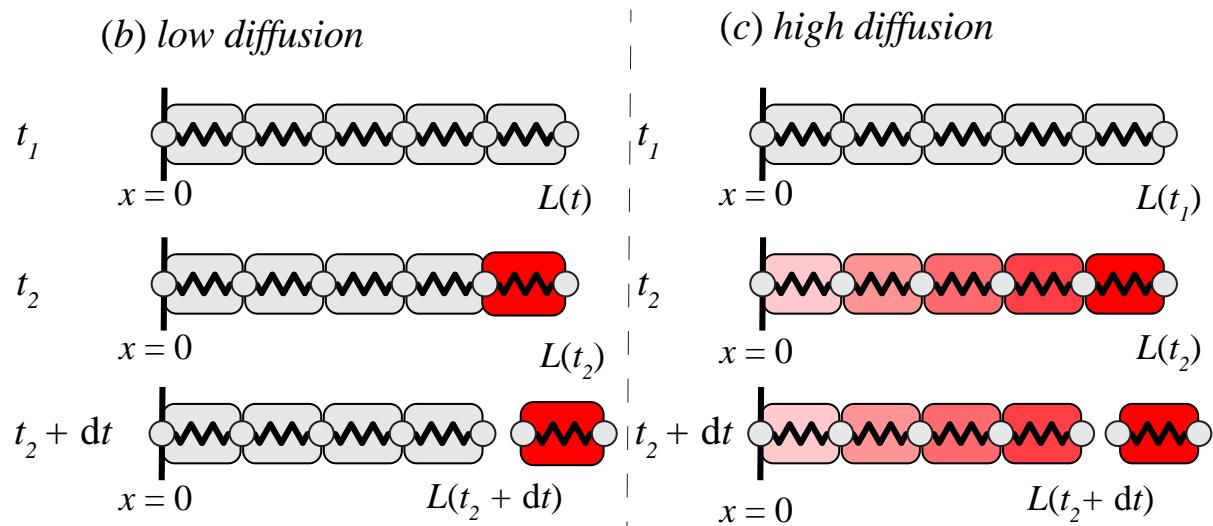
where, for simplicity, we use  $f_i(t) = k(l_i(t) - a)$  to represent a linear Hookean force law, and  $N(t)$  evolves in time due to proliferation and cell detachment.

We assume any cell in the tissue is able to proliferate via a stochastic cell proliferation mechanism that may depend on cell length. We assume that cell  $i$  proliferates with probability  $P(l_i)dt$  in the small time interval  $[t, t + dt]$ , where the function  $P(\cdot)$  depends on the proliferation mechanism considered, and  $l_i$  is the current cell length (Figure 5.2(b)). When a proliferation event occurs in a cell of length  $l_i$ , the cell divides into two cells of length  $l_i/2$ , and any chemical

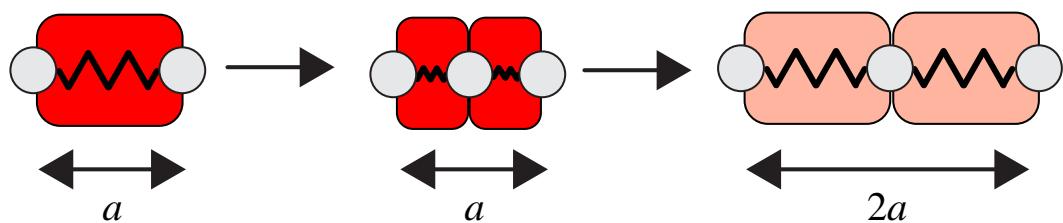
## (a) Cell detachment at a constant rate



## Cell detachment driven by chemically dependent EMT



## (d) Impact of proliferation on size and chemical concentration



**Figure 5.1:** Schematic for models of EMT and cell detachment. (a) Boundary cell detachment at a constant rate. (b)-(c) Boundary cell detachment driven by chemically-dependent EMT with (b) low diffusion and (c) high diffusion. The EMT-inducing chemical diffuses inwards from the external environment through the cell at the free boundary. In (b) with low diffusion the boundary cell contains most of the EMT-inducing chemical whereas in (c) with high diffusion the chemical is spread throughout the boundary and internal cells. Chemical concentration is shown with colouring: low concentration below the cell-detachment chemical threshold (light red) to a higher concentration above the cell-detachment chemical threshold (dark red). (d) Proliferation produces two identical daughter cells. Each daughter cell mechanically relaxes to the resting cell length and the concentration changes accordingly.

inside the cell is conserved and divided equally between the two daughter cells. The chemical concentration subsequently dilutes as the daughter cells mechanically relax to their resting cell lengths (Figure 5.1(d)).

We assume an external source of EMT-inducing chemical and suppose that linear diffusion is the key transport mechanism by which the chemical is transported into and through the epithelial tissue. To model linear diffusion we consider the chain of cells to be a non-uniform lattice on which we can simulate a point-jump process for molecules of the chemical. As it is computationally expensive to track many individual particles, we focus on the chemical concentration. In the epithelial tissue, each cell  $i$  has a chemical concentration  $c_i(t) = M_i(t)/l_i(t)$ , where  $M_i(t)$  is the number of molecules of the chemical in cell  $i$  at time  $t$ , and  $l_i(t)$  is the length of cell  $i$  at time  $t$ . Therefore, the concentration in each cell is assumed to be well-mixed. Then the equations governing chemical concentration are [252]

$$l_1 \frac{dc_1}{dt} = - \underbrace{c_1 \frac{dl_1}{dt}}_{\text{dilution}} + \underbrace{T_2^- c_2 l_2 - T_1^+ c_1 l_1}_{\text{diffusion}}, \quad (5.3.4)$$

$$l_i \frac{dc_i}{dt} = - \underbrace{c_i \frac{dl_i}{dt}}_{\text{dilution}} + \underbrace{T_{i-1}^+ c_{i-1} l_{i-1} - (T_i^+ + T_i^-) c_i l_i + T_{i+1}^- c_{i+1} l_{i+1}}_{\text{diffusion}}, \quad i = 2, 3, \dots, N-1, \quad (5.3.5)$$

$$l_N \frac{dc_N}{dt} = - \underbrace{c_N \frac{dl_N}{dt}}_{\text{dilution}} + \underbrace{T_{N-1}^+ c_{N-1} l_{N-1} - T_N^- c_N l_N}_{\text{diffusion}} + \underbrace{S}_{\text{source}}, \quad (5.3.6)$$

where  $T_i^\pm$  are left and right transition rates that model linear diffusion of chemical molecules between neighbouring cells with diffusivity  $D$ , respectively. How these transition rates are chosen requires great care and is detailed in Supplementary Material 5A.2 where we introduce a new method, called the *Interval-Voronoi* method. The dilution term in Equations (5.3.4)-(5.3.6) represents the fact that chemical concentrations increase/decrease as the cell length reduces/increases. To mimic a chemical diffusing into the tissue from an external source, we assume that a constant number of molecules per unit time,  $S$ , is provided to cell  $N$  from the external source. We further assume that the chemical cannot diffuse across the boundary at the rear of the tissue ( $x = 0$ ). This assumption corresponds to modelling only the right hand side of an epithelial tissue where  $x = 0$  is the middle of the tissue.

Given the chemical concentrations in each cell, we introduce cell detachment driven by EMT which is a key feature of the discrete mechanical model. Here, we consider cell detachment to be a two-step process. The first step models the EMT process itself as a cell-state transition whereby cells acquire an invasive phenotype. The boundary cell gains the invasive phenotype when its chemical concentration is above a constant threshold,  $C$  (Figure

5.2) [41, 137, 226]. If at any time the chemical concentration drops below the threshold the cell loses its invasive phenotype, which it can only regain once the chemical concentration increases above  $C$ . The ability of the cell to gain and lose its invasive phenotype is associated with epithelial-mesenchymal plasticity [110]. The second part of the process is where the boundary cell, once it acquires the invasive phenotype, detaches from the tissue [134] in  $[t, t + dt)$  with probability  $\omega(c_N(t)) dt$  where  $c_N(t)$  is the chemical concentration in the boundary cell  $N$  at time  $t$ . Once a cell detaches we no longer consider its dynamics and we assume it moves away from the epithelial tissue.

As we are interested in whether tumours grow or shrink, we can consider the evolution of the total number of cells,  $N(t)$ , which depends on the balance between proliferation and EMT. For an individual realisation of this discrete model,  $N(t)$  is expected to increase when

$$\sum_{i=1}^{N(t)} P(l_i) > \omega(c_N). \quad (5.3.7)$$

We numerically simulate the discrete model governed by Equations (5.3.1)–(5.3.6), and prescribe initial conditions for the cell positions,  $x_i(0)$ , the mechanical cell properties  $k$  and  $a$ , drag coefficient  $\eta$ , as well as proliferation properties, and assume that there is initially no chemical inside any cell in the tissue (Supplementary Material 5A.2).

### 5.3.2 Continuum model

We now derive the corresponding free boundary continuum model for cell detachment driven by chemically-dependent EMT. Components of this model have been derived in our previous studies, and where this is the case we state the equation and provide a reference to the reader for full details (Chapters 2 and 3, [19, 157, 158, 225]).

The cell density,  $q(x, t) > 0$ , which is the number of cells per unit length and the continuous analogue of  $1/l_i$ , evolves according to the following nonlinear moving boundary problem (Chapter 4, [158])

$$\frac{\partial q(x, t)}{\partial t} = \underbrace{-\frac{1}{\eta} \frac{\partial^2 f(x, t)}{\partial x^2}}_{\text{mechanical relaxation}} + \underbrace{q(x, t)P\left(\frac{1}{q(x, t)}\right)}_{\text{proliferation}}, \quad 0 < x < L(t), \quad (5.3.8)$$

where  $f(x, t) = k(1/q(x, t) - a)$  is the continuous analogue of the discrete Hookean force law and  $P(1/q(x, t))$  is the continuous analogue of  $P(l_i)$ .

The fixed boundary condition at  $x = 0$ , corresponding to Equation (5.3.1), is (Chapter 2 and

3, [157])

$$\frac{\partial q(0, t)}{\partial x} = 0, \quad (5.3.9)$$

and the mechanical relaxation condition at the free boundary,  $x = L(t)$ , gives rise to a nonlinear boundary condition [19, 225]

$$\frac{1}{2q(L(t), t)} \frac{\partial f(L(t), t)}{\partial x} = -f(L(t), t). \quad (5.3.10)$$

The boundary conditions in Equations (5.3.9) and (5.3.10) ensure that no cells are lost by crossing the tissue boundaries but cells can still detach at the free boundary,  $x = L(t)$ . In the continuum model this corresponds to loss of tissue material at a moving interface [34]. To capture cells detaching from the tissue at  $x = L(t)$ , we consider conservation of mass and derive the following evolution equation for the free boundary (Supplementary Material 5A.1)

$$\frac{dL(t)}{dt} = \underbrace{\frac{1}{\eta q(L(t), t)} \frac{\partial f(L(t), t)}{\partial x}}_{\text{mechanical relaxation}} - \underbrace{\frac{\omega(c(L(t), t))}{q(L(t), t)}}_{\text{cells lost due to detachment}}. \quad (5.3.11)$$

The chemical concentration is governed by the following advection-diffusion equation [20, 48, 49, 225]

$$\frac{\partial c(x, t)}{\partial t} + \frac{\partial}{\partial x} (u(x, t)c(x, t)) = D \frac{\partial^2 c(x, t)}{\partial x^2}, \quad 0 < x < L(t), \quad (5.3.12)$$

where  $D$  is the diffusion coefficient, and the cell velocity,  $u(x, t)$ , determined from Equation (5.3.8), is

$$u(x, t) = \frac{1}{\eta q(x, t)} \frac{\partial f(x, t)}{\partial x}. \quad (5.3.13)$$

The boundary at  $x = 0$  is fixed, and because there is no chemical transport across this boundary in the discrete model, we impose the following no-flux boundary condition

$$\frac{\partial c(0, t)}{\partial x} = 0. \quad (5.3.14)$$

At the free boundary,  $x = L(t)$ , the only transport of chemical in the discrete model is the supply of a constant number of molecules per unit time,  $S$ , from the external environment. The

corresponding boundary condition at  $x = L(t)$  in the continuum model is

$$D \frac{\partial c(L(t), t)}{\partial x} = S, \quad (5.3.15)$$

obtained by enforcing that the total flux of  $c(x, t)$  in the frame of reference co-moving with  $L(t)$  is equal to  $-S$  [225].

We supplement the continuum model with initial conditions for tissue length,  $L(0)$ , density  $q(x, 0)$  for  $0 < x < L(0)$ , and chemical concentration  $c(x, 0)$  for  $0 < x < L(0)$ . Then Equations (5.3.8)–(5.3.10), (5.3.11), and (5.3.12)–(5.3.15) are solved numerically using a boundary fixing transformation [120] and an implicit finite difference approximation, see Supplementary Material 5A.3 for further details.

## 5.4 Results and discussion

The evolution of the number of cells in the epithelial tissue,  $N(t)$ , the length of the epithelial tissue,  $L(t)$ , and the number of cells that undergo EMT and detach from the epithelial tissue are coupled. Biologically this coupling is of great interest. In the context of cancer, a primary tumour site without EMT is a localised problem, whereas a single tumour site with EMT and cell detachments may result in many secondary tumour sites that can be a greater problem as 90% of cancer related deaths are associated with metastatic spread [42]. We are interested in how mechanical interactions influence EMT and subsequently how tumours grow or shrink. Therefore, we choose parameters to explore the range of behaviours that our new mathematical model of EMT predicts. Further, as the continuum model is useful to analyse possible behaviours, we seek to understand when the continuum model is a good description of the underlying discrete model by considering initial populations with low numbers of cells. Our parameter choices are also consistent with experimental observations that can vary greatly depending on the cell type and driving mechanisms: a cell can proliferate on the order of once every 12 hours to once every few days; EMT can occur over the course of hours, a few days [6], or many days (e.g. 9–12 days [112]); the rate of mechanical relaxation is faster than the rate of proliferation and EMT, with  $\eta/k \approx 5\text{--}16$  minutes being a typical experimental value [28]; and a typical experimental value for the resting cell length being  $a \approx 10 \mu\text{m}$  [68].

### 5.4.1 Cell-length-independent proliferation and chemically-independent EMT

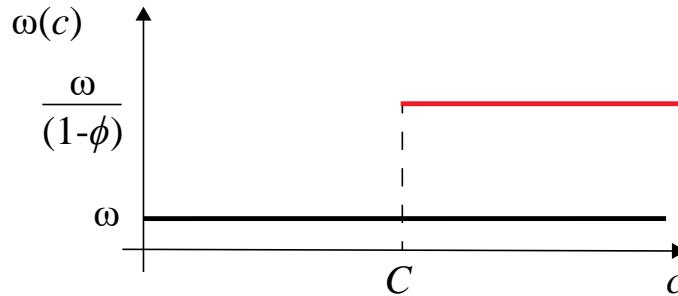
The simplest model is chemically-independent cell detachment of the boundary cell at a constant rate,  $\omega$ , with cell-length-independent proliferation for each cell at a constant rate,  $\beta$  (Figure 5.2(a),(b)). It is useful to first examine this problem with the continuum model. Conservation of mass (Equation (5A.1.1)) gives a simple ordinary differential equation for the evolution of  $N(t)$ ,

$$\frac{dN(t)}{dt} = \beta N(t) - \omega, \text{ with solution } N(t) = \frac{\omega}{\beta} + \left( N(0) - \frac{\omega}{\beta} \right) \exp(\beta t). \quad (5.4.1)$$

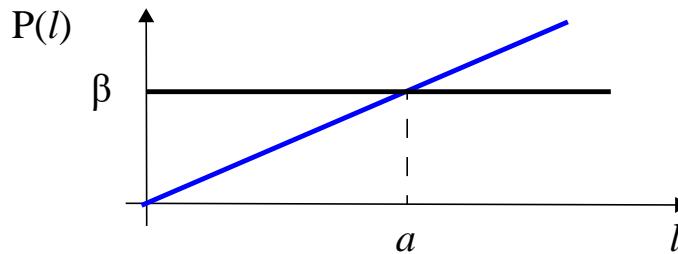
Then, depending on the initial number of cells,  $N(0)$ , and the critical cell number,  $\omega/\beta$ , there are three possible long-term outcomes: i)  $N(t) \rightarrow 0$  in finite time, which we refer to as extinction, when  $N(0) < \omega/\beta$ ; ii)  $N(t)$  remains constant for all  $t$  when  $N(0) = \omega/\beta$ ; iii)  $N(t) \rightarrow \infty$  as  $t \rightarrow \infty$  when  $N(0) > \omega/\beta$ . It is clear from Equation (5.4.1) that in all cases mechanical interactions do not influence  $N(t)$ . However, each cell mechanically relaxes towards its equilibrium length as  $N(t)$  evolves over time. To capture the evolution of the tissue length,  $L(t)$ , and cell density,  $q(x, t)$  we solve the full continuum model, governed by Equations (5.3.8)–(5.3.10) and

(5.3.11) (Figure 5.3). The total number of cells which detach grows linearly with time at rate  $\omega$  for chemically-independent EMT when the tissue is not close to extinction, and plateaus if extinction occurs (Figure 5A.4).

(a) Cell detachment rate dependent on chemical concentration,  $c$ .



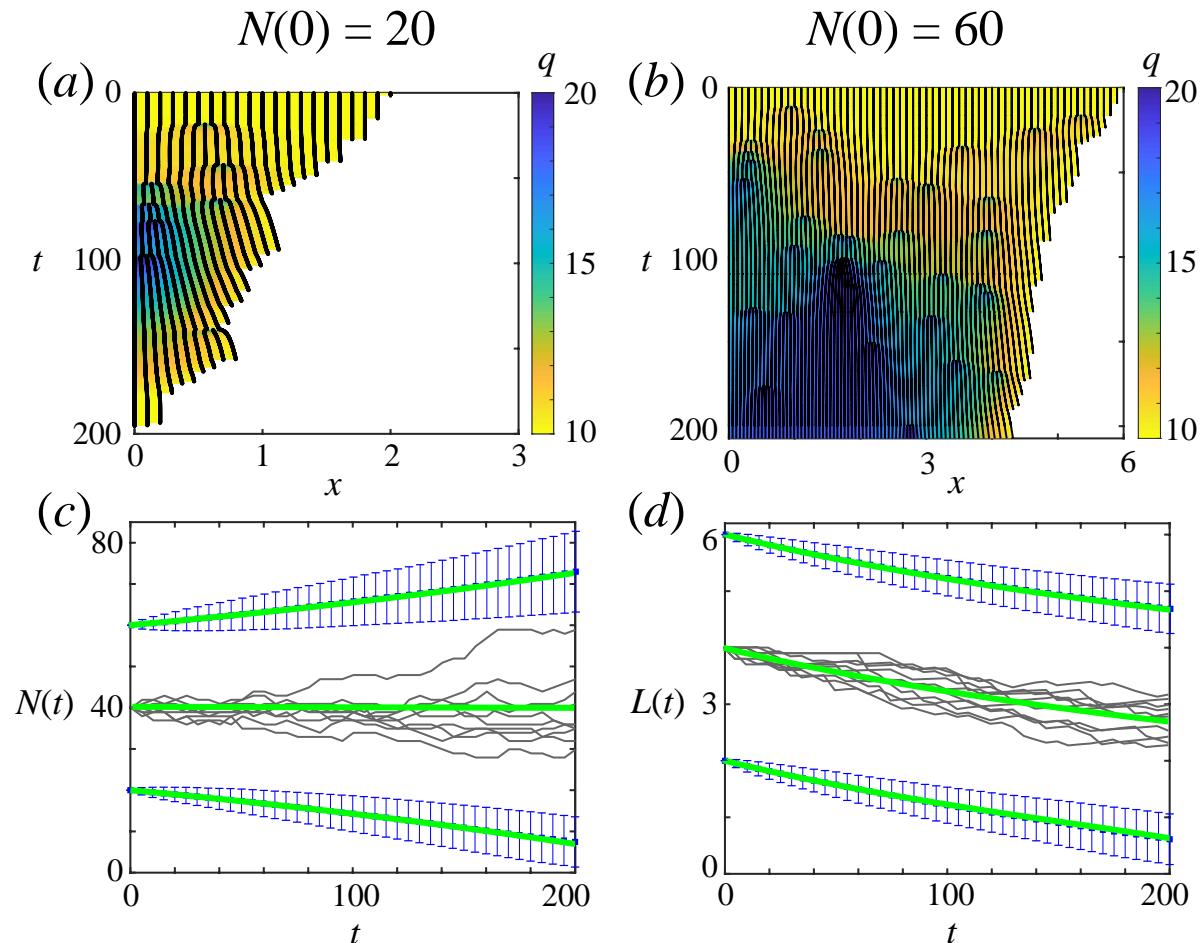
(b) Proliferation rate dependent on cell length,  $l$ .



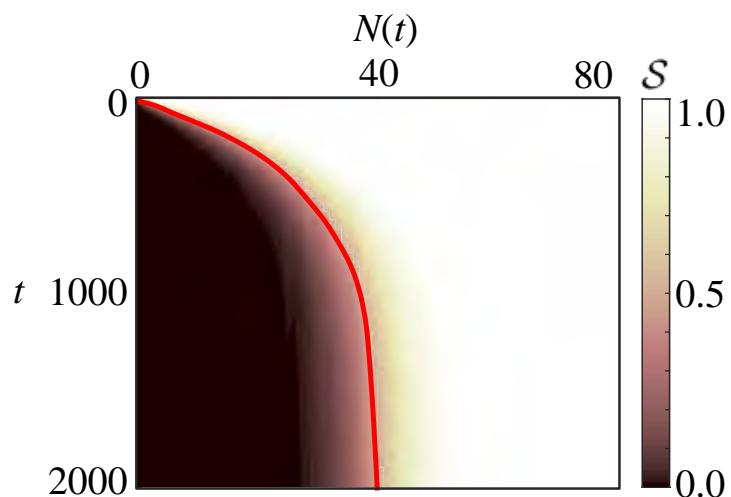
**Figure 5.2:** Cell detachment driven by chemically-dependent EMT and proliferation mechanisms. (a) Cell detachment mechanisms: cell detachment at a constant rate  $\omega$  independent of chemical concentration (black); cell detachment at a constant rate  $\omega/(1 - \phi)$  when the concentration,  $c$ , is above a concentration threshold,  $C$  (red). (b) Cell-length-dependent proliferation mechanisms: independent of cell length at rate  $\beta$  (black); linearly dependent on cell-length (blue) defined by ensuring  $P(0) = 0$  and  $P(a) = \beta$ . In this main manuscript we set  $\omega = 0.1$ ,  $\phi = 0.9$ ,  $C = 500$  and  $\beta = 0.0025$ , and vary  $\phi$  in supplementary material.

In general, the solution of the continuum model provides an accurate approximation for the evolution of  $N(t)$ ,  $L(t)$  and  $q(x, t)$  when compared to appropriately averaged quantities from many discrete realisations (Figures 5.3, 5A.5-5A.8). This correspondence between the discrete and continuum model holds provided that the rate of mechanical relaxation, determined by the ratio of cell stiffness to drag coefficient,  $k/\eta$ , is sufficiently fast relative to the rate of cell proliferation, see Section 3.3 of Chapter 3 [158]; and that  $N(t)$  is sufficiently large to define a continuum, see Supplementary Material 5A.4.4. However, when  $N(0)$  is close to  $\omega/\beta$  the behaviour of the discrete and the continuum model may differ. Extinction behaviour of the continuum model is deterministic and solely determined by  $N(0)$ , whereas stochastic effects in the discrete model imply that different realisations for the same  $N(0)$  may sometimes result in extinction and sometimes in unbounded growth (Figures 5.3(c), Equation (5.3.7)) [220]. To quantify this difference between the models, we simulate many identically-prepared realisations of the discrete model and calculate the survival probability of the tissue: the probability

that an individual realisation is not extinct at a certain time. By comparing the survival probability of the tissue from the discrete and continuum models for a range of  $N(0)$ , Figure 5.4 shows that when  $N(0)$  is close to  $\omega/\beta$  and when  $N(0)$  is close to extinction, results from the continuum model may not reflect the behaviour of individual discrete realisations.



**Figure 5.3:** Chemically-independent cell detachment at a constant rate with a cell-length-independent proliferation mechanism. (a)-(b) Kymographs showing evolution of cell boundaries (black curves, note bifurcations when cells divide) for one discrete realisation with cell density,  $q(x, t)$ , colouring for (a)  $N(0) = 20$  and (b)  $N(0) = 60$ . Note domain size in (b) is double that of (a). (c)-(d) Three initial cell populations starting at mechanical equilibrium,  $L(0) = N(0)a$ . For  $N(0) = 20, 60$  the average of 2000 discrete realisations (blue) are compared with the continuum model (green). For  $N(0) = \omega/\beta = 40$ , 10 discrete realisations (grey) are compared with the continuum model (green). (c) Evolution of number of cells,  $N(t)$ . (d) Evolution of tissue length,  $L(t)$ . Mechanical parameters:  $k = 1$ ,  $a = 0.1$ ,  $\eta = 1$ .

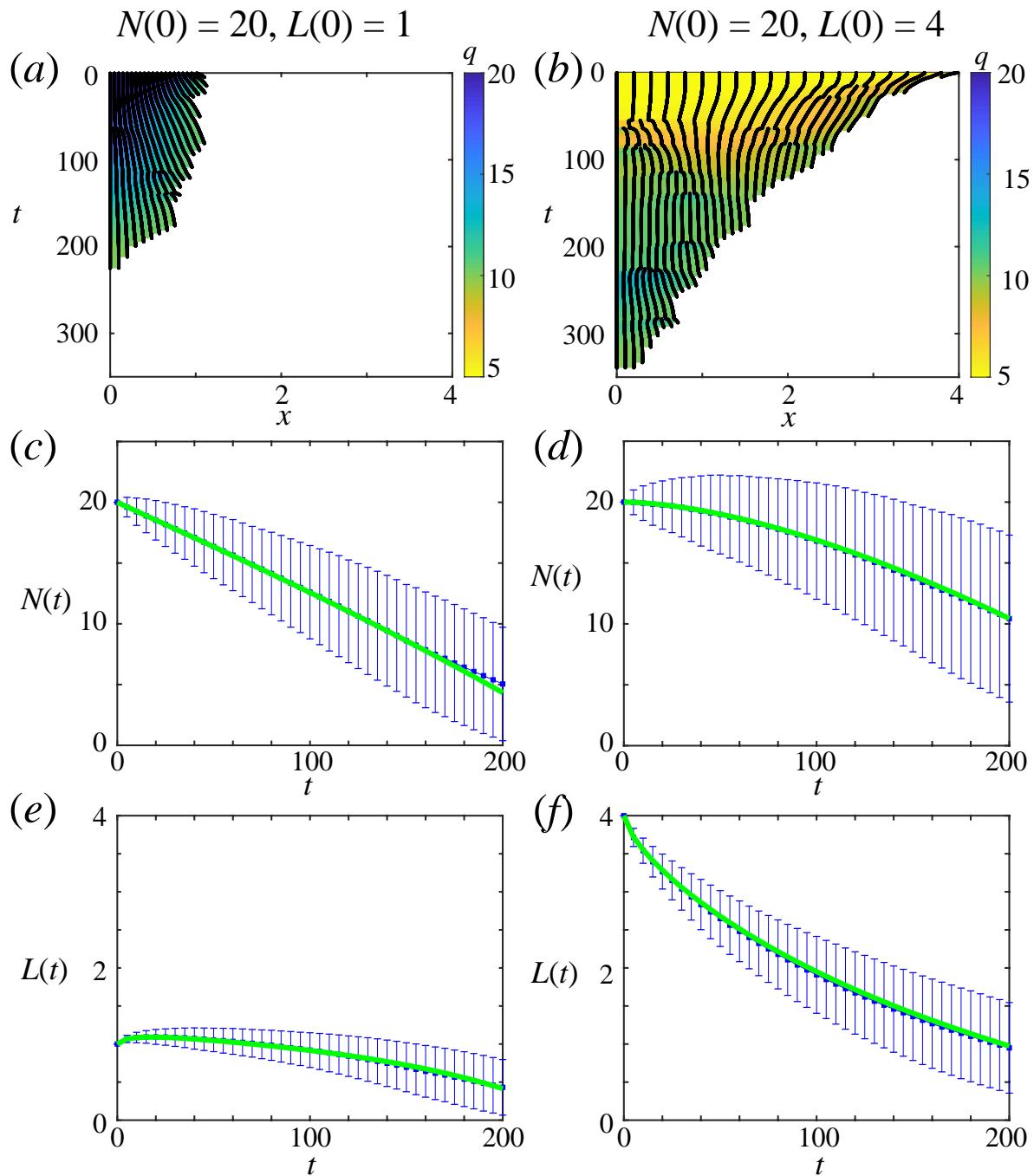


**Figure 5.4:** Survival probability of the tissue,  $S$ , for the cell-length-independent proliferation mechanism and chemically-independent cell detachment with  $N(0) = 1, 2, \dots, 80$ . Comparison between the deterministic continuum model (red solid line), and the average of 2000 realisations of the stochastic discrete model (shading). Here,  $w/\beta = 40$ .

### 5.4.2 Cell-length-dependent proliferation and chemically-independent EMT

With a constant cell-length-independent proliferation rate,  $N(0)$  determines the long-term solution of the continuum model, whereas with cell-length-dependent proliferation we must also consider the initial cell lengths,  $l_i(0)$ , resting cell length,  $a$ , and the ratio of cell stiffness to drag coefficient,  $k/\eta$ . If we consider a linear proliferation mechanism  $P(l_i) = \beta l_i/a$ , shown in Figure 5.2(b), then the critical tissue length is  $\omega a/\beta$  (Equation (5.3.7)). Therefore in this case mechanical interactions between the cells are important. Parameter combinations that lead to extinction with cell-length-independent proliferation may now grow without bound with cell-length-dependent proliferation (Section 5A.4.3). Similarly, parameter combinations that lead to unbounded growth with cell-length-independent proliferation may now lead to extinction with cell-length-dependent proliferation. Further, the model predicts that compressed tissues can go extinct faster (Figure 5.5(a),(c),(e)) than stretched tissues (Figure 5.5(b),(d),(f)). Good agreement between the continuum model and appropriately averaged quantities from many discrete realisations is also observed when considering cell-length-dependent proliferation (Supplementary Material 5A.4.7).

When we consider cell-length-dependent proliferation, the long term outcome of the model depends upon the mechanical properties,  $k/\eta$ , and rate of cell proliferation and detachment. In general, when  $k/\eta$  is large compared to  $\beta$ , the outcome of the model is similar to the simpler cell-length-independent proliferation case. As before, the solution of the continuum model is a good approximation of appropriately averaged data from the discrete model, except when  $N(t)$  is low (Figure 5.5). However, whereas for cell-length-independent proliferation stochastic effects are important when the initial number of cells  $N(0)$  is close to  $\omega/\beta$ , for cell-length-dependent proliferation stochastic effects are important whenever the current tissue length,  $L(t)$ , is close to the critical tissue length,  $\omega a/\beta$ , as the epithelial tissue may go extinct in some realisations while the epithelial tissue may grow in other realisations (Supplementary Material 5A.4.3). As we are considering chemically-independent EMT the total number of cells which detach grows linearly with time at rate  $\omega$  when the tissue is not close to extinction (Figure 5A.4).



**Figure 5.5:** Chemically-independent cell detachment at a constant rate with a linear cell-length-dependent proliferation mechanism. Two initial cell populations with  $N(0) = 20$ , the first uniformly compressed with  $L(0) = 1$  and the second uniformly stretched with  $L(0) = 4$ . (a)-(b) Kymographs with density,  $q(x, t)$ , colouring. (c)-(f) The average of 2000 discrete realisations (blue) are compared with the continuum model (green). (c)-(d) Evolution of total cell number,  $N(t)$ . (e)-(f) Evolution of tissue length,  $L(t)$ . Mechanical parameters:  $k = 1, a = 0.1, \eta = 1$ . Critical tissue length is  $\omega a / \beta = 4$ .

### 5.4.3 Chemically-dependent EMT with small diffusivity

We now consider a general EMT-inducing chemical, with TGF- $\beta$  being one such example of many candidate signalling molecules. As different EMT-inducing chemicals may have different diffusivities, we will consider a range of possible diffusivities from a very small diffusivity to assuming the chemical in the tissue is in diffusive equilibrium at all times. To begin, we assume very small diffusivity and cell-length-independent proliferation. Assuming that the tissue initially contains no chemical, the chemical is provided to the boundary cell only, and small diffusivity, then the chemical is mostly concentrated in the boundary cell. To compare the chemically-dependent cell detachment model with the chemically-independent cell detachment model (Section 5.4.1) we choose parameters so that the average rate of cell detachment is the same in both models, provided the boundary cell is close to its resting cell length. Chemically-dependent cell detachment is a two-step process: i) the boundary cell gains an invasive phenotype when its chemical concentration is above the chemical threshold,  $C$ , ii) the boundary cell detaches. So we introduce a parameter  $\phi \in [0, 1]$  which defines the ratio of the average time in process i) as  $\phi/\omega$  and the average time in process ii) as  $(1 - \phi)/\omega$  (Figure 5.2). Note that  $\phi = 0$  corresponds to the chemically-independent model we explore in Sections 5.4.1-5.4.2. As before, the total number of cells which detach grows linearly with time at rate  $\omega$  when the tissue is not close to extinction (Figure 5A.4).

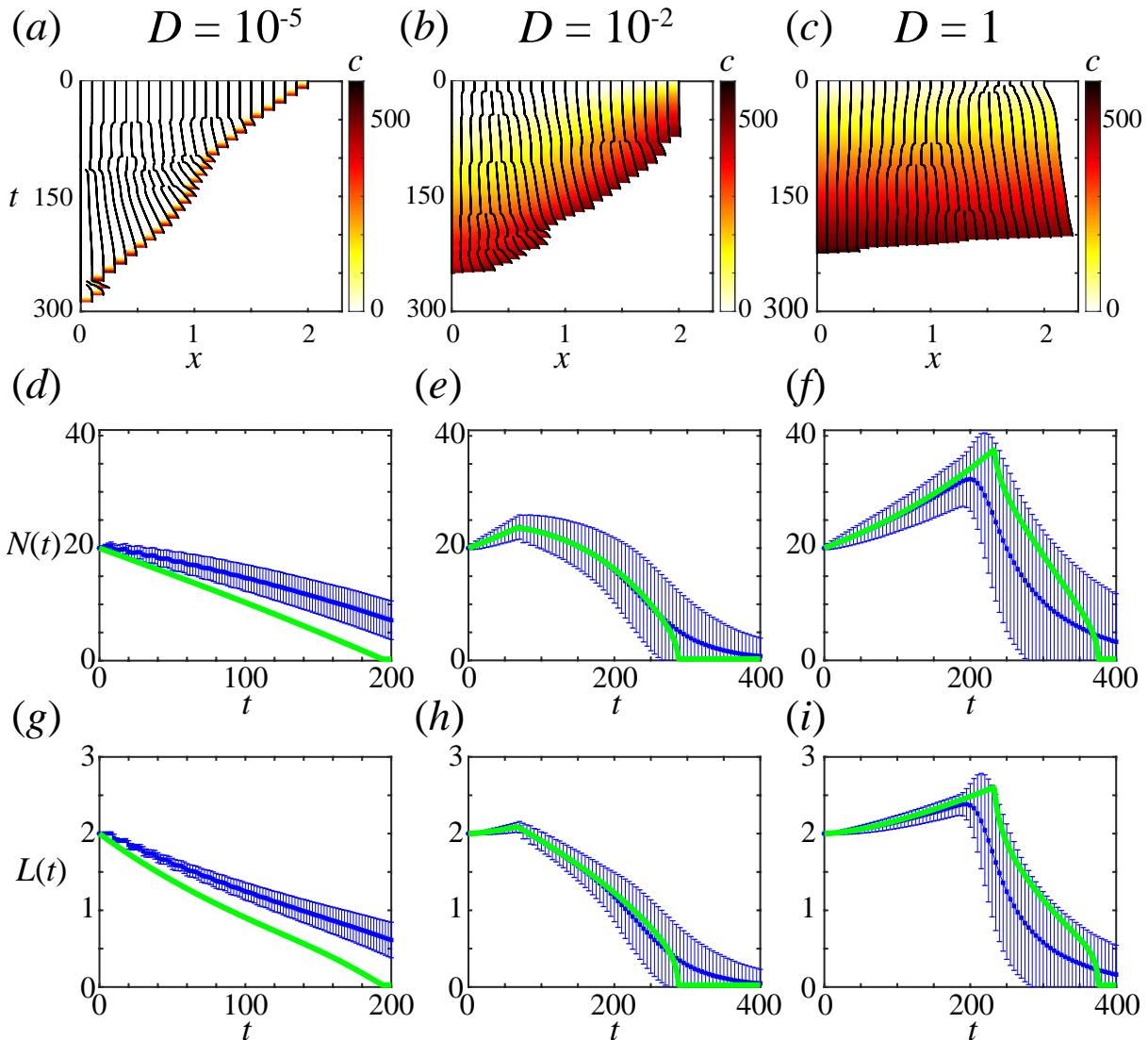
We find that agreement between results from the discrete model and corresponding continuum model is not as accurate as before for large values of  $\phi$  (Figures 5.6(a),(d),(g) for  $\phi = 0.9$ , Supplementary Material 5A.4.5). In the discrete model we assume that a constant number of molecules of the chemical are supplied to the boundary cell (Equation (5.3.6)), to mimic a chemical diffusing in from the external environment, and assume that the concentration in every cell is well-mixed. However, for low diffusivities, here  $D = 10^{-5}$ , the well-mixed assumption is not valid. So the rate of cell detachment,  $\omega(c_N(t))$ , should be updated to account for intracellular chemical concentration gradients. In contrast to the discrete model, the continuum model does capture intracellular concentration gradients and the rate of cell detachment is determined by  $c(L(T), t)$ , which is the concentration at the right edge of the boundary cell. With low diffusivity, the chemical concentration is localised to  $L(t)$  in the continuum model, rather than spread throughout the cell, so the continuum model reaches the concentration threshold faster than realisations of the discrete model. This explains the difference in Figures 5.6(d),(g). When  $\phi$  is small or diffusivity is increased these differences are smaller (Supplementary Material 5A.4.6,5.4.4).

#### 5.4.4 Chemically-dependent EMT with higher diffusivity

Higher chemical diffusivity results in the boundary cell having a lower chemical concentration, on average, than the same simulation with lower chemical diffusivity. This means that the time for the first cell to undergo EMT and detach increases (Figures 5.6(b),(c)). The delay of the first cell detachment can be sufficient to result in a transient rise in total cell number. However, as diffusivity is high the chemical concentration inside internal cells is close to the chemical concentration inside the boundary cell. Therefore, after the first cell detaches the new boundary cell may already be above or close to the concentration threshold, and hence quickly gains the invasive phenotype. This can lead to a rapid sequence of cell detachments, which was not seen with the models in the previous sections (Figures 5.6(b),(c), 5A.4(e)-(g)).

Results in Figures 5.6(e),(f),(h),(i) show good agreement between the continuum model and the appropriately averaged quantities from many discrete realisations. The difference in Figures 5.6(e),(h) for  $t \geq 250$  is due to low  $N(t)$  near extinction (Section 5.4.1, Supplementary Material 5A.4.4). The difference in Figures 5.6(f),(i) at  $t \geq 200$  is due stochastic effects in the discrete model, including the number of cells and tissue length. These are more prominent for  $D = 1$  (Figures 5.6(f),(i)) in comparison to  $D = 10^{-3}$  (Figures 5.6(e),(h)) due to the increased time to reach the concentration threshold (Supplementary Material 5A.4.5).

Special cases assuming the chemical in the tissue is at diffusive equilibrium at all times and instantaneous mechanical relaxation result in four possible behaviours (Supplementary Material (5A.5)): unbounded tissue growth without EMT and cell detachment; unbounded tissue growth with some EMT; eventual tissue homeostasis and constant EMT; and eventual tissue extinction due to EMT.



**Figure 5.6:** Increasing diffusion delays first EMT. Cell detachment driven by chemically-dependent EMT with varying diffusivities and a cell-length-independent proliferation mechanism. Cells are initially at their resting cell lengths with initial cell populations  $N(0) = 20$ . Kymographs with chemical concentration,  $c(x, t)$ , colouring shown for (a)  $D = 10^{-5}$ , (b)  $D = 10^{-2}$ , (c)  $D = 1$ . (d)-(i) The average of 2000 discrete realisations (blue) are compared with the continuum model (green). (d)-(f) Evolution of total cell number,  $N(t)$ . (g)-(i) Evolution of tissue length,  $L(t)$ . Mechanical parameters:  $k = 1$ ,  $a = 0.1$ ,  $\eta = 1$ .

## 5.5 Conclusion

In this work we seek to explore how mechanical interactions influence the evolution of an epithelial tissue. Using our mathematical model, we ask when do tumours grow or shrink, and how mechanical interactions and an EMT-inducing chemical could influence the rate of cell detachment. Starting with a stochastic cell-based discrete model describing mechanical relaxation, cell-length-dependent proliferation, and cell detachment driven by chemically-dependent EMT, we derive the corresponding deterministic continuum description which takes the form of a novel nonlinear free boundary problem. In contrast to previous free boundary models we derive the boundary condition from cell-level biological processes and incorporate EMT. Both the discrete and continuum models useful information: discrete models show the important role of stochastic effects while continuum models help classify possible behaviours. Our results show good agreement between the continuum model and appropriately averaged quantities from many discrete realisations. However, as can be expected, there are occasions when the deterministic continuum model does not capture the fact that, due to stochastic proliferation and EMT in the discrete model, different identically prepared individual discrete realisations may result in different long-term behaviour.

Our models suggest that the coupling of mechanical interactions with EMT is important, can change the probability of long-term extinction significantly, and give rise to different rates of cell detachment [12, 187]. Using our model we postulate that to prevent cell detachment driven by EMT and delay the start of the metastatic cascade, one could chemically alter the speed of mechanical relaxation to encourage the tissue length to increase and hence cause the EMT-inducing chemical concentration to decrease. However, if the tissue length increases then proliferation may be more likely and the number of cancer cells in primary tumour would increase, which is also not desirable. Therefore, there is a delicate trade off between proliferation and EMT that should be considered when seeking to prevent cancer development. Furthermore, the model predicts that if EMT is delayed then the tissue may rapidly collapse due to many cells detaching in quick succession, which is undesirable as it may encourage metastasis. In contrast, for wound healing we may prefer cell-detachment driven by EMT and more proliferation to encourage the wound to heal faster. It will be useful to explore these ideas by extending this model to track individual cells or clusters of cells [26] that detach from the tissue in a two-regime model [216] or multi-organ model [69, 70], and incorporating mesenchymal–epithelial transitions (MET) [84, 102, 115]. Furthermore, the time taken for a cell to proliferate can be on the order of once every 12 hours to once every few days and EMT can occur over the course of hours, a few days [6], or many days (e.g. 9–12 days [112]), depending on the cell

type and driving mechanisms. Therefore, the critical cell number  $\omega/\beta$  and critical tissue length  $\omega a/\beta \mu\text{m}$  we consider in this work are of a similar order of magnitude to that expected *in vitro* and may be interesting to test experimentally.

The mathematical framework that we develop here, and in related studies (Chapters 2, 3, and 4, [19, 157–159, 225]), is well-suited to incorporate additional biological mechanisms and explore different modelling assumptions. One modelling assumption we could change would be to allow the EMT-inducing chemical to diffuse across the boundary at the rear of the tissue, which may prevent a build up of chemical. Introducing intracellular diffusion in the discrete model would also resolve the issue around the well-mixed assumption not being valid for low diffusivities. We also assume, to illustrate a potential role for mechanochemical coupling, that a single chemical drives EMT. In reality, many biological processes at the level of proteins, mRNAs, and miRNAs occur and it may be interesting to incorporate regulatory networks which govern these processes into each cell in the discrete model [137, 226]. While linear diffusion of the EMT-inducing chemical between cells is arguably the simplest approach to include chemical transport and some experimental evidence exists for intercellular chemical transport [113], other mechanisms may be more biologically realistic, such as chemicals diffusing externally and being uptaken by cells [25] and cell adhesion regulated by interactions between E-cadherin and  $\beta$ -catenin [187]. It is an interesting question to ask whether other transport mechanisms are well approximated by the linear diffusion model we consider here.

The one-dimensional approach taken in this work has many advantages in its predictive power, interpretability, and relative computational simplicity in comparison to two- or three-dimensional models. Furthermore, cell-length-dependent proliferation may be thought of as an approximation for cell-volume-dependent proliferation which occurs for cells that move in three-dimensional environments. However, real cells can also spread without changing volume, so it may be beneficial to explore the role of the cell cycle in this one-dimensional framework [234]. A significant extension of this work would be to consider higher dimensions. The discrete model could be extended by considering a cell-centre or vertex model which introduces questions regarding cell shape and how neighbours can be identified, along with increased computational expense [67, 175, 179]. A corresponding continuum model in higher dimensions is less clear. The one-dimensional model enforces an ordering of neighbouring cells, which is important when deriving a continuum model [68, 163]. However, in higher dimensions cells can change their neighbours which poses significant challenges [68, 163]. We leave this extension for future consideration.



## **Chapter 5A**

# **Supplementary Material for Chapter 5**

<b>Contents</b>	<b>Page No.</b>
S1 Continuum model: Evolution of free boundary equation derivation	179
S2 Numerical methods: Discrete model	180
S3 Numerical methods: Continuum model	189
S3.1 Initial conditions	194
S4 Additional results	195
S4.1 Counting the total number of cells that detach	195
S4.2 Cell-length-independent proliferation	197
S4.3 Cell-length-dependent proliferation	201
S4.4 Differences between the discrete and continuum models: low $N(t)$	203
S4.5 Chemically-dependent EMT: cell-length-independent proliferation	205
S4.5.1 Small diffusivity, $D = 10^{-5}$	205
S4.5.2 Starting closer to chemical threshold with $D = 1$	207
S4.6 Sensitivity to $\phi$	209
S4.7 Chemically-dependent EMT: linear proliferation	211
S5 Diffusive equilibrium at all times	213

## 5A.1 Continuum model: Evolution of free boundary equation derivation

The discrete model allows for cells to detach at the free boundary,  $x = L(t)$ . In the continuum model this corresponds to loss of tissue material at a moving interface [34]. By considering conservation of mass for the total number of cells,  $N(t)$ , the rate of change of  $N(t)$  due to proliferation and invasion is

$$\frac{dN(t)}{dt} = \underbrace{\int_{x=0}^{x=L(t)} q(x, t) P\left(\frac{1}{q(x, t)}\right) dx}_{\text{growth due to proliferation}} - \underbrace{\omega(c(L(t), t))}_{\text{loss due to cell detachment}}. \quad (5A.1.1)$$

The rate of change of  $N(t)$  can also be written in terms of the cell density,  $q(x, t)$ , as follows

$$\frac{dN(t)}{dt} = \frac{d}{dt} \left( \int_{x=0}^{x=L(t)} q(x, t) dx \right). \quad (5A.1.2)$$

Differentiating the right hand side of Equation (5A.1.2) with respect to  $t$  and applying Equation (5.3.8) for the cell density and Equation (5.3.9) for the boundary condition at  $x = 0$  gives

$$\begin{aligned} \frac{dN(t)}{dt} = & \frac{dL(t)}{dt} q(L(t), t) - \frac{1}{\eta} \frac{\partial f(L(t), t)}{\partial x} \\ & + \int_{x=0}^{x=L(t)} q(x, t) P\left(\frac{1}{q(x, t)}\right) dx. \end{aligned} \quad (5A.1.3)$$

Equating (5A.1.1) and (5A.1.3) and rearranging we obtain evolution of the free boundary equation

$$\frac{dL(t)}{dt} = \frac{1}{q(L(t), t)} \left[ \frac{1}{\eta} \frac{\partial f(L(t), t)}{\partial x} - \omega(c(L(t), t)) \right]. \quad (5A.1.4)$$

Substituting Equation (5.3.10) into Equation (5A.1.4) we can obtain a different form for the evolution of the free boundary equation as

$$\begin{aligned} \frac{dL(t)}{dt} = & \underbrace{\frac{k(s(t), t)}{\eta} \left[ a(L(t), t) - \frac{1}{q(L(t), t)} - \frac{1}{2q(L(t), t)^3} \frac{\partial q(x, t)}{\partial x} \right]}_{\text{mechanical relaxation}} \\ & - \underbrace{\frac{\omega(c(L(t), t))}{q(L(t), t)}}_{\text{loss due to cell detachment}}. \end{aligned} \quad (5A.1.5)$$

## 5A.2 Numerical methods: Discrete model

The key technical challenge to overcome to numerically simulate the discrete model concerns how to implement linear chemical diffusion in an evolving population of cells with variable cell lengths. This is the primary focus of the first half of this section. Previous models have implemented diffusion on growing domains [20, 48, 49, 203, 207–209, 252], however, what is unique to this work is that we are interested in the chemical concentration inside individual cells when the positions of cell boundaries are known and evolve in time. Furthermore, previous studies tend to consider uniform growth throughout the tissue whereas here, due to mechanical interactions and proliferation, we have non-uniform growth throughout the tissue. As we will show these complications require a new numerical method.

To model the diffusion of molecules of a chemical we have a choice between microscopic or mesoscopic individual-based models or macroscopic population-based models. Microscopic individual-based models are often posed as a population of particles undergoing Brownian motion. Mesoscopic individual-based models are often posed as a population of particles undergoing a random walk on a lattice or as a position-jump process on a lattice. However, individual-based models tend to be more computationally expensive than population-level models and can be mathematically intractable. Macroscopic approaches can be simpler to write down and are often easier and faster to simulate for a large number of particles.

Macroscopically, linear diffusion of particles on a fixed domain,  $0 < x < L$ , can be modelled with the following classical partial differential equation

$$\frac{\partial c(x, t)}{\partial t} = D \frac{\partial^2 c(x, t)}{\partial x^2}, \quad 0 < x < L, \quad (5A.2.1)$$

where  $c(x, t)$  is the particle density, or equivalently the chemical concentration, and  $D$  is the macroscopic diffusion coefficient. Whereas on a domain whose length,  $L(t)$ , evolves in time, conservation of mass arguments and applying Reynolds transport theorem gives the following partial differential equation for the evolution of  $c(x, t)$ , [20]

$$\frac{\partial c(x, t)}{\partial t} + \frac{\partial}{\partial x} (u(x, t)c(x, t)) = D \frac{\partial^2 c(x, t)}{\partial x^2}, \quad 0 < x < L(t), \quad (5A.2.2)$$

where  $u(x, t)$  is a velocity field prescribed by domain growth [20, 48, 49, 252]. In this work, Equation (5A.2.2) is the same as Equation (5.3.13) in the main manuscript with  $u(x, t) = 1/(\eta q(x, t)) \partial f(x, t)/\partial x$  determined by mechanical interactions between the cells.

Previous studies have demonstrated that a stochastic individual-based model incorporating

domain growth, taking the form of a position-jump model on a uniform lattice, is equivalent to the continuum model in Equation (5A.2.2) [20]. However, domain growth was implemented by instantaneous doubling and dividing of underlying lattice sites which always results in a uniform lattice, corresponding to all cells have the same length at all times. This is not the case for this work. We obtain a non-uniform lattice, as cell lengths vary due to the effects of mechanical interactions between cells and proliferation. Therefore, to discuss the method we will focus on a non-uniform lattice and initially assume that positions of the cell boundaries are fixed and known.

To model diffusion on a non-uniform lattice Yates et al. [252] make clear that one must be careful, and suggest two methods which we refer to as method A and method B. To explain the methods we first state the two key sets of points: i) positions of cell boundaries,  $x_i$  for  $i = 1, 2, \dots, N + 1$  (represented as circles in schematics in Figures 5.1, 5A.1-5A.3); ii) resident points,  $y_i$  for  $i = 1, 2, \dots, N$ , satisfying  $x_i < y_i < x_{i+1}$ , which defines the location in cell  $i$  where the particles of the chemical are considered to be positioned (represented as crosses in schematics in Figures 5A.1-5A.3). In method A (Figure 5A.1(a)), Yates et al. [252] assume the resident points,  $y_i$ , are chosen first and then the cell boundaries,  $x_i$ , are defined in a Voronoi neighbourhood sense: a point is in cell  $i$  if it is closer to the resident point associated with cell  $i$ , given by  $y_i$ , rather than any other resident point  $y_j$ . They show this method can be used to accurately model linear diffusion due to the Voronoi partition (see Supplementary Material Section 1 of Yates et al. [252]). However, in this work the positions of the cell boundaries are already known from the mechanical interactions (Equations (5.3.1)-(5.3.3)) so we cannot use method A and instead consider method B. In method B (Figure 5A.1(b)), Yates et al. [252] first prescribe the position of the cell boundaries,  $x_i$ , which is what we require, and then they choose the resident points to be the position of the centre of cell  $i$ , so  $y_i = (x_i + x_{i+1})/2$ . They show this method does not accurately model linear diffusion as there is not a Voronoi partition. When all cells are the same size, resulting in a uniform lattice, methods A and B are equivalent.

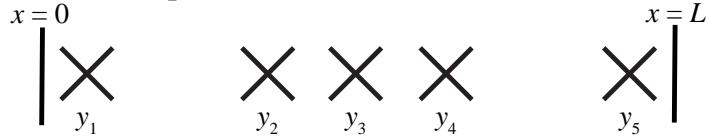
Before proceeding with our new method, which combines and extends methods A and B, we briefly discuss the underlying microscopic diffusion model and how it relates to the mesoscale position-jump process model of diffusion [252]. In the microscopic model of diffusion, the position of an individual particle which undergoes Brownian motion is governed by a stochastic differential equation,

$$dX(t) = \sqrt{2D} dW_t, \quad (5A.2.3)$$

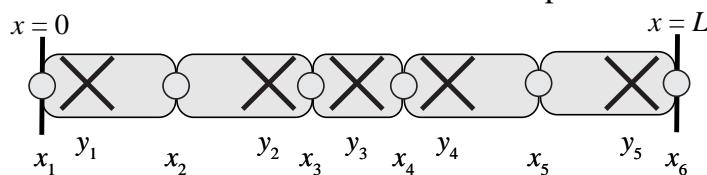
where  $dW_t$  is a standard Wiener process and  $D$  is the macroscopic diffusion coefficient. For

**(a) Method A**

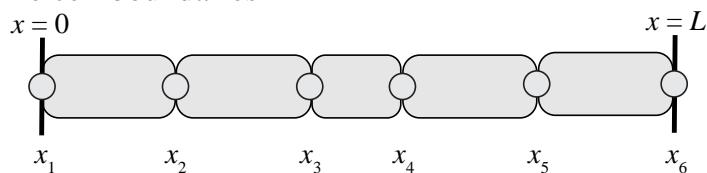
i) Define resident points



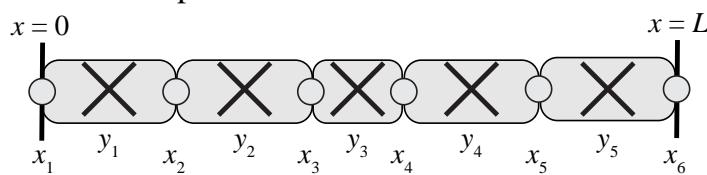
ii) Define cell boundaries based on a Voronoi partition

**(b) Method B**

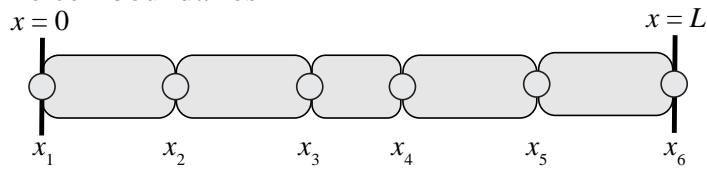
i) Define cell boundaries



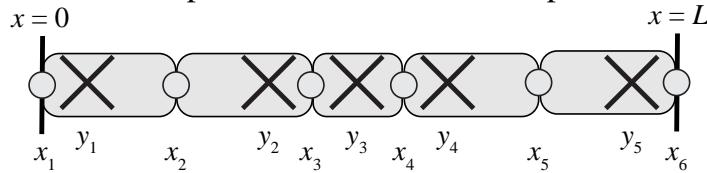
ii) Define resident points as cell centres

**(c) Interval-Voronoi method**

i) Define cell boundaries



ii) Choose resident points to define a Voronoi partition



**Figure 5A.1:** Different methods to model linear diffusion for variable cells sizes. (a) Method A: define resident points,  $y_i$ , first and then define cell boundaries,  $x_i$ , to form a Voronoi partition. (b) Method B: define cell boundaries first and then define resident points at the cell centres. (c) New Interval-Voronoi method: define cell boundaries first and then choose resident points to define a Voronoi partition. Circles represent cell boundaries,  $x_i$ , and crosses represent resident points,  $y_i$ . Shown for  $N = 5$  cells.

the mesoscale position-jump process model of diffusion we seek the rates at which a particle at resident point,  $y_i$ , moves to the neighbouring left or right resident points at  $y_{i-1}$  or  $y_{i+1}$ , respectively. These transition rates,  $T_i^\pm$ , can be found by initialising a particle at  $y_i$  solving a first passage time problem on the domain  $y_{i-1} < x < y_{i+1}$  [190, 252]. Then the transition rates are [59, 252]

$$T_1^+ = \frac{2D}{(y_2-y_1)(y_1+y_2)}, \quad (5A.2.4)$$

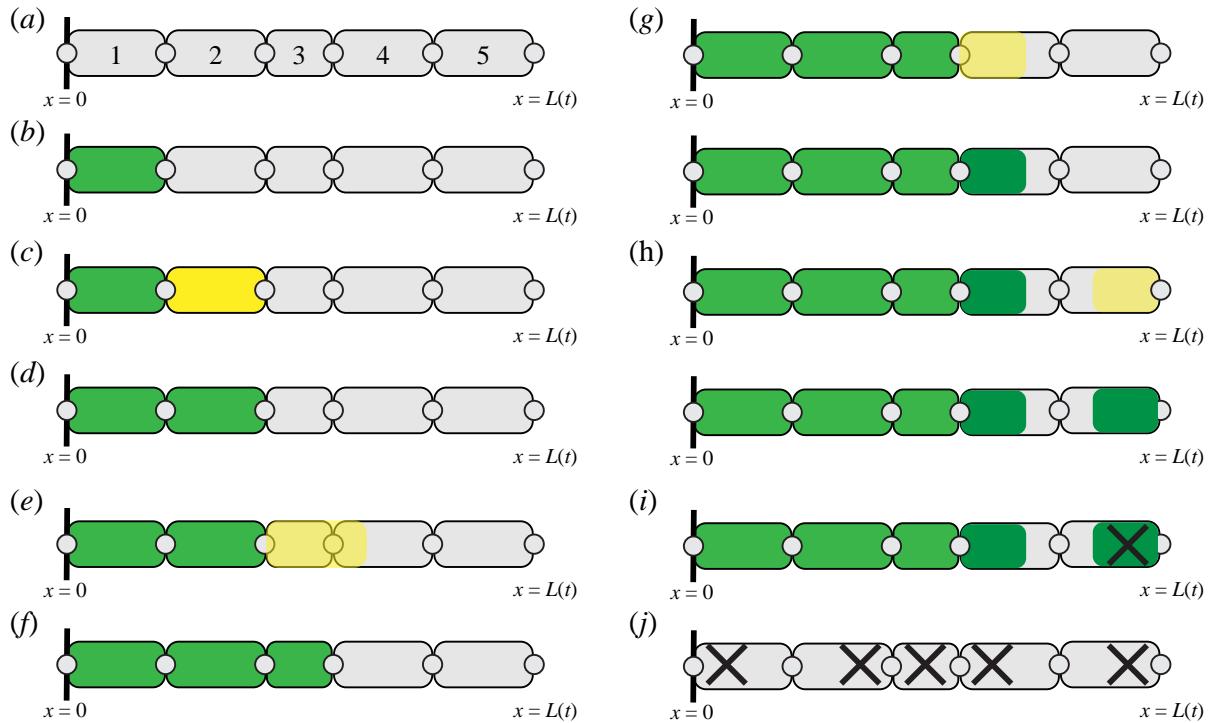
$$T_i^+ = \frac{2D}{(y_{i+1}-y_i)(y_{i+1}+y_{i-1})} \quad i = 2, 3, \dots, N-1, \quad (5A.2.5)$$

$$T_i^- = \frac{2D}{(y_i-y_{i-1})(y_{i+1}+y_{i-1})} \quad i = 2, 3, \dots, N-1, \quad (5A.2.6)$$

$$T_N^- = \frac{2D}{(y_N-y_{N-1})(2(L(t)-y_N)+y_N-y_{N-1})}, \quad (5A.2.7)$$

where  $D$  is the macroscopic diffusion coefficient. When all cells are the same length,  $l$ , these transition rates simplify to  $D/l^2$ .

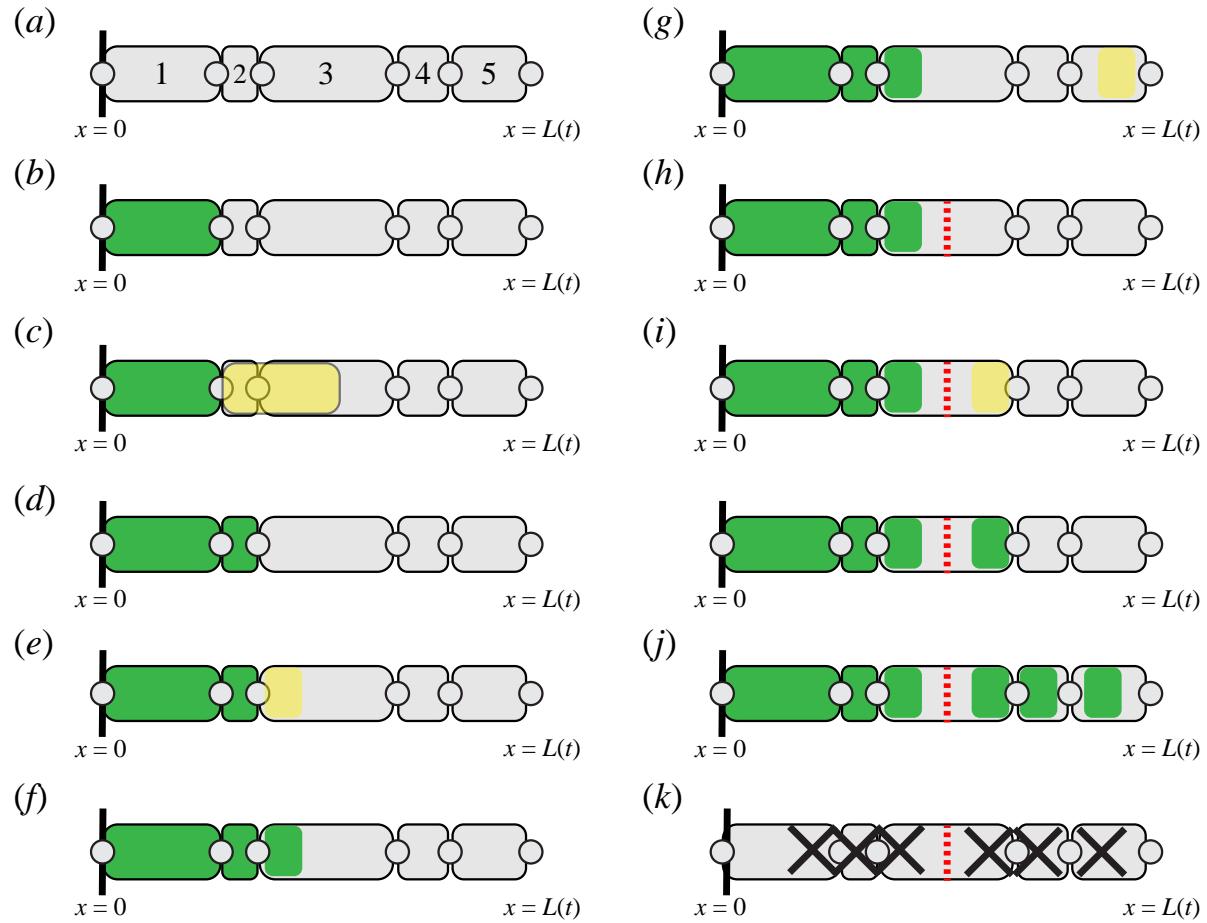
To accurately model linear diffusion in our work we combine and extend methods A and B. Specifically, we first define the positions of the cell boundaries,  $x_i$ , and then we choose the resident points,  $y_i$ , so that the position of the bisection of neighbouring resident points is the position of a cell boundary, i.e. choose  $y_i$  such that  $(y_i + y_{i+1})/2 = x_i$  for  $i = 2, 3, \dots, N$  (Figure 5A.1(c)). This results in a Voronoi partition on the set of  $y_i$ , where the edges of the Voronoi partition coincide with the cell boundaries. We now explain how to choose the  $y_i$  in such a manner by following Figure 5A.2. We assume an initial cell configuration (Figure 5A.2(a)) and will work from the leftmost cell to the rightmost cell. Initially, the resident point of the leftmost cell could be placed anywhere in this cell, which we call the possible region of the first cell and show in green (Figure 5A.2(b)). Next we reflect the possible region of the first cell about the right boundary of the first cell, and colour this yellow (Figure 5A.2(c)). Then intersecting the reflected possible region of the first cell with the interval occupied by the second cell gives the possible region for the second cell, which we indicate in green (Figure 5A.2(d)). We repeat until reaching the rightmost cell (Figure 5A.2(e)-(h)). Given possible regions for all of the cells we choose the resident point of the rightmost cell to be the midpoint of the possible region of the rightmost cell (Figure 5A.2(i)). Then reflecting the resident point about the left boundary of the rightmost cell we obtain the location of the resident point of the penultimate cell. We repeat until we have obtained the resident points for all cells (Figure 5A.2(j)). This method gives us a set of resident points which can be used to accurately model linear diffusion. However, this method only works when a Voronoi partition can be defined, which occurs when cells are similar lengths.



**Figure 5A.2:** Schematic for *Interval-Voronoi* method for cells of similar size. Circles represent cell boundaries,  $x_i$ , crosses represent resident points,  $y_i$ , and green represent possible regions for a cell where a resident point could be placed. Details discussed in the text.

Depending on the configuration of the positions of the cell boundaries it is not always possible to define the resident points as above. For example, following Figure 5A.3, the possible regions for the first three cells can be defined as before (Figure 5A.3(a)-(f)). However, when we reflect the possible region of the third cell and intersect this with the interval occupied by the fourth cell we obtain an empty set. Therefore, there is no location in the third cell where we will be able to place a resident point to define a Voronoi partition. To resolve this problem, we now divide the third cell into two compartments (Figure 5A.3(h)). To do so, we first assume that the current possible region for the third cell is now the possible region for the left compartment of the third cell. Then, we choose the position of the compartment boundary which divides the cell (red-dashed line in Figure 5A.3(h)) so that the position of the right boundary of the right compartment of the third cell is equal to the position of the right boundary of the third cell, the possible region of the left compartment of the third cell does not overlap with the possible region of the right compartment of the third cell, and the possible region of the right compartment of the third cell is maximised. In Figure 5A.3(h) this corresponds to dividing the cell into two compartments of equal length. We can then proceed as before (Figure 5A.3(i)-(j)). We can then define resident points for each compartment and subsequently for each cell (Figure 5A.3(k)). When numerically simulating the discrete model, as described further below, we update the chemical concentration using a well-mixed assumption at each time step

so this method does not result in intracellular chemical concentration gradients. This is a valid approximation when  $D$  is sufficiently large.



**Figure 5A.3:** Schematic for *Interval-Voronoi* with compartments per cell. This method is required when a Voronoi partition cannot be defined on the initial cell configuration, which occurs when cell lengths are not similar. Circles represent cell boundaries,  $x_i$ , crosses represent resident points,  $y_i$ , green represent possible regions for a cell where a resident point could be placed, red-dashed represents the a compartment boundary within a cell. Details discussed in the text.

Being able to define a Voronoi partition is necessary [252] but results from simulations show that it is not always sufficient to accurately model linear diffusion. The distances between neighbouring resident points should also be the same order of magnitude throughout the population. If we suppose the resident points are  $y_i$  for  $i = 1, 2, \dots, N$  then if

$$\log_{10} (\max(\{y_{i+1} - y_i : i = 1, 2, \dots, I - 1\})) - \log_{10} (\min(\{y_{i+1} - y_i : i = 1, 2, \dots, N - 1\})) < \gamma, \quad (5A.2.8)$$

the distances between neighbouring resident points are approximately the same order and we can proceed. The condition in Equation (5A.2.8) and the value  $\gamma = 0.8$  were found to be good choices for results in this work. If the condition in Equation (5A.2.8) is not satisfied we introduce

a maximum compartment length equal to half of the minimum length of the current cell sizes. Then for any cells whose length is greater than the maximum compartment length, we divide those cells into compartments so that the maximum length of any compartment is less than the maximum compartment length. We then determine the resident points as above.

The method described above, which we name the *Interval-Voronoi* method, can now be used to accurately model linear diffusion on a fixed domain with variable cell lengths. In this work, all of the cell boundaries, and consequently the length of the domain,  $L(t)$ , are evolving in time. However, when we numerically simulate a single realisation of the discrete model we discretise time with a constant time step  $\Delta t$ . Then for each time interval  $[t, t + \Delta t)$  we assume the cell boundaries are fixed and the domain is fixed and apply the *Interval-Voronoi* method. Further details are now shown.

Let us consider a single realisation of the discrete model. The epithelial tissue is initialised with  $N$  cells each with cell stiffness  $k$  and resting cell length  $a$ . The position of each cell boundary  $x_i(t)$  for  $i = 1, 2, \dots, N$  is defined. Every cell in the tissue is prescribed with the same proliferation mechanism and proliferation parameter,  $\beta$ . The chemical concentration is initially set to zero in each cell. To simulate the model, we discretise time with a constant time step  $\Delta t$ . Then for each time interval  $[t, t + \Delta t)$  we: i) update cell positions according to mechanical interactions; ii) update the chemical concentrations in each cell; iii) implement proliferation or cell detachment if it occurs.

We update the positions of each of the cell boundaries,  $x_i(t)$ , by integrating Equations (5.3.1)-(5.3.3) using a forward Euler approximation,

$$x_1(t + \Delta t) = 0, \quad (5A.2.9)$$

$$x_i(t + \Delta t) = x_i(t) + \frac{\Delta t}{\eta} \left[ k(x_{i+1} - x_i - a) \right. \quad (5A.2.10)$$

$$\left. -k(x_i - x_{i-1} - a) \right], \quad i = 2, \dots, N-1,$$

$$x_{N+1}(t + \Delta t) = x_{N+1}(t) - \frac{\Delta t}{\eta} k(x_{N+1} - x_N - a). \quad (5A.2.11)$$

Next we seek to update the chemical concentration by applying the *Interval-Voronoi* method. Using the updated positions of cell boundaries,  $x_i(t + \Delta t)$  for  $i = 1, 2, \dots, N+1$ , we determine the resident points  $y_i$  for  $i = 1, 2, \dots, \hat{N}$ . Then we can calculate the boundaries of compartments  $\hat{x}_i$  for  $i = 1, 2, \dots, \hat{N}+1$ , in a Voronoi neighbourhood sense, where  $\hat{N} \geq N$  is the number of resident points. If the *Interval-Voronoi* method requires at least one cell to be divided into compartments then  $\hat{N} > N$  in order for a Voronoi partition to be defined. Then given the resident points we determine the transition rates,  $\hat{T}_i^\pm$ , in terms of compartment lengths,

$\hat{l}_i(t + \Delta t) = \hat{x}_{i+1}(t + \Delta t) - \hat{x}_i(t + \Delta t)$ . We now have the information required to update the chemical concentration for the cells. First we update the chemical concentration for all compartments, which we denote as  $\hat{c}_i(t)$  for  $i = 1, 2, \dots, \hat{N}$  by integrating Equations (5.3.4)-(5.3.6) using forward Euler approximations,

$$\begin{aligned} \hat{c}_1(t + \Delta t) = & \hat{c}_1(t) + \Delta t \left[ \frac{1}{\hat{l}_1(t+\Delta t)} \left( \hat{T}_2^- \hat{c}_2(t) \hat{l}_2(t + \Delta t) - \hat{T}_1^+ \hat{c}_1(t) \hat{l}_1(t + \Delta t) \right) \right] \\ & - \frac{\hat{c}_1(t)}{\hat{l}_1(t+\Delta t)} \left( \hat{l}_1(t + \Delta t) - \hat{l}_1(t) \right) \end{aligned} \quad (5A.2.12)$$

$$\begin{aligned} \hat{c}_i(t + \Delta t) = & \hat{c}_i(t) + \Delta t \left[ \frac{1}{\hat{l}_i(t+\Delta t)} \left( \hat{T}_{i-1}^+ \hat{c}_{i-1}(t) \hat{l}_{i-1}(t + \Delta t) \right. \right. \\ & \left. \left. - \left( \hat{T}_i^- + \hat{T}_i^+ \right) \hat{c}_i(t) \hat{l}_i(t + \Delta t) + \hat{T}_{i+1}^- \hat{c}_{i+1}(t) \hat{l}_{i+1}(t + \Delta t) \right) \right] \\ & - \frac{\hat{c}_i(t)}{\hat{l}_i(t+\Delta t)} \left( \hat{l}_i(t + \Delta t) - \hat{l}_i(t) \right), \quad i = 2, 3, \dots, \hat{N} - 1, \end{aligned} \quad (5A.2.13)$$

$$\begin{aligned} \hat{c}_{\hat{N}}(t + \Delta t) = & \hat{c}_N(t) + \Delta t \left[ \frac{1}{\hat{l}_{\hat{N}}(t+\Delta t)} \left( \hat{T}_{N-1}^+ \hat{c}_{N-1}(t) \hat{l}_{N-1}(t + \Delta t) - \hat{T}_N^- \hat{c}_N(t) \hat{l}_N(t + \Delta t) \right) \right] \\ & - \frac{\hat{c}_{\hat{N}}(t)}{\hat{l}_{\hat{N}}(t+\Delta t)} \left( \hat{l}_{\hat{N}}(t + \Delta t) - \hat{l}_{\hat{N}}(t) \right) + \frac{S\Delta t}{l_{\hat{N}}(t+\Delta t)}. \end{aligned} \quad (5A.2.14)$$

If the *Interval-Voronoi* method does not introduce any compartments per cell, i.e. if  $\hat{N} = N$ , then  $\hat{T}_i^\pm = T_i^\pm$ ,  $\hat{l}_i(t + \Delta t) = l_i(t + \Delta t)$ , and  $\hat{c}_i(t) = c_i(t)$ . Hence, if  $\hat{N} = N$ , using Equations (5A.2.13)-(5A.2.14), we directly determine the chemical concentrations  $c_i(t + \Delta t)$  for  $i = 1, 2, \dots, N$  and we can proceed to incorporating if a cell proliferation or detachment event occurs in the time step. However, if  $\hat{N} > N$  before proceeding we apply the well-mixed assumption to each cell. Specifically, if the *Interval-Voronoi* method introduces  $j$  compartments into cell  $i$  with concentrations  $\hat{c}_k(t + \Delta t), \hat{c}_{k+1}(t + \Delta t), \dots, \hat{c}_{k+j-1}(t + \Delta t)$  then the chemical concentration of cell  $i$  is set to be  $c_i(t + \Delta t) = (1/l_i) \sum_{r=1}^j \hat{c}_i(t + \Delta t) \hat{l}_i(t + \Delta t)$ .

Next, given the cell positions and chemical concentrations at time  $t + \Delta t$ , we determine whether a cell proliferation event or a cell detachment event occurs during the time interval  $[t, t + \Delta t]$ . Analogous to Murphy et al. [158] we proceed using rejection sampling [74] where we generate three independent random numbers from a uniform distribution,  $r_1 \sim U[0, 1], r_2 \sim U[0, 1], r_3 \sim U[0, 1]$ . Then a cell event, which could be either a proliferation event or a cell detachment event, occurs when

$$r_1 < \omega(c_N)\Delta t + \sum_{i=1}^{N(t)} P(l_i^N)\Delta t. \quad (5A.2.15)$$

Given that a cell event occurs, a proliferation event occurs if

$$r_2 < \frac{\sum_{i=1}^{N(t)} P(l_i^N)}{\sum_{i=1}^{N(t)} P(l_i^N) + \omega(c_N)}. \quad (5A.2.16)$$

Otherwise we have a cell detachment event and we remove the boundary cell. If a proliferation event occurs, to determine which cell is proliferating we find the index  $j$  which satisfies

$$\frac{\sum_{i=1}^j P(l_i^N)}{\sum_{i=1}^{N(t)} P(l_i^N)} < r_3 \leq \frac{\sum_{i=1}^{j+1} P(l_i^N)}{\sum_{i=1}^{N(t)} P(l_i^N)}. \quad (5A.2.17)$$

We then divide the parent cell into two equally sized daughter cells with the same chemical concentration and mechanical properties as the parent cell.

We repeat the above steps for each time step until the final time. This method requires that at most one cell event can occur within each time step so  $\Delta t$  should be chosen sufficiently small.

### 5A.3 Numerical methods: Continuum model

The continuum model is solved numerically using a boundary fixing transformation [120], finite difference approximations, and the Newton-Raphson method [44] with adaptive time stepping. Full details now follow.

For completeness, we rewrite the governing equations for the cell density,  $q(x, t)$ , from Equation (5.3.8) and chemical concentration,  $c(x, t)$ , from Equation (5.3.12), where we have used the definitions of  $f(x, t)$ , and  $u(x, t)$  from Equation (5.3.13),

$$\frac{\partial q(x, t)}{\partial t} = -\frac{1}{\eta} \frac{\partial^2}{\partial x^2} \left[ k \left( \frac{1}{q(x, t)} - a \right) \right] + q(x, t) P \left( \frac{1}{q(x, t)} \right), \quad 0 < x < L \quad (5A.3.1)$$

$$\frac{\partial c(x, t)}{\partial t} + \frac{\partial}{\partial x} \left( \frac{1}{\eta q(x, t)} \frac{\partial}{\partial x} \left[ k \left( \frac{1}{q(x, t)} - a \right) \right] c(x, t) \right) = D \frac{\partial^2 c(x, t)}{\partial x^2}, \quad 0 < x < L \quad (5A.3.2)$$

These governing equations are solved with the following boundary conditions from Equations (5.3.9), (5.3.10), (5.3.14), (5.3.15),

$$\frac{\partial}{\partial x} \left[ k \left( \frac{1}{q(0, t)} - a \right) \right] = 0, \quad (5A.3.3)$$

$$-k \left( \frac{1}{q(L(t), t)} - a \right) = \frac{1}{2q(L(t), t)} \frac{\partial}{\partial x} \left[ k \left( \frac{1}{q(L(t), t)} - a \right) \right], \quad (5A.3.4)$$

$$\frac{\partial c(0, t)}{\partial x} = 0, \quad (5A.3.5)$$

$$D \frac{\partial c(L(t), t)}{\partial x} = S, \quad (5A.3.6)$$

and the boundary position,  $L(t)$ , evolves according to Equation (5.3.11) and we can apply Equation (5A.3.4) to give

$$\frac{dL(t)}{dt} = \frac{-2k}{\eta} \left( \frac{1}{q(L(t), t)} - a \right) - \frac{\omega}{q(L(t), t)}. \quad (5A.3.7)$$

Equations (5A.3.1)-(5A.3.7) form a moving boundary problem for coupled non-linear partial differential equations. To proceed we apply a standard boundary fixing transformation [120] by setting  $\xi = x/L(t)$  to transform the evolving domain  $0 \leq x \leq L(t)$  to a fixed domain  $0 \leq \xi \leq 1$ . Equations (5A.3.1)-(5A.3.7) then become

$$\begin{aligned} \frac{\partial q(\xi, t)}{\partial t} = & -\frac{1}{L(t)^2 \eta} \frac{\partial^2}{\partial \xi^2} \left[ k \left( \frac{1}{q(\xi, t)} - a \right) \right] + q(\xi, t) P \left( \frac{1}{q(\xi, t)} \right) \\ & + \frac{\xi}{L(t)} \frac{dL(t)}{dt} \frac{\partial q(\xi, t)}{\partial \xi}, \quad 0 < \xi < 1, \end{aligned} \quad (5A.3.8)$$

$$\begin{aligned} \frac{\partial c(\xi, t)}{\partial t} = & -\frac{1}{L(t)^2} \frac{\partial}{\partial \xi} \left( \frac{1}{\eta q(\xi, t)} \frac{\partial}{\partial \xi} \left[ k \left( \frac{1}{q(\xi, t)} - a \right) \right] c(\xi, t) \right) + \frac{D}{L(t)^2} \frac{\partial^2 c(\xi, t)}{\partial \xi^2} \\ & + \frac{\xi}{L(t)} \frac{dL(t)}{dt} \frac{\partial q(\xi, t)}{\partial \xi}, \quad 0 < \xi < 1, \end{aligned} \quad (5A.3.9)$$

with the following boundary conditions

$$\frac{\partial}{\partial \xi} \left[ k \left( \frac{1}{q(0, t)} - a \right) \right] = 0, \quad (5A.3.10)$$

$$-k \left( \frac{1}{q(1, t)} - a \right) = \frac{1}{2q(1, t)} \frac{\partial}{\partial \xi} \left[ k \left( \frac{1}{q(1, t)} - a \right) \right], \quad (5A.3.11)$$

$$\frac{\partial c(0, t)}{\partial \xi} = 0, \quad (5A.3.12)$$

$$\frac{D}{L(t)} \frac{\partial c(1, t)}{\partial \xi} = S, \quad (5A.3.13)$$

and the boundary position,  $L(t)$ , evolves according to

$$\frac{dL(t)}{dt} = \frac{-2k}{\eta} \left( \frac{1}{q(1, t)} - a \right) - \frac{\omega}{q(1, t)}. \quad (5A.3.14)$$

Next, we discretise the domain  $0 \leq \xi \leq 1$  with a uniform mesh with spatial step  $\Delta \xi$  and use the subscript  $j = 1, 2, \dots, J$  to represent the index of the spatial nodes. We discretise time with a uniform mesh with time step  $\Delta t$  and use the superscript  $n = 1, 2, \dots, T$  to represent temporal step. Second-order spatial derivatives are approximated by standard central differences. First-order spatial derivatives are approximated by standard upwind differences. A standard implicit finite difference approximation is used to approximate temporal derivatives.

Equation (5A.3.14) governing the evolution of  $L(t)$  becomes

$$L^{n+1} = L^n + \frac{-2k\Delta t}{\eta} \left( \frac{1}{q_J^n} - a \right) - \frac{\Delta t \omega}{q_J^n}. \quad (5A.3.15)$$

Finite difference approximations of the cell density equations give, for internal spatial nodes

$j = 2, \dots, J - 1$ ,

$$0 = -q_j^{n+1} + q_j^n \quad (5A.3.16)$$

$$-\frac{\Delta t}{\eta(L^{n+1}\Delta\xi)^2} \left[ k \left( \frac{1}{q_{j-1}^{n+1}} - a \right) - 2k \left( \frac{1}{q_j^{n+1}} - a \right) + k \left( \frac{1}{q_{j+1}^{n+1}} - a \right) \right] \quad (5A.3.17)$$

$$+ \Delta t q_j^{n+1} P \left( \frac{1}{q_j^{n+1}} \right) + \frac{\xi \Delta t}{2\Delta\xi L^{n+1}} (L^{n+1} - L^n) (q_{j+1}^{n+1} - q_j^{n+1}). \quad (5A.3.18)$$

At  $\xi = 0$ , corresponding to spatial node 1, the cell density is updated with

$$0 = -k \left( \frac{1}{q_1^{n+1}} - a \right) + k \left( \frac{1}{q_2^{n+1}} - a \right). \quad (5A.3.19)$$

At  $\xi = 1$ , corresponding to spatial node  $J$ , the cell density is updated with

$$0 = \frac{1}{2q_J^{n+1} L^{n+1} \Delta\xi} \left[ k \left( \frac{1}{q_J^{n+1}} - a \right) - k \left( \frac{1}{q_{J-1}^{n+1}} - a \right) \right] + k \left( \frac{1}{q_J^{n+1}} - a \right). \quad (5A.3.20)$$

To proceed we calculate the velocity at each node, given by

$$v_1 = \frac{1}{\Delta\xi \eta q_1^{n+1} (L^{n+1})^2} \left[ k \left( \frac{1}{q_2^{n+1}} - a \right) - k \left( \frac{1}{q_1^{n+1}} - a \right) \right] - \frac{\xi}{L^{n+1}} \frac{L^{n+1} - L^n}{\Delta t}, \quad (5A.3.21)$$

$$v_j = \frac{1}{2\Delta\xi \eta q_i^{n+1} (L^{n+1})^2} \left[ k \left( \frac{1}{q_{j+1}^{n+1}} - a \right) - k \left( \frac{1}{q_{j-1}^{n+1}} - a \right) \right] - \frac{\xi}{L^{n+1}} \frac{L^{n+1} - L^n}{\Delta t}, \quad (5A.3.22)$$

$j = 2, \dots, J - 1,$

$$v_J = \frac{1}{\Delta\xi \eta q_J^{n+1} (L^{n+1})^2} \left[ k \left( \frac{1}{q_J^{n+1}} - a \right) - k \left( \frac{1}{q_{J-1}^{n+1}} - a \right) \right] - \frac{\xi}{L^{n+1}} \frac{L^{n+1} - L^n}{\Delta t}, \quad (5A.3.23)$$

where forward and backward differences applied at the left and right boundaries, respectively.

Considering the chemical concentration for internal nodes  $j = 2, 3, \dots, J - 1$  and assuming

$v_j \geq 0$  so that first order spatial derivatives are upwinded, finite difference approximations give

$$\begin{aligned} 0 = & -c_j^{n+1} + c_j^n \\ & + \frac{\Delta t}{\Delta \xi} v_j (c_j^{n+1} - c_{j-1}^{n+1}) \\ & - \Delta t c_j^{n+1} \frac{1}{2\eta(L^{n+1}\Delta\xi)^2} \left[ \left( \frac{1}{q_{j+1}^{n+1}} + \frac{1}{q_j^{n+1}} \right) \left( k \left( \frac{1}{q_{j+1}^{n+1}} - a \right) + k \left( \frac{1}{q_j^{n+1}} - a \right) \right) \right. \\ & \left. \left( \frac{1}{q_j^{n+1}} + \frac{1}{q_{j-1}^{n+1}} \right) \left( k \left( \frac{1}{q_j^{n+1}} - a \right) + k \left( \frac{1}{q_{j-1}^{n+1}} - a \right) \right) \right] \\ & \frac{D\Delta t}{(\Delta\xi L^{n+1})^2} (c_{j-1}^{n+1} - 2c_j^{n+1} + c_{j+1}^{n+1}). \end{aligned} \quad (5A.3.24)$$

Similarly, if  $v_j < 0$  upwinding is applied to the first order spatial derivative. At  $\xi = 0$  we apply a forward difference approximation to Equation (5A.3.12) governing the chemical concentration which gives

$$0 = -c_1^{n+1} + c_2^{n+1}. \quad (5A.3.25)$$

At  $\xi = 1$  we apply a backwards difference approximation to Equation (5A.3.13) governing the chemical concentration which gives

$$0 = -c_J^{n+1} + c_{J-1}^{n+1} + \frac{S\Delta\xi L^{n+1}}{D}. \quad (5A.3.26)$$

Equations (5A.3.15)-(5A.3.26) form a system of nonlinear algebraic equations for the cell density, chemical concentration, and evolution of  $L(t)$ . We solve these equations using the Newton-Raphson method [44, 158]. In each Newton-Raphson iteration we first update  $L(t)$ , according to Equation (5A.3.15), then the cell density, governed by Equations (5A.3.16)-(5A.3.20), then the chemical concentration, governed by Equations (5A.3.24)-(5A.3.26). Newton-Raphson iterations are performed at each time point until the infinity norm of the the difference between successive estimates of  $\{q_1^{n+1}, q_2^{n+1}, \dots, q_J^{n+1}\}$  and  $\{c_1^{n+1}, c_2^{n+1}, \dots, c_J^{n+1}\}$  is below a specified tolerance,  $\epsilon$ . To ensure that the Newton-Raphson iteration converges we apply adaptive time stepping. To implement adaptive time stepping we introduce a maximum number of iterations. When the maximum number of iterations is reached without the tolerance being met we divide the timestep by ten and repeat. Once the tolerance is met we reset the time step for the next temporal node. In our results adaptive time stepping is important when the chemical concentration first reaches the chemical threshold at  $L(t)$ , and it reduces the computational time required to obtain the numerical solution. We use the Thomas algorithm to solve the linear systems which arise from the Newton-Raphson method. To ensure all numerical results

are grid-independent we set  $\Delta\xi = 10^{-5}$ , initially set  $\Delta t = 10^{-3}$ , set the maximum number of iterations for each time step to ten, and set  $\epsilon = 10^{-8}$ .

Key algorithms used to generate results are available on Github.

### 5A.3.1 Initial conditions

In this work, we refer to three types of initial conditions: compressed, mechanical equilibrium, and stretched. Here, we state these for the discrete and continuum model.

In the discrete model we choose every cell to initially have the same length,  $l_i(0) = L(0)/N(0)$ . If  $l_i < a$  then each cell is compressed and the tissue is compressed. If  $l_i = a$  then each cell is at its resting cell length and the tissue is at mechanical equilibrium. If  $l_i > a$  then each cell is stretched and the tissue is stretched. The corresponding initial conditions in the continuum model are obtained using  $q(x, 0) = N(0)/L(0)$  for  $0 < x < L(0)$ .

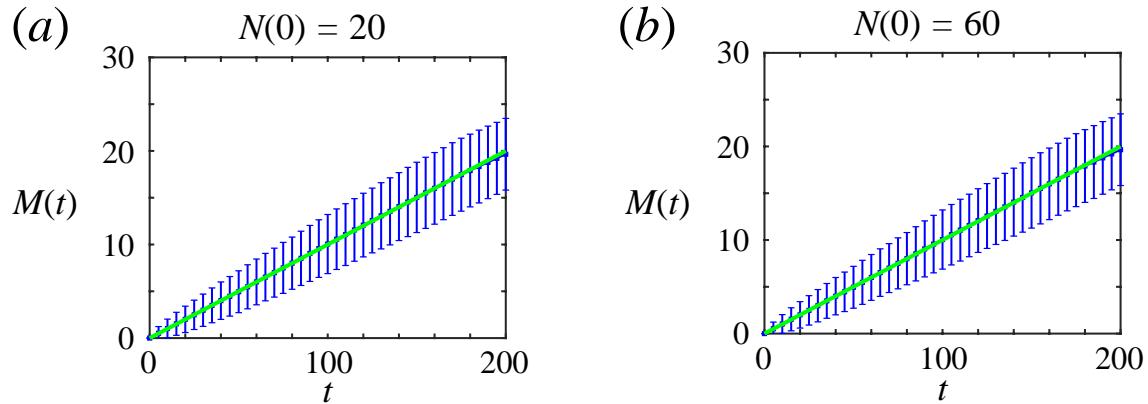
## 5A.4 Additional results

### 5A.4.1 Counting the total number of cells that detach

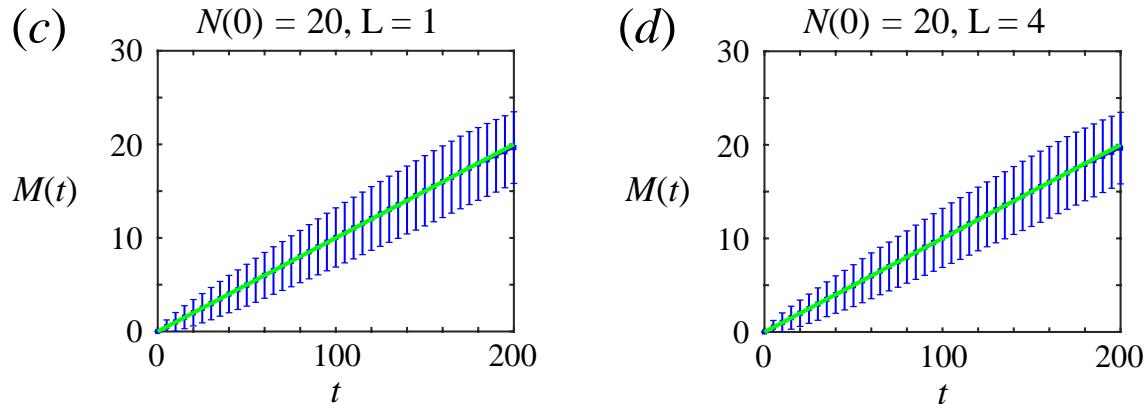
We first define  $M(t)$  as the total number of cells which have detached by time  $t$ . For chemically-independent EMT, cell detachment occurs at a constant rate  $\omega$ . Therefore,  $M(t) = \omega t$  while  $N(t) > 0$ , i.e. the total number of cells which have detached increases linearly with time while the tissue still contains cells. If  $N(t)$  reaches zero then  $M(t)$  plateaus. In Figures 5A.4(a)-(d), we show the good agreement between the results of the average of many discrete realisations and  $M(t) = \omega t$ . This holds for later times in Figure 5A.4(b) as the tissue does not go extinct. In contrast for Figures 5A.4(a),(c),(d)  $M(t)$  eventually plateaus at later times due to extinction.

For chemically-dependent EMT and  $D = 10^{-5}$  in Figure 5A.4(e),  $M(t)$  is very similar to the chemically-independent EMT case due to how parameters are chosen, as discussed in Section 5.4.3. Also note that as  $\phi = 0.9$  in Figure 5A.4(e) the cells require some time to reach the chemical threshold before detaching rapidly. Hence, we observe reduced noise for chemically-dependent EMT with  $D = 10^{-5}$  in comparison to results for chemically-independent EMT in Figures 5A.4(a)-(d). As  $\phi \rightarrow 0$  the noise in  $M(t)$  will increase, whereas for  $\phi = 1$  there will be no noise in  $M(t)$ . For chemically-dependent EMT with higher diffusivities of  $D = 10^{-2}$  and  $D = 1$ , Figures 5A.4(f),(g), respectively, cell detachment is initially delayed, then cells detach rapidly until extinction occurs.

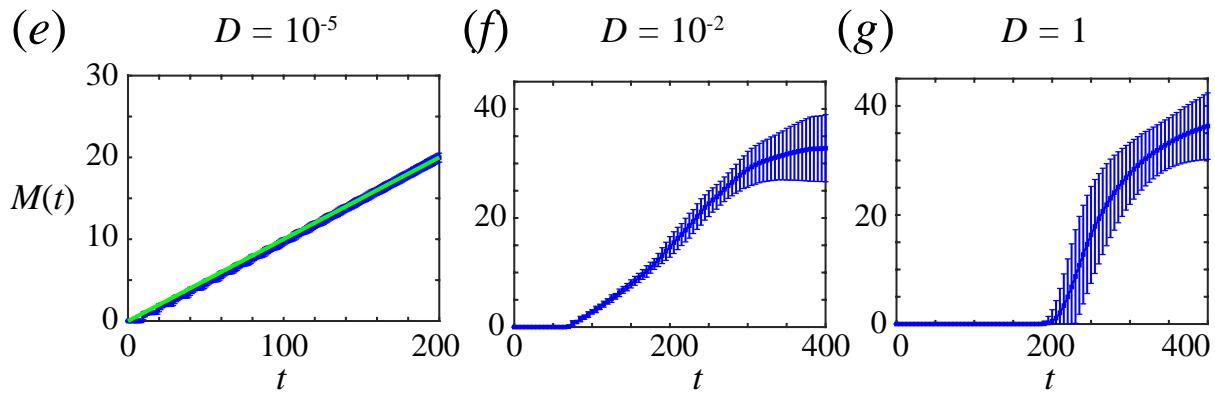
### 3.1 Chemically-independent EMT and cell-length-independent proliferation



### 3.2 Chemically-independent EMT and cell-length-dependent proliferation



### 3.3 Chemically-dependent EMT and cell-length-independent proliferation

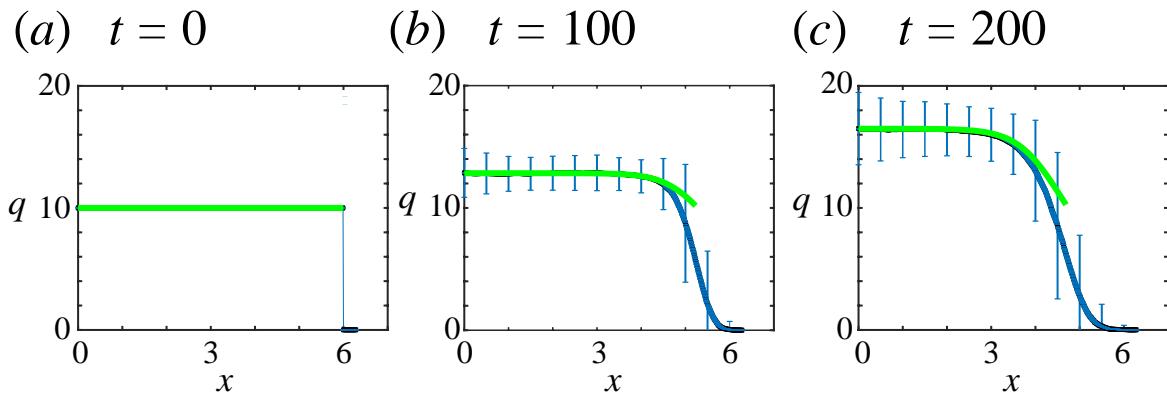


**Figure 5A.4:** Evolution of total number of cells that detach,  $M(t)$ , for examples presented in the main manuscript. (a)-(b) Chemically-independent EMT and cell-length independent proliferation (from Section 5.4.1). (c)-(d) Chemically-independent EMT and linear cell-length-dependent EMT (from Section 5.4.2). (e)-(g) Chemically-dependent EMT and cell-length-independent proliferation for (e)  $D = 10^{-5}$ , (f)  $D = 10^{-2}$ , (g)  $D = 1$  (from Section 5.4.3-5.4.4). The average of 2000 discrete realisations (blue) are compared with  $M(t) = \omega t$  (green).

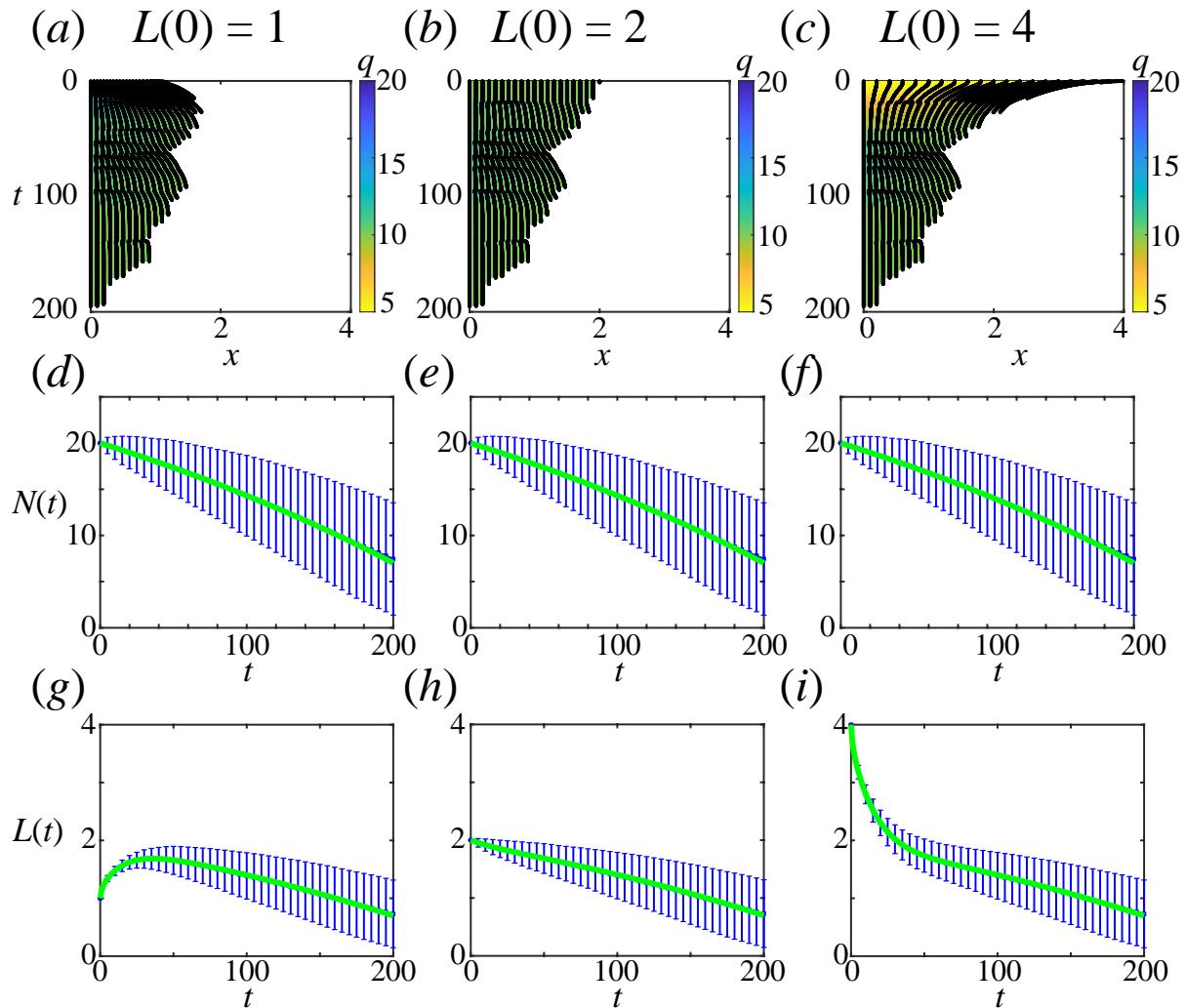
### 5A.4.2 Cell-length-independent proliferation

In the main manuscript we present results for the evolution of  $N(t)$  and  $L(t)$  with  $k = 1$  starting with cells initially at mechanical equilibrium. In Figures 5A.6 and 5A.7 we present results for  $k = 10$  and tissues that are initially compressed or stretched. Good agreement is observed between the continuum model and the average of many discrete realisations.

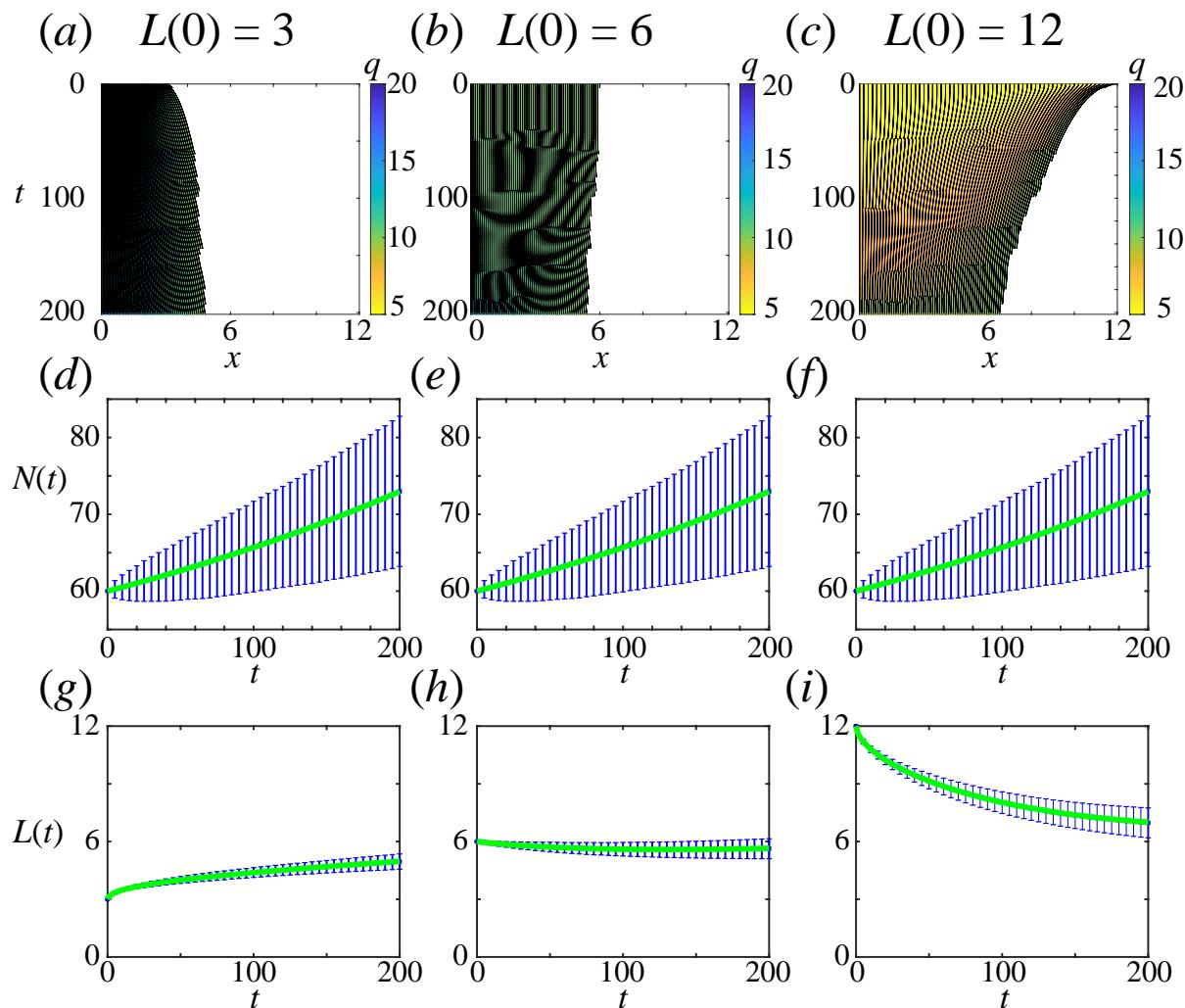
In Figures 5A.5 and 5A.8 we compare snapshots of the density,  $q(x, t)$ , from the continuum model to the average of many discrete realisations and observe good agreement. Differences at the free boundary  $x = L(t)$  are a result of the length in each discrete realisation being different due to the stochastic proliferation and cell detachment mechanisms.



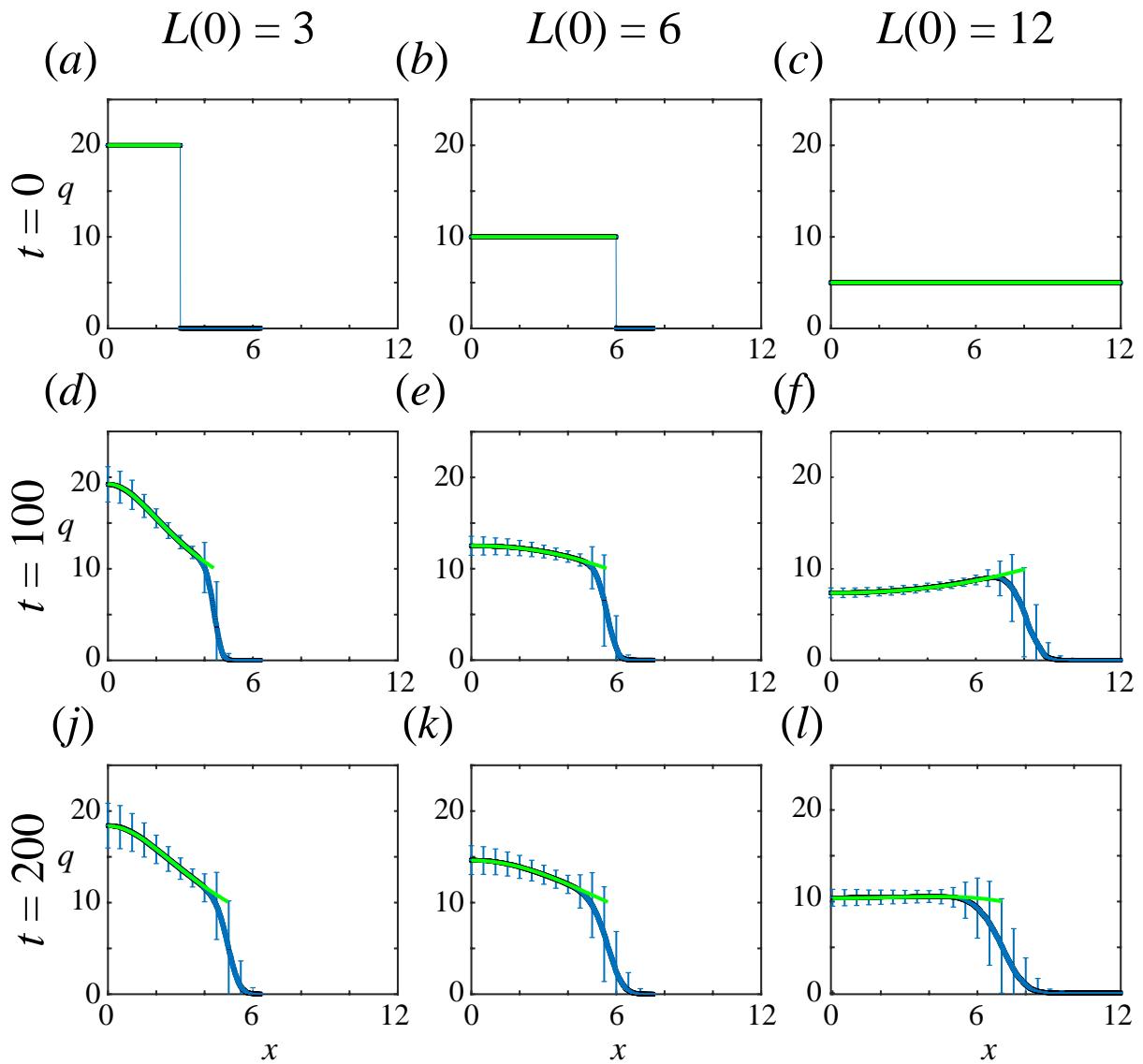
**Figure 5A.5:** Density snapshots corresponding to  $N(0) = 60$  in Figure 5.3. Cells are initially at their resting cell lengths. The average of 2000 discrete realisations (blue) are compared with the continuum model (green) at times (a)  $t = 0$ , (b)  $t = 100$ , (c)  $t = 200$ . Mechanical parameters:  $k = 1$ ,  $a = 0.1$ ,  $\eta = 1$ .



**Figure 5A.6:** Cell-length-independent proliferation and chemically-independent cell detachment for  $N(0) = 20$  and  $k = 10$  with initial conditions: (a),(d),(g) compressed, (b),(e),(h) mechanical equilibrium, (c),(f),(i) stretched. (a)-(c) Kymographs with density,  $q(x, t)$ , colouring. The average of 2000 discrete realisations (blue) are compared with the continuum model (green). (d)-(f) Evolution of total cell number,  $N(t)$ . (g)-(i) Evolution of tissue length,  $L(t)$ . Mechanical parameters:  $k = 10$ ,  $a = 0.1$ ,  $\eta = 1$ .



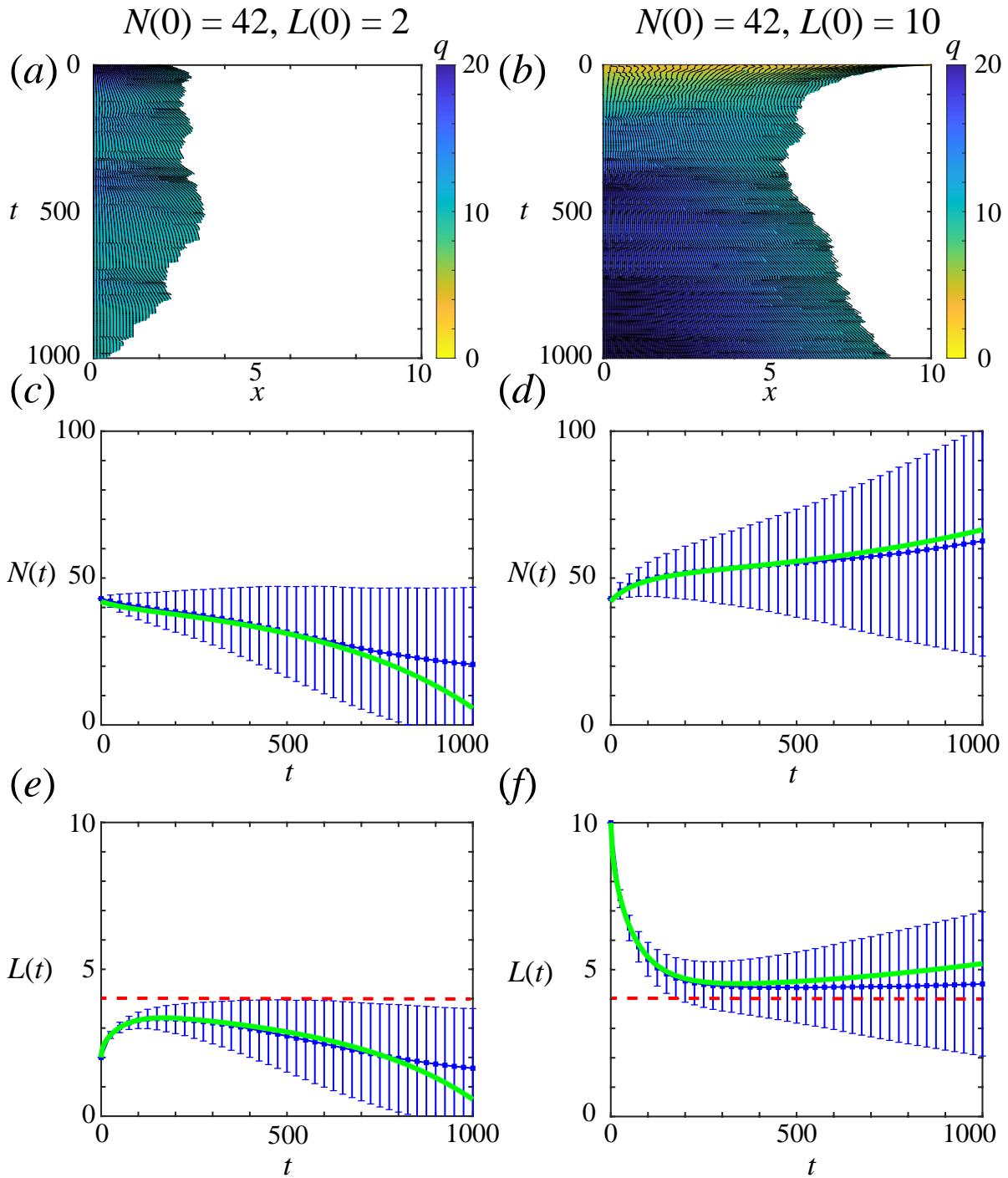
**Figure 5A.7:** Cell-length-independent proliferation and chemically-independent cell detachment for  $N(0) = 60$  and  $k = 10$  with initial conditions: (a),(d),(g) compressed, (b),(e),(h) mechanical equilibrium, (c),(f),(i) stretched. (a)-(c) Kymographs with density,  $q(x, t)$ , colouring. The average of 2000 discrete realisations (blue) are compared with the continuum model (green). (d)-(f) Evolution of total cell number,  $N(t)$ . (g)-(i) Evolution of tissue length,  $L(t)$ . Mechanical parameters:  $k = 10$ ,  $a = 0.1$ ,  $\eta = 1$ .



**Figure 5A.8:** Density snapshots corresponding to Figure 5A.7. The average of 2000 discrete realisations (blue) are compared with the continuum model (green) at times  $t = 0, 100, 200$  for initial tissue lengths,  $L(0) = 3, 6, 12$ . Mechanical parameters:  $k = 10, a = 0.1, \eta = 1$ .

### 5A.4.3 Cell-length-dependent proliferation

Previously, with cell-length-independent proliferation starting with  $N(0) = 42$  leads to unbounded growth in the continuum model, and mostly unbounded growth but sometimes extinction for realisations of the discrete model. Now we consider cell-length-dependent proliferation, where the initial tissue length can influence the long-term behaviour. Here we also assume chemically-independent cell detachment. Now with  $N(0) = 42$ , if the tissue is initially compressed then extinction is more likely (Figure 5A.9a,c,e) whereas if the tissue is initially stretched then the tissue is more likely to eventually grow without bound (Figure 5A.9(b),(d),(f)). Good agreement is observed between the average of many discrete realisations and the continuum description with differences due to being close to extinction (Figures 5A.9(a),(c),(e)) and due to some discrete realisations crossing the critical length threshold while others do not (Figures 5A.9(b),(d),(f)).



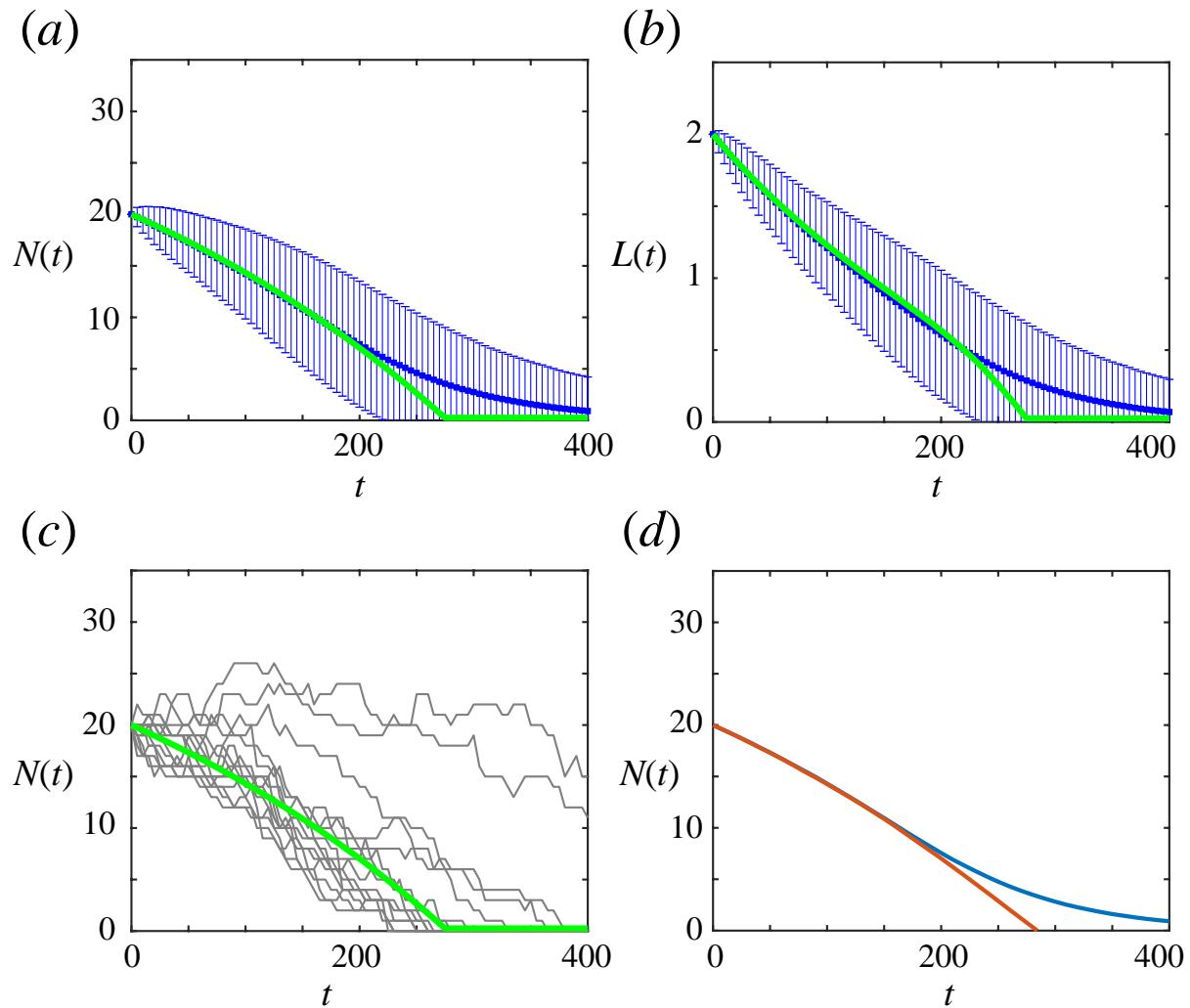
**Figure 5A.9:** Chemically-independent cell detachment with linear cell-length-dependent proliferation mechanism. Two initial cell populations with  $N(0) = 42$ , the first uniformly compressed with  $L(0) = 2$  and the second uniformly stretched with  $L(0) = 10$  (a)-(b) Kymographs with density,  $q(x, t)$ , colouring. The average of 2000 discrete realisations (blue) are compared with the continuum model (green). (c)-(d) Evolution of total cell number,  $N(t)$ . (e)-(f) Evolution of tissue length,  $L(t)$ . Red dashed line in (e)-(f) corresponds to the critical tissue length,  $\omega/\beta$ . (e)-(f) Evolution of total cell number,  $N(t)$ . Mechanical parameters:  $k = 1, a = 0.1, \eta = 1$ .

#### 5A.4.4 Differences between the cell-based and continuum models: low $N(t)$

In the main manuscript, we state that the average of many cell-based realisations does not agree with the solution of the continuum model when  $N(t)$  is low and close to extinction. Here we provide further explanation.

In Figure 5A.10(a), we extend the results from Figure 5.3 to  $t = 400$ . We observe that after approximately  $t = 200$  there is a difference between the average of many discrete realisations and the continuum model. Similar behaviour is observed for the evolution of  $L(t)$  in Figure 5A.10(b). These differences do not reduce when more simulations are performed. In Figure 5A.10(c) we compare the solution of the continuum model with fifteen realisations of the discrete model and observe that many discrete realisations go extinct before the continuum model reaches  $N(t) = 0$ .

In Figure 5A.10(d) we simulate only the total cell number of the discrete model, which evolves stochastically according to  $N(t + dt) = \beta N(t) - \omega$ , with two methods: i) whenever  $N(t) = 0$  in an individual realisation it is stopped; ii) allow  $N(t) < 0$ . We find that if we allow  $N(t) < 0$  then the discrete model matches results from the continuum model. However, allowing  $N(t) < 0$  is physically unrealistic. Instead, stopping individual realisations when  $N(t) = 0$  is physically realistic. Therefore, we must accept that the continuum model does not faithfully replicate the behaviour of the discrete model near extinction. Hence, the discrete model should be used when considering populations with low  $N(t)$ .

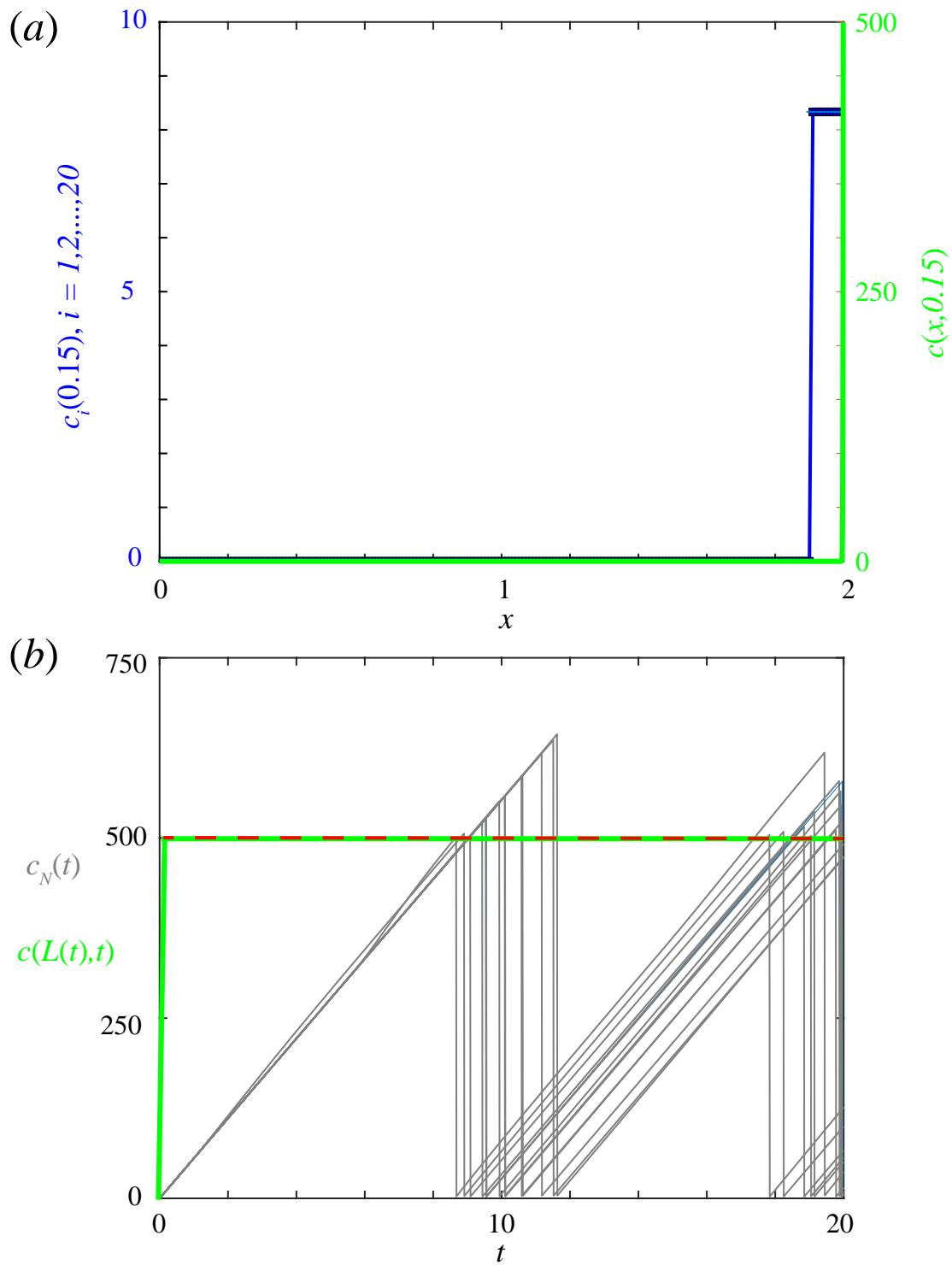


**Figure 5A.10:** Differences near extinction between discrete and continuum results for total cell number and interface position. The average of 2000 discrete realisations (blue) are compared with the continuum model (green). (a) Evolution of total cell number,  $N(t)$ . (b) Evolution of tissue length,  $L(t)$ . (c) Evolution of number of cells,  $N(t)$ , with 15 discrete realisations compared to solution from the continuum model. Mechanical parameters:  $k = 10, a = 0.1, \eta = 1$ . (d) Difference between discrete realisations stopped when  $N(t) = 0$  (blue), which is physically realistic, and those where  $N(t) < 0$  was allowed which is physically unrealistic (red).

### 5A.4.5 Chemically-dependent EMT: cell-length-independent proliferation

**Small diffusivity,  $D = 10^{-5}$ .**

In Figure 5.6(d),(g) of the manuscript we observe that there is a difference between the average of many discrete realisations and the continuum model. We explain that this is due to the well-mixed assumption for chemical concentration inside cells not being valid for small diffusivities. We show this in Figure 5A.11. In Figure 5A.11(a) we compare a snapshot of the chemical concentration from the continuum model and a discrete realisation at very early time,  $t = 0.15$ . In Figure 5A.11(b) we compare the  $c(L(t), t)$  from the continuum model with  $c_N(t)$  from fifteen realisations of the discrete model, as these are used to calculate the rate of cell detachment. These results show that the continuum model reaches the concentration threshold much earlier than the discrete model which causes the number of cells in the continuum model to reduce faster than in the discrete model, hence explaining the difference in Figure 5.6(d),(g).

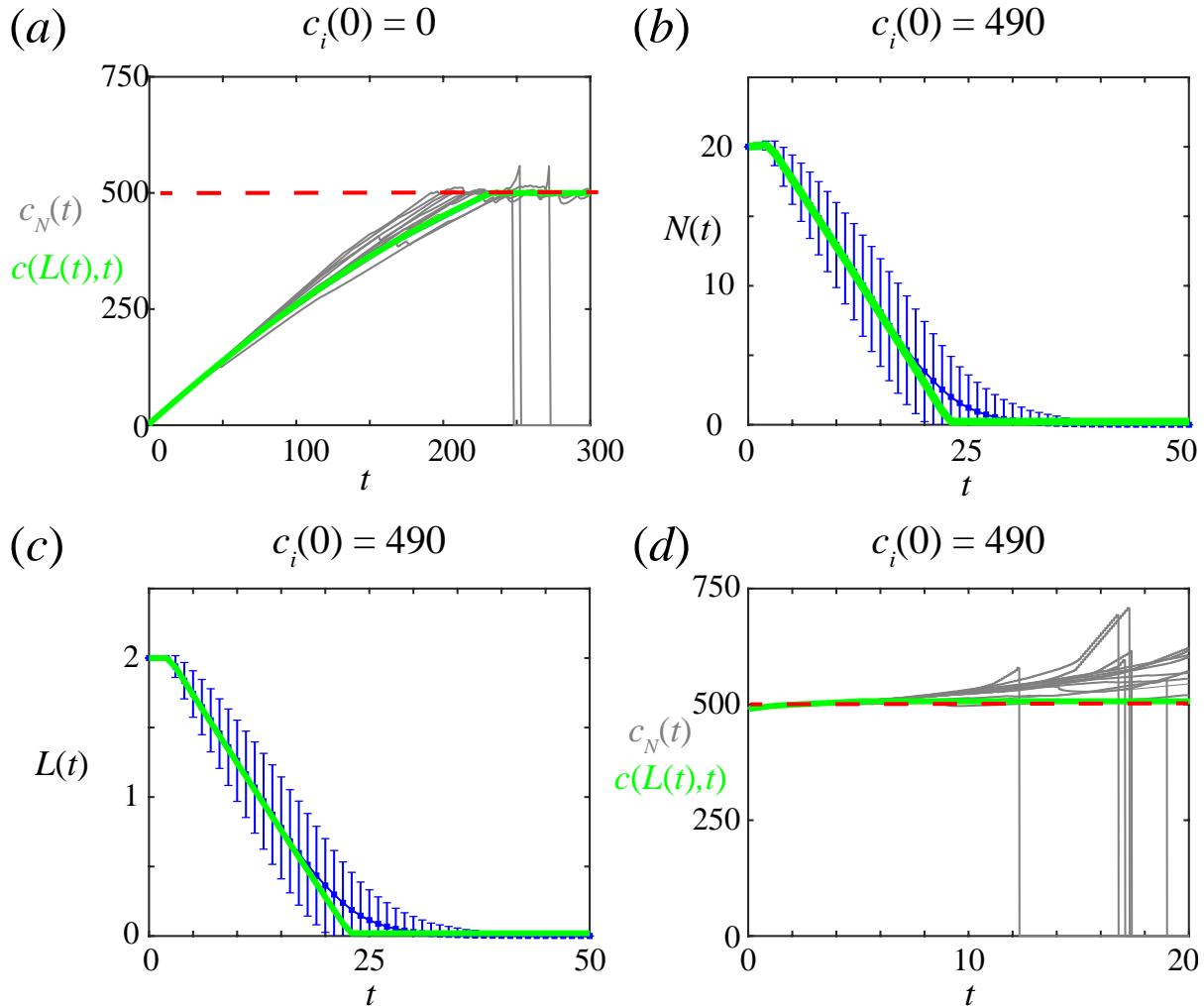


**Figure 5A.11:** Continuum model (green) reaches concentration threshold (red-dashed) faster than individual realisations of the discrete model (grey). Chemically-dependent EMT with  $D = 10^{-5}$  and with cell-length-independent proliferation. (a) Concentration snapshot from the continuum model at time  $t = 0.15$ ,  $c(x, 0.15)$  for  $0 < x < L(1)$ , compared to one realisation of the discrete model, where the chemical concentration in cell  $i$  is  $c_i(0.15)$  for  $i = 1, 2, \dots, N$ . (b) Chemical concentration at the boundary node of the continuum model,  $c(L(t), t)$ , compared to the chemical concentration of the boundary cell,  $c_N(t)$ , for fifteen realisations of the discrete model. Here  $N(0) = 20$ . Mechanical parameters:  $k = 1, a = 0.1, \eta = 1$

**Starting closer to the chemical threshold with  $D = 1$** 

In Figure 5.6(f),(i) of the manuscript we observe that there is a difference for  $D = 1$ . Here, we show that this difference is due to stochastic effects in realisations of the discrete model. Each discrete realisation has a different tissue length,  $L(t)$ , resulting in the concentration threshold being reached at different times, which does not occur in the continuum model (Figure 5A.12(a)).

If instead, we compare results when starting close to the chemical threshold,  $C$ , we find an improved match. Specifically, starting with  $D = 1$  and  $c_i = 490$  for  $i = 1, 2, \dots, 20$ , and  $C = 500$ , rather than with  $c_i = 0$  for  $i = 1, 2, \dots, 20$ , we find an improved match (Figure 5A.12(b)-(c)). We observe that different realisations of the discrete model reach the concentration threshold,  $C = 500$ , at approximately the same time as the continuum model (Figure 5A.12(d)). This is due to the reduced time for stochastic effects in  $N(t)$  and  $L(t)$  to play a role.

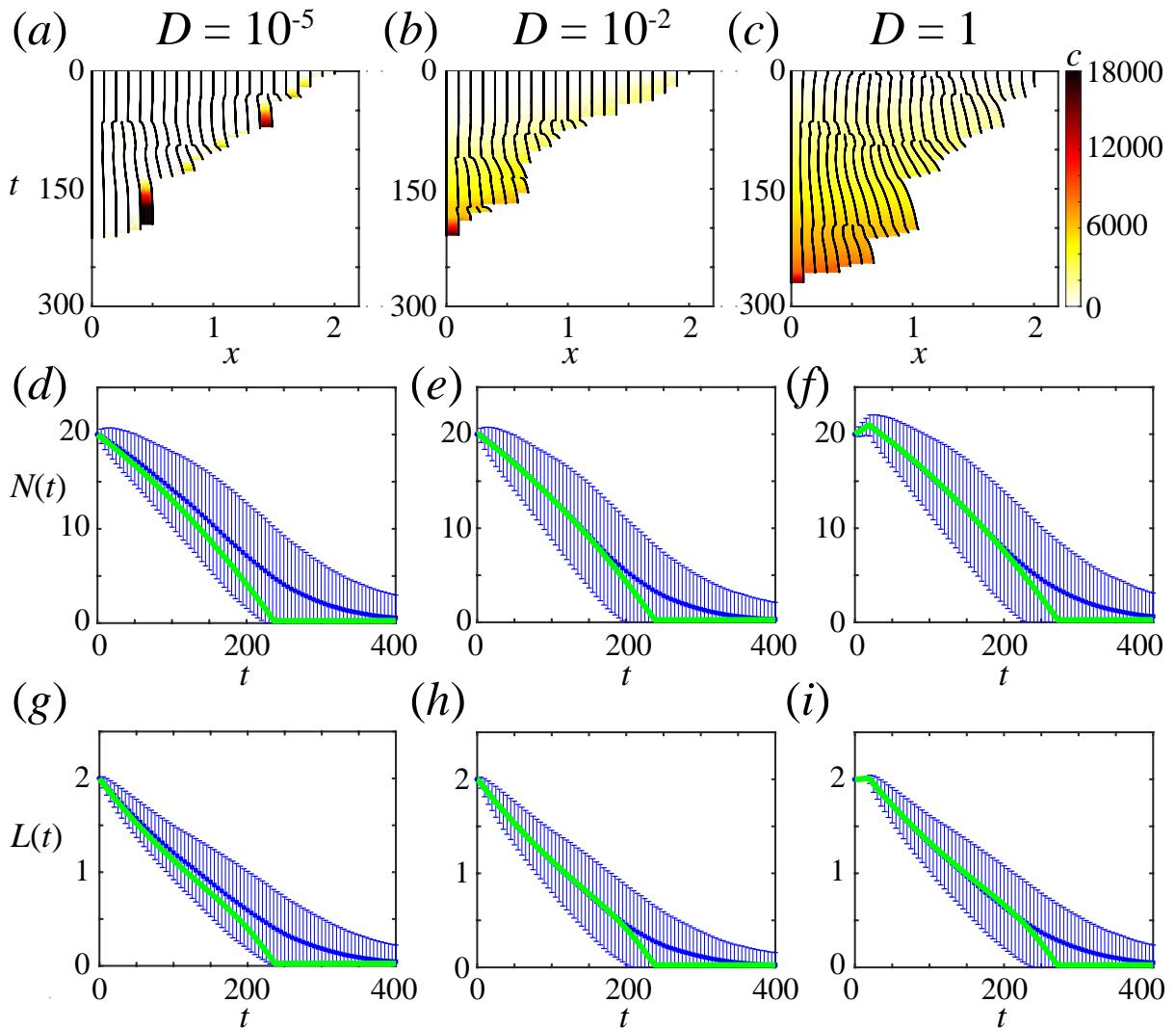


**Figure 5A.12:** Comparison of results from continuum model and the average of many realisations of the discrete model for chemically-dependent EMT with  $D = 1$  and cell-length-independent proliferation. (a) Evolution of the chemical concentration in the final cell,  $c_N(t)$ , when  $c_i(0) = 0$  for  $i = 1, 2, \dots, 20$ . Some discrete realisations reach the concentration threshold,  $C = 500$  (red-dashed), earlier than the continuum model. (b)-(c) The average of 2000 discrete realisations (blue) are compared with the continuum model (green). (b) Evolution of total cell number,  $N(t)$ . (c) Evolution of tissue length,  $L(t)$ . (d) Evolution of the chemical concentration in the final cell,  $c_N(t)$ , when  $c_i(0) = 490$  for  $i = 1, 2, \dots, 20$ . Realisations of the discrete model reach the concentration threshold,  $C = 500$  (red-dashed), at approximately the same time as the continuum model. Fifteen individual realisations of the discrete model are shown in (a) and (d). Mechanical parameters:  $k = 1$ ,  $a = 0.1$ ,  $\eta = 1$ .

### 5A.4.6 Sensitivity to $\phi$

Chemically-dependent cell detachment is a two-step process: i) the boundary cell gains an invasive phenotype when the chemical concentration inside the boundary cell is above the chemical threshold,  $C$ ; ii) the boundary cell detaches. We introduce a parameter  $\phi \in [0, 1]$  which defines the ratio of the average time in process i) as  $\phi/\omega$  and the average time in process ii) as  $(1-\phi)/\omega$  (Figure 5.2). In the manuscript we present results for  $\phi = 0.9$  (Figure 5.6). Here, in Figure 5A.13 we present results for  $\phi = 0.1$ . As before, we choose parameters so that the average rate of cell detachment is the same as in previous models. To do this we keep the chemical threshold,  $C$ , fixed and vary the constant number of molecules per unit time supplied to the boundary cell from the external environment,  $S$ . Results for  $\phi = 0.1$  show improved agreement between the continuum model and the average of many discrete realisations in comparison to  $\phi = 0.9$ . This is because the time to reach the chemical threshold is quicker and the time to stochastically detach is longer. Results for  $D = 10^{-5}$  (Figures 5A.13(a),(d),(g)) and  $D = 10^{-2}$  (Figures 5A.13(b),(e),(h)) look similar as they both reach the chemical threshold after a very short time. Note that when  $\phi = 0$  the chemically-dependent model is the same as the chemically-independent model.

In the above, we keep  $C$  fixed and vary  $S$ . Alternatively, one could vary  $C$  and keep  $S$  fixed, or vary both  $C$  and  $S$ . Furthermore, one could also keep both  $C$  and  $S$  fixed, and varying the rate of cell detachment. However, we do not show this here as then the average time for the boundary cell to detach is not the same as the average time to detach for previous models, which would not be a fair comparison.

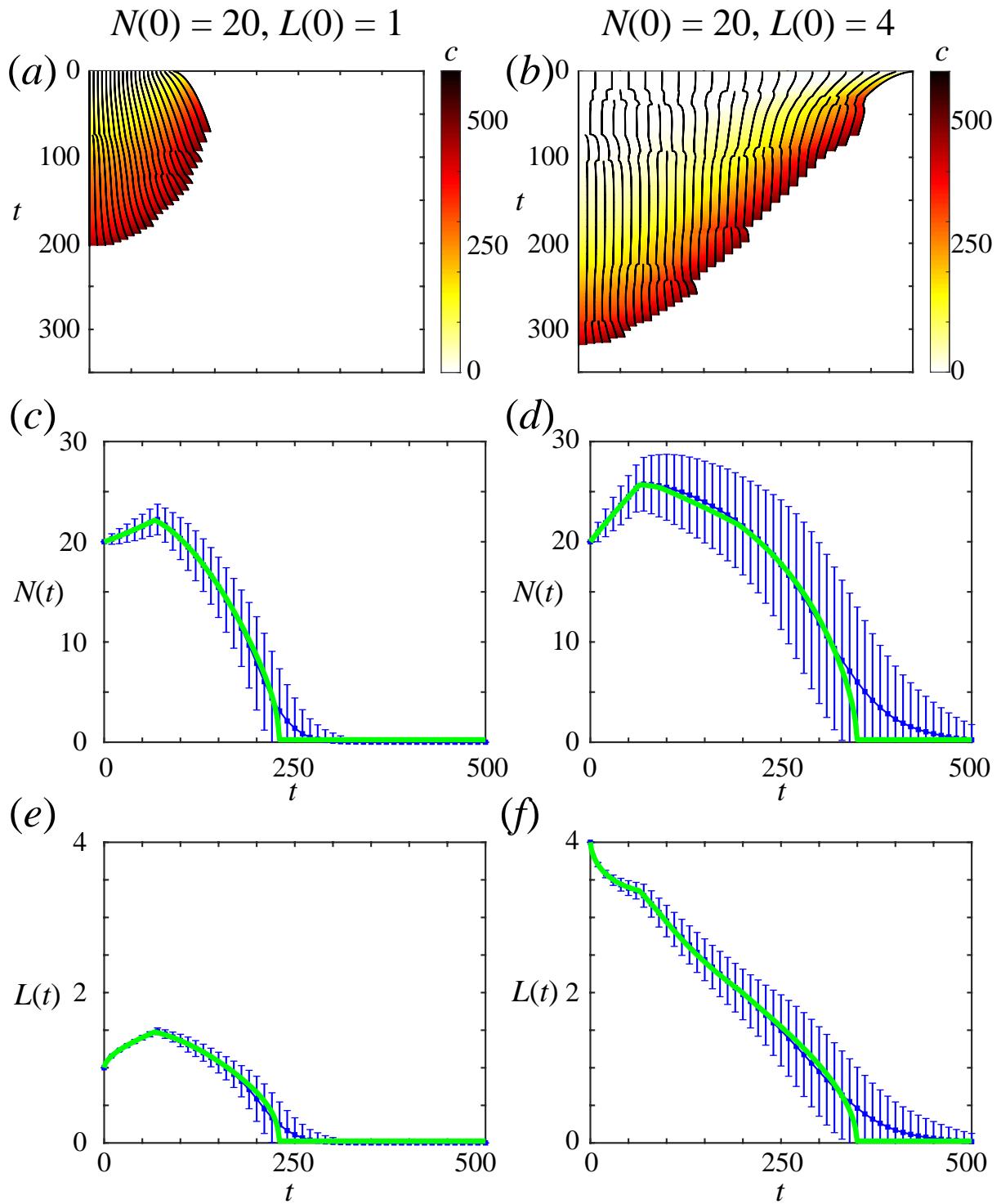


**Figure 5A.13:** Sensitivity to  $\phi$ . Figure 5.6, exploring diffusion delaying first EMT event, repeated with  $\phi = 0.1$ . Cell detachment driven by chemically-dependent EMT with varying diffusivities and cell-length-independent proliferation mechanism. Cells initially at their resting cell lengths with initial cell populations  $N(0) = 20$ . Kymographs with chemical concentration,  $c(x, t)$ , colouring shown for (a)  $D = 10^{-5}$ , (b)  $D = 10^{-2}$ , (c)  $D = 1$ . (d)-(i) The average of 2000 discrete realisations (blue) are compared with the continuum model (green). (d)-(f) Evolution of total cell number,  $N(t)$ . (g)-(i) Evolution of tissue length,  $L(t)$ . Mechanical parameters:  $k = 1, a = 0.1, \eta = 1$ .

### 5A.4.7 Chemically-dependent EMT: linear proliferation

In the manuscript we present results for cell detachment driven by chemically-dependent EMT and a cell-length-independent proliferation mechanism. Here we present results with a linear cell-length-dependent proliferation mechanism.

In Figure 5A.14 we observe that when comparing an initially compressed tissue to an initially stretched tissue, with the same  $N(0)$ , the time to reach the chemical threshold and the time for all cells to detach is shorter for the initially compressed tissue. This is due to proliferation being less likely in the compressed tissue with cell-length-dependent proliferation. This is also due to the chemical reaching the concentration threshold faster in the compressed tissue than in the initially stretched tissue, as there is less space for the chemical to diffuse in the compressed tissue.



**Figure 5A.14:** Cell detachment driven by chemically-dependent EMT with  $D = 10^{-2}$  and with linear cell-length-dependent proliferation. Two initial cell populations with  $N(0) = 20$ , the first uniformly compressed with  $L(0) = 1$  and the second uniformly stretched with  $L(0) = 4$ . (a)-(b) Kymographs with density,  $q(x,t)$ , colouring. (c)-(f) The average of 2000 discrete realisations (blue) are compared with the continuum model (green). (c)-(d) Evolution of total cell number,  $N(t)$ . (e)-(f) Evolution of tissue length,  $L(t)$ . Mechanical parameters:  $k = 1, a = 0.1, \eta = 1$ .

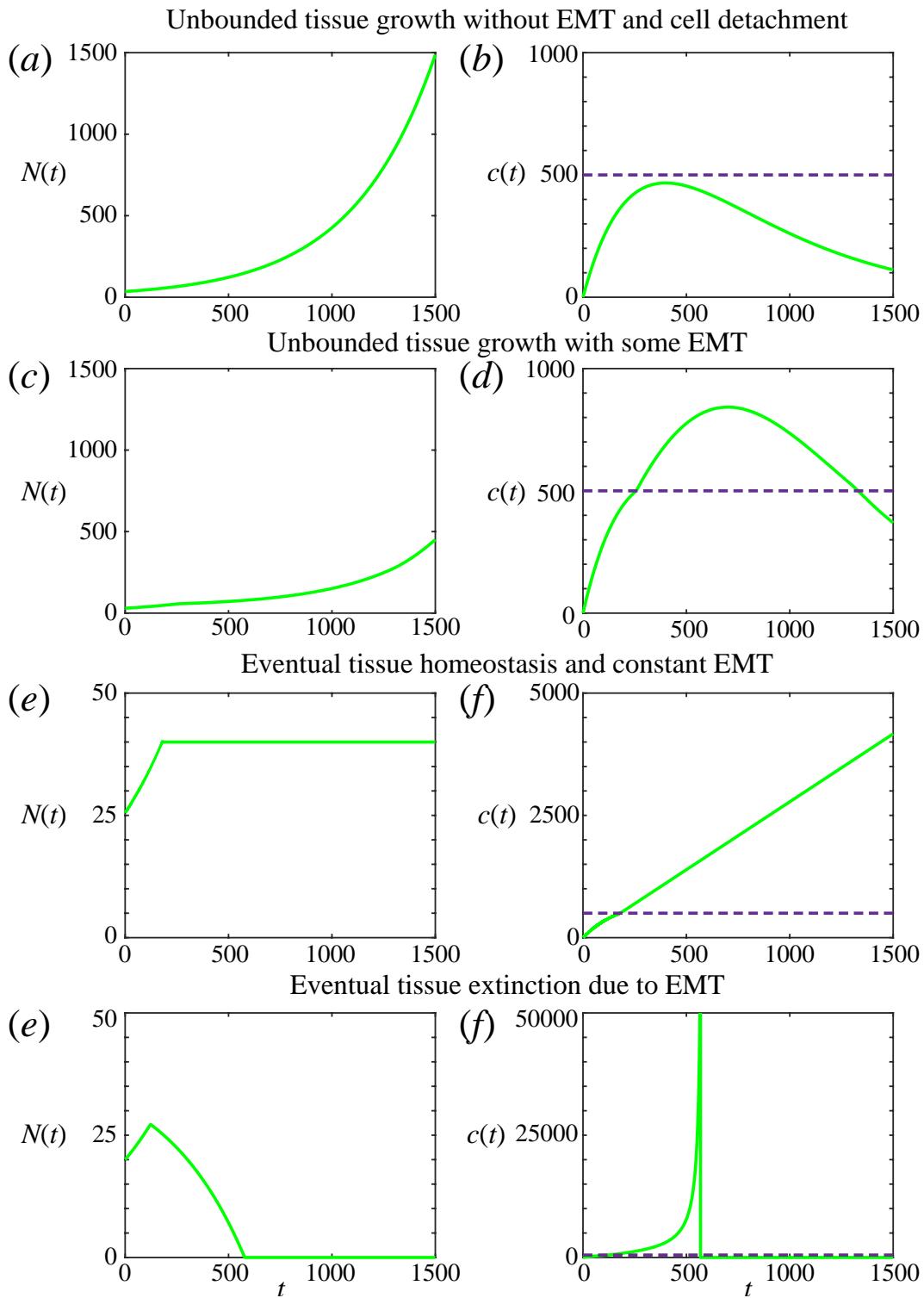
## 5A.5 Diffusive equilibrium at all times

If the diffusivity of the EMT-inducing chemical is very high it is reasonable to assume that the chemical in the tissue is at diffusive equilibrium at all times. This simplifies the analysis as every cell in the tissue experiences the same concentration at time  $t$ , which we denote  $c(t)$ . This can be useful to understand possible long-term behaviours.

To proceed, we make a further assumption that cells are always at mechanical equilibrium and consider the continuum model. As cells are at mechanical equilibrium, the cell-length-independent and cell-length-dependent proliferation mechanisms are equivalent, so before any cell detachment events occur  $N(t) = N(0) \exp(\beta t)$ ,  $L(t) = N(t)a$ , and  $c(t) = St/L(t)$ . Then for cell detachment to occur at least once we require that  $c(t) > C$  which is equivalent to requiring  $N(0) < S/(Ca\beta \exp(1)) = N_I$ .

So if  $N(0) > N_I$  the tissue grows without bound and there is no EMT and no cell detachment, and  $N(t)$  evolves according to  $dN(t)/dt = \beta N(t)$ . However, if  $N(0) < N_I$  the time to reach the concentration threshold,  $t_C$ , is the solution of  $t_C/\exp(\beta t_C) = CN(0)a/S$ . While  $c(t) > C$ ,  $N(t)$  evolves according to  $dN(t)/dt = \beta N(t) - \phi\omega$  and if the concentration decreases below  $C$  then  $N(t)$  evolves according to  $dN(t)/dt = \beta N(t)$ . This results in four possible behaviours:

1. unbounded tissue growth without EMT and cell detachment (Figure 5A.15(a),(b));
2. unbounded tissue growth with some EMT (Figure 5A.15(c),(d));
3. eventual tissue homeostasis and constant EMT (Figure 5A.15(e),(f));
4. eventual tissue extinction due to EMT (Figure 5A.15(g),(h)).



**Figure 5A.15:** For diffusive equilibrium at all times and instantaneous mechanical relaxation there are four possible behaviours. (a) Unbounded tissue growth without EMT and cell detachment. (b) Unbounded tissue growth with some EMT. (c) Eventual tissue homeostasis and constant EMT. (d) Eventual tissue extinction due to EMT. Purple dashed line corresponds to chemical threshold  $C$ . Results shown for  $S = 50/9, C = 500, \phi = 0.9, k = 1, a = 0.05, \beta = 0.0025, \omega = 0.1$ .

## **Part II**

# **Experimental design and mathematical modelling**



## **Chapter 6**

# **Designing and interpreting 4D tumour spheroid experiments**

## 6.0 Preamble

### An article under consideration at *Nature Communications*

**Murphy RJ**, Browning AP, Gunasingh G, Haass NK, Simpson MJ (2021). Designing and interpreting 4D tumour spheroid experiments. *Under consideration at Nature Communications*. bioRxiv preprint.

In this chapter we transition to Part 2 of this thesis exploring how mathematical modelling can improve experimental designs, in particular for tumour spheroid experiments that are routinely performed to study cancer progression and treatment. We address objective 5 and research question 5. We focus on avascular tumour growth and directly quantitatively connect experimental data, that I collect in the wet laboratory (Figure 6.0), to mathematical modelling using statistical analysis. Since many mathematical models have already been proposed to study avascular tumour growth, but few have been experimentally validated, we use the seminal Greenspan mathematical model [79]. Greenspan's model, due to a spherical symmetry assumption, describes the growth of the tumour with equations with one spatial dimension, namely the tumour radius. Note that models in Chapters 2-5 also have one spatial dimension. Also, similarly to Part 1 of this thesis, Greenspan's model is a mechanochemical model. However, in Greenspan's model mechanical interactions are assumed to maintain the tumour as a compact solid mass, rather than being explored explicitly as in Chapters 2-5. Furthermore, chemical diffusion is important in Greenspan's model to determine the time-evolution of the tumour internal structure, whereas in Chapter 5 we explore chemical diffusion in relation to epithelial-mesenchymal transitions.

Key results for this chapter include: performing tumour spheroid experiments with real-time cell cycle imaging to collect an abundance of experimental data across a range of experimental designs; verification of the Greenspan's mathematical model to that experimental data; and development of an objective mathematical modelling framework with statistical identifiability analysis to quantitatively compare experimental designs and identify design choices that produce reliable biological insight to provide recommendations for future studies.

Due to journal formatting requirements where this article is currently under consideration, the layout of this chapter is different to previous chapters. This chapter is structured as follows: Introduction; Results; Discussion; and then Methods. The Supplementary Material for this Chapter, included in Chapter 6A, is extensive including: further details of mathematical modelling methods; further details of statistical identifiability analysis methods; experimental data; additional results for different cell lines; and additional results with synthetic data.



**Figure 6.0:** To perform the tumour spheroid experiments to collect the experimental data for this chapter I was trained, starting from no experience in a laboratory prior to this PhD. I performed tumour spheroid experiments from start to finish including: cell culturing; spheroid formation; spheroid harvesting; spheroid fixing and mounting in preparation for imaging; confocal microscopy; and, image acquisition, processing, and analysis. Images show me in the laboratory during the spheroid formation stage of an experiment (further details are shown in Section 6.5.3 Experimental methods and the experimental protocol is detailed in [218]).

### Statement of Contribution of Co-Authors for Thesis by Published Paper

The authors listed below have certified that:

1. they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
2. they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
3. there are no other authors of the publication according to these criteria;
4. potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit, and
5. they agree to the use of the publication in the student's thesis and its publication on the QUT's ePrints site consistent with any limitations set by publisher requirements.

In this case of this chapter:

Murphy RJ, Browning AP, Gunasingh G, Haass NK, Simpson MJ (2021). Designing and interpreting 4D tumour spheroid experiments. biorxiv preprint

Contributor	Statement of Contribution
Ryan J. Murphy (Candidate)  <i>R. Murphy</i> 02/09/2021	Conceived and designed the study, performed biological experiments, developed the codes for numerical simulation continuum model and statistical analysis, performed numerical simulations, generated results, interpreted results, drafted the manuscript.
Alexander P. Browning	Performed biological experiments. Provided comments and gave final approval for publication.
Gency Gunasingh	Performed biological experiments. Provided comments and gave final approval for publication.
Nikolas K. Haass	Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research.
Matthew J. Simpson	Conceived and designed the study and provided comments and gave final approval for publication. Supervised the research.

#### Principal Supervisor Confirmation

I have sighted email or other correspondence from all Co-authors confirming their certifying authorship.

Matthew Simpson

2/9/2021

Name

Signature

Date

## 6.1 Abstract

Tumour spheroid experiments are routinely used to study cancer progression and treatment. Various and inconsistent experimental designs are used, leading to challenges in interpretation and reproducibility. Using multiple experimental designs, live-dead cell staining, and real-time cell cycle imaging, we measure necrotic and proliferation-inhibited regions in over 1000 4D tumour spheroids (3D space plus cell cycle status). By intentionally varying the initial spheroid size and temporal sampling frequencies across multiple cell lines, we collect an abundance of measurements of internal spheroid structure. These data are difficult to compare and interpret. However, using an objective mathematical modelling framework and statistical identifiability analysis we quantitatively compare experimental designs and identify design choices that produce reliable biological insight. Measurements of internal spheroid structure provide the most insight, whereas varying initial spheroid size and temporal measurement frequency is less important. Our general framework applies to spheroids grown in different conditions and with different cell types.

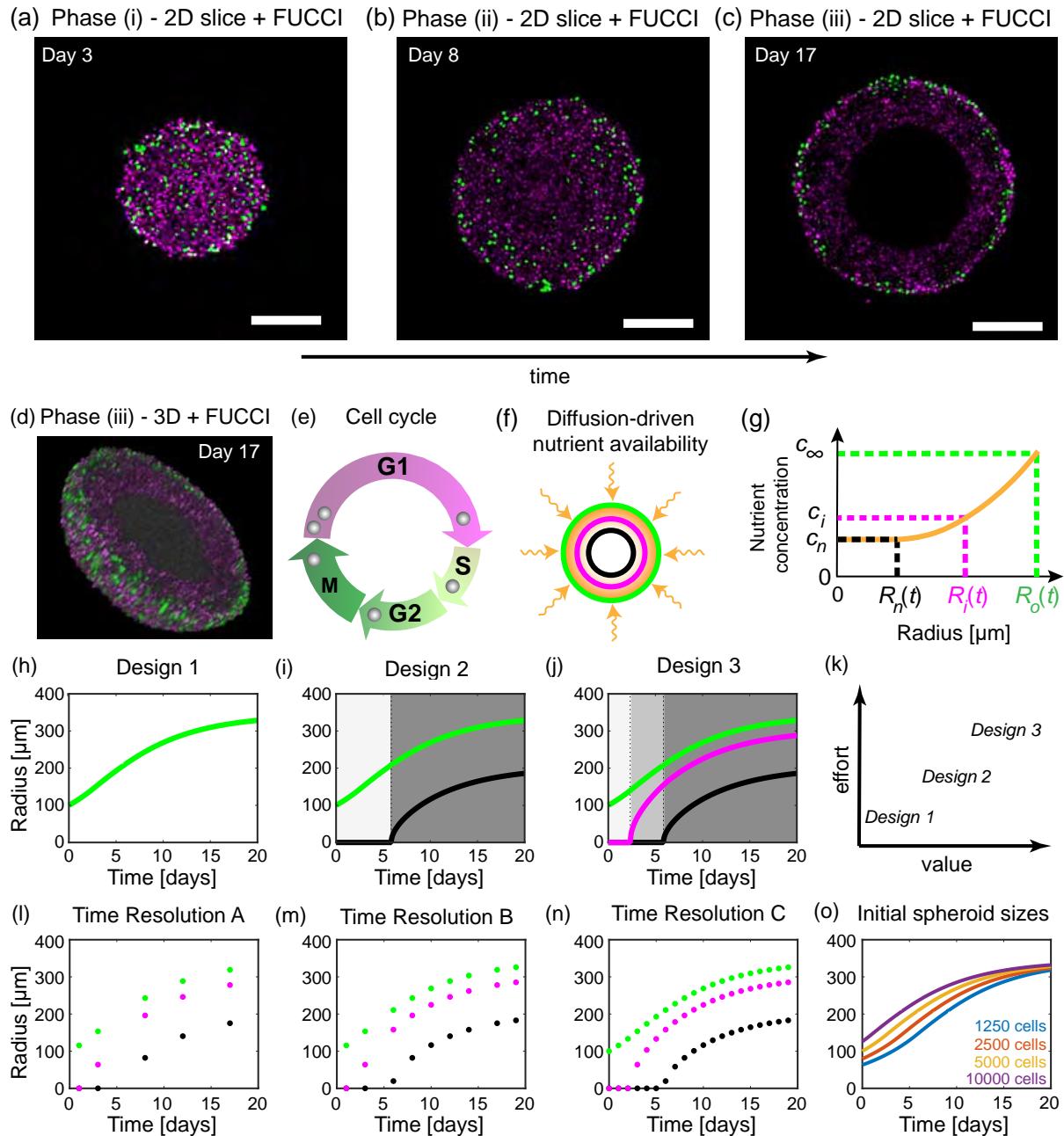
## 6.2 Introduction

Tumour spheroid experiments are an important *in vitro* tool routinely used since the 1970s to understand avascular tumour growth, cancer progression, develop cancer treatments, and reduce animal experimentation [46, 71, 99, 119, 150, 169, 197, 215, 217, 245]. However, a vast range of experimental designs are employed, leading to inconsistencies in: i) the times when measurements are taken; ii) experimental durations, ranging from a few days to over a month [8, 18, 83, 121, 168, 177]; iii) the initial number of cells used to form spheroids [8, 18, 54, 83, 88, 121, 168, 177], commonly between 300 to 20,000 cells [54, 168]; and, iv) the type of experimental measurements that are taken [8, 18, 54, 83, 88, 121, 168, 177]. This variability in experimental protocols makes comparing different studies very difficult, and introduces challenges in both interpretation and reproducibility of these experiments.

Mathematical modelling provides a powerful tool to provide such interpretation through model calibration and mechanism deduction. Simple mathematical models calibrated to outer radius measurements, such as Gompertzian growth models, have been used for decades to predict the growth of tumours [22, 173]. However, these simple mathematical models do not provide information about the internal spheroid structure over time. In response, many mathematical models of varying complexity have been developed to explore the internal structure of spheroids [7, 16, 17, 36, 37, 40, 65, 79, 85, 105, 106, 108, 117, 122, 135, 148, 149, 152, 191, 242, 243]. Here, we revisit the seminal Greenspan mathematical model for avascular tumour spheroid growth and quantitatively directly connect it to data for the first time [79]. Greenspan's mathematical model was the first to describe the three phases of avascular tumour spheroid growth: in phase (i) cells throughout the spheroid can proliferate; in phase (ii) cells near the periphery proliferate while a central region of living cells cannot proliferate, referred to as the inhibited region; and in phase (iii) there is an outer region of proliferative cells, an intermediate region of living inhibited cells, and a central necrotic region composed of dead cells and cellular material in various stages of disintegration and dissolution (Figure 6.1a-d, Methods 6.5.1). These various regions of cellular behaviour are thought to arise as a result of nutrient availability, such as oxygen, that is driven by diffusion and uptake.

In this study, we systematically explore a range of experimental designs and measurements. The first and simplest measurements we obtain are of the outer radius of the spheroid. Next, using live-dead cell staining we obtain measurements of the necrotic region. Measurements of the inhibited region are harder to obtain using traditional techniques. We use fluorescent ubiquitination-based cell cycle indicator (FUCCI) transduced cell lines [39, 87, 192, 237, 238, 251]. The nuclei of these cells fluoresce red when cells are in the gap 1 (G1) phase of

the cell cycle and green when cells are in the synthesis (S), gap 2 (G2) and mitotic (M) phases of the cell cycle (Figure 6.1e). For clarity, we choose to show cells in the gap 1 (G1) phase in magenta instead of red. These data are collected for human melanoma cell lines established from primary (WM793b) and metastatic cancer sites (WM983b, WM164) [87, 100, 213, 214], with endogenously low (WM793b) and high (WM983b, WM164) microphthalmia-associated transcription factor which is a master regulator of melanocyte biology [219]. Analysing these data provides real-time visualisation of the cell cycle throughout tumour spheroids and powerfully reveals the time evolution of the inhibited region (Figure 6.1a-d). This additional dimension of information that we capture in our experiments, namely the cell cycle status, together with the three-spatial dimensions of the tumour spheroid give rise to the term 4D tumour spheroid experiments. Given an abundance of measurements of the outer radius, inhibited radius, and necrotic radius for tumour spheroids across multiple initial spheroid sizes, time points, and cell lines, we calculate maximum likelihood estimates (MLE) and form approximate 95% confidence intervals for the parameters of the Greenspan model. This allows us to quantitatively elucidate how modifying experimental designs can extract more information from experiments. Furthermore, this approach identifies the experimental design choices that are important and lead to reliable biological insight.



**Figure 6.1:** Tumour spheroid growth and the Greenspan mathematical model. Tumour spheroids experience three phases of growth. (a)-(d) Confocal microscopy reveals different phases of tumour growth. Fluorescent ubiquitination-based cell cycle indicator (FUCCI) transduced cells allow visualisation of each cell's stage in the cell cycle. (a)-(c) 2D equatorial plane images of WM793b human melanoma tumour spheroids, formed with 5000 cells per spheroid, on days 3, 8, and 17 after formation. Scale bar 200 μm. (d) 3D representation of half of a WM793b human melanoma tumour spheroid on day 17 after formation, additional 3D representations are shown in supplementary material 6A.3.3. (e) Cell cycle schematic coloured with respect to FUCCI signal. (f) Schematic for Greenspan mathematical model. Nutrient diffuses within the tumour spheroid and is consumed by living cells. (g) Snapshot of nutrient concentration,  $c(r, t)$  for  $0 < r < R_o(t)$ , for a tumour spheroid in phase (iii) and where  $R_o(t)$  is the tumour spheroids outer radius. External nutrient concentration is  $c_\infty$ . Inhibited radius,  $R_i(t)$ , and necrotic radius,  $R_n(t)$ , are defined as the radius where the nutrient concentration first reaches the thresholds  $c_i$  and  $c_n$ , respectively. (h)-(j) Three experimental designs varying by measurement type. Design 1 considers only the outer radius (green). Design 2 considers the outer (green) and necrotic radius (black). Design 3 considers the outer (green), necrotic (black), and inhibited (magenta) radius. (k) Comparison of experimental designs with respect to their value and effort required. (l)-(n) Three experimental designs that vary due to the time resolution at which measurements are taken. (o) Four experimental designs that vary the number of cells used to form each spheroid.

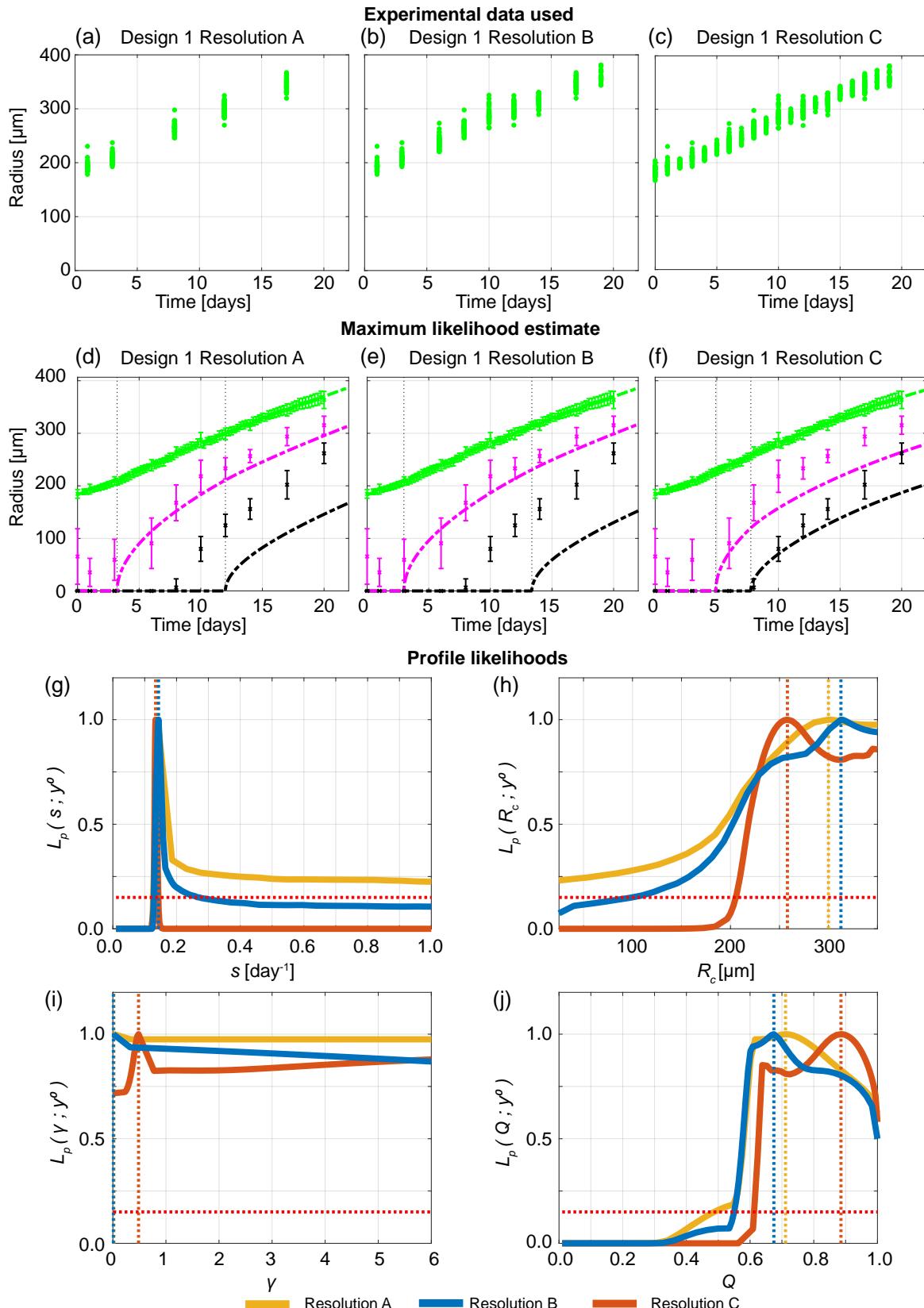
## 6.3 Results

The results in this main document are for spheroids formed with the WM793b human melanoma cell line [87, 100, 213, 214]. Additional results in Supplementary Material 6A.6 and 6A.7 show results for two other cell lines.

### 6.3.1 Outer radius measurements are not sufficient to predict inhibited and necrotic radii

Tumour outer radius measurements are simple to obtain and have been used for decades to quantify tumour growth [22, 173]. Modern technology enables these measurements to be obtained more frequently, easily, and accurately. For example, the IncuCyte S3 live cell imaging system (Sartorius, Goettingen, Germany) enables automated image acquisition and processing to measure spheroids every minute throughout an experiment providing a large number of measurements with ease. However, it is unclear whether these measurements provide sufficient information to understand and probe the internal structure of tumour spheroids and accurately predict tumour growth. Furthermore, it is unclear when measurements should be taken and the frequency of measurement. Performing experiments with WM793b spheroids formed with 5000 cells per spheroid, a typical choice in many experiments [83, 217–219], 24 spheroids are imaged every six hours. We monitor the time evolution of the outer radius to determine when spheroid formation ends and growth begins, which we call day 0 and occurs four days after seeding (Supplementary Material 6A.3.1), and to decide when to terminate the experiment, which we choose to be day 20. These measurements, supplemented with additional outer radius measurements from spheroids harvested for confocal imaging (Supplementary Material 6A.3.2-6A.3.3), provide an abundance of data. We now compare three experimental designs with increasing temporal resolution: (i) Resolution A, using measurements from days 1, 3, 8, 12, 17 (Figure 6.1l, 6.2a); (ii) Resolution B, using measurements from days 1, 3, 6, 8, 10, 12, 14, 17, 19 (Figure 6.1m, 6.2b); and, (iii) Resolution C, using daily measurements from day 0 to day 19 (Figure 6.1n, 6.2c). Excluding the final day(s) of measurements from these temporal resolutions allows a predictive check to be performed. Note that all these temporal resolutions are low relative to the capability of the automated imaging system but are high relative to the number of measurements typically taken in standard experiments [83, 87, 121, 138, 177].

To understand the influence of the choice of temporal resolution we now qualitatively and quantitatively compare the results. Across the three temporal resolutions in Figures 6.2d-f we observe excellent agreement between the full set of outer radius measurements, collected



**Figure 6.2:** Increasing the temporal resolution when the outer radius is measured is not sufficient to predict necrotic and inhibited radii. (a)-(c) Experimental data used in Design 1 with temporal resolutions A, B, and C. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set, where error bars show standard deviation of the experimental data. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, blue, orange lines in (g)-(j) represent profile likelihoods from Design 1 with temporal resolutions A, B, and C, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. Results shown for WM793b spheroids formed with 5000 cells per spheroid.

every six hours, and the predicted outer radius from the Greenspan model simulated with the MLE (Methods 6.5.1-6.5.2). However, it is clear that the prediction of the inhibited and necrotic radius is poor with Resolutions A and B (Figures 6.2d-e). With Resolution C, the prediction of the inhibited and necrotic radius appears to have improved (Figure 6.2f) but we will show that it is misleading to suggest that increasing the temporal resolution is always beneficial. While MLE point estimates are insightful, it is unclear whether a similarly excellent match to the outer radius measurements could be obtained with different parameter values in the mathematical model. To answer this question we undertake a profile likelihood analysis of the five parameters that govern the behaviour of the mathematical model (Methods 6.5.1):

1.  $s$  [day $^{-1}$ ], the rate at which cell volume is produced by mitosis per unit volume of living cells (Figure 6.2g),
2.  $R_c$  [ $\mu\text{m}$ ], the outer radius when the necrotic region first forms (Figure 6.2h),
3.  $\gamma = \lambda/s$  [-], the proportionality constant given by the rate at which cell volume is lost from the necrotic core,  $\lambda$ , divided by the rate at which cell volume is produced by mitosis per unit volume of living cells,  $s$ , (Figure 6.2i),
4.  $Q^2 = (c_\infty - c_i) / (c_\infty - c_n)$  [-], the ratio of the difference between the inhibited nutrient concentration threshold,  $c_i$ , and external nutrient threshold,  $c_\infty$  to the difference between the necrotic nutrient concentration threshold,  $c_n$ , and external nutrient threshold,  $c_\infty$  (Figure 6.2j),
5.  $R_o(0)$  [ $\mu\text{m}$ ], the initial outer radius (Supplementary Material 6A.4.3).

Profile likelihoods are a powerful tool to visualise and analyse how many parameter values give a similar match to the experimental data in comparison to the MLE. Furthermore, we use profile likelihoods to compute approximate 95% confidence intervals for each parameter (Supplementary Material 6A.4.1). Narrow approximate 95% confidence intervals indicate parameters are identifiable and that few parameters give a similar match to the data as the MLE. In contrast, wide approximate 95% confidence intervals suggest that parameters are not identifiable, that many parameters give a similar match to the experimental data, and that additional information is required to confidently estimate the parameters.

The profile likelihoods for  $s$  across all three temporal resolutions (Figure 6.2g) lead to a peak that is close to  $s = 0.14$  [day $^{-1}$ ]. These peaks correspond to the MLEs. While there is a wide 95% approximate confidence interval for  $s$  with Resolution A, there are narrow approximate 95% confidence intervals for  $s$  with Resolutions B and C. The profile likelihoods for the other

parameters,  $R_c$ ,  $\gamma$ , and  $Q$ , are wide and do not change significantly using different temporal resolutions (Figures 6.2h-j). For example, the profile likelihoods for  $\gamma$  across all three temporal resolutions (Figure 6.2i) are approximately flat and equal to one. These profile likelihoods for  $R_c$ ,  $\gamma$ , and  $Q$  suggest that increasing the temporal resolution does not provide significant additional information. These results are consistent with additional results using synthetic data (Supplementary Material 6A.5). Additional results for different initial spheroid sizes (Supplementary Material 6A.4) and results for the WM983b cell line also clearly show that increasing the temporal resolution may result in a worse prediction from the MLE for the time evolution of the internal structure (Supplementary Material 6A.6). Overall, these results suggest that Design 1 (Figure 6.1h) is not a reliable design to identify the true parameter values and cannot be used to determine details of the internal structure of tumour spheroids. This is important because this is the most standard measurement [18, 119, 138, 168, 173, 177].

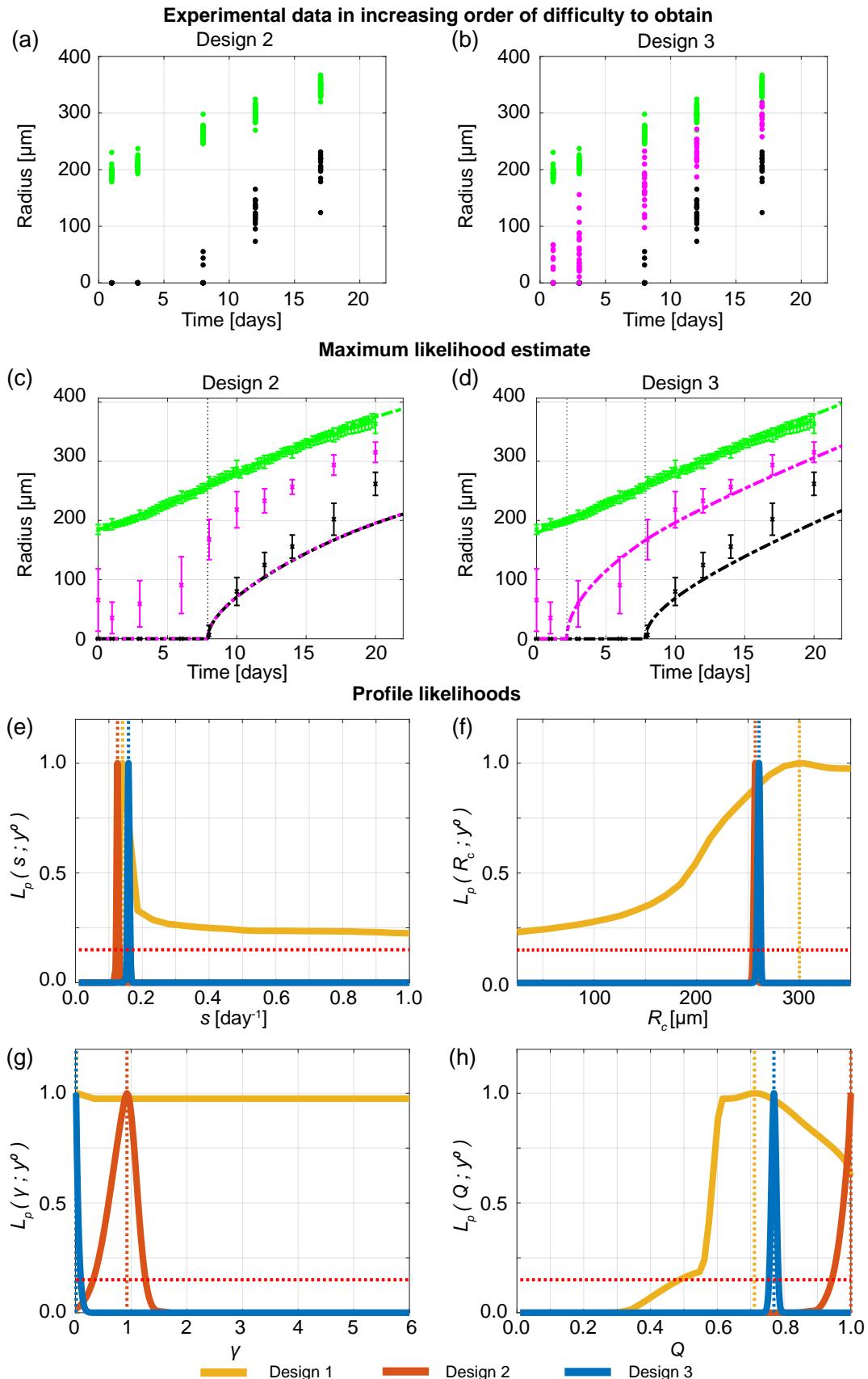
### 6.3.2 Cell cycle data are informative

Given that measuring the outer radius of tumour spheroids alone (Design 1) is insufficient to determine details of the internal spheroid structure, we now examine which measurements are required to provide reliable estimates. The next simplest measurements to obtain are both the outer radius and necrotic core radius, which we refer to as Design 2 (Figure 6.1i). However, Design 2 requires far more experimental effort since necrotic core measurements are more time-consuming involving harvesting, fixing, staining procedures, confocal microscopy or cryosectioning, and image processing. In addition, necrotic core measurements are end point measurements only, meaning that many spheroids are required to collect many data points. While intuitively we may anticipate that more effort leads to more insight, it is impossible to quantify the value of this additional effort without a mathematical modelling and uncertainty quantification framework such that we employ here.

Using Design 2, with low temporal resolution, for spheroids formed with 5000 cells per spheroid, we do not observe a necrotic core until approximately day 8 (Figure 6.3a, Supplementary Material 6A.3.3). The Greenspan model simulated with the MLE excellently matches the growth of the outer radius, as before, and now captures the formation and growth of the necrotic core (Figure 6.3c). Interestingly, the MLE suggests that the inhibited region is very small, so  $R_i(t)$  is very close to  $R_n(t)$ . However, experimental measurements of the inhibited radius not only suggest that an inhibited region exists, but that it forms prior to the formation of the necrotic core (Supplementary Material 6A.3.3). Profile likelihoods for each parameter are relatively narrow, and because the profile for  $Q$  is peaked and close to  $Q = 1$ , these profiles are consistent with either the absence of an inhibited region or a very small inhibited region (Figures 6.3e-h). Therefore, these data do not identify the true parameter values since the calibrated mathematical model is inconsistent with the experimental observations that clearly show the formation of an inhibited region. This inconsistency does not mean that the mathematical model is incorrect. Our interpretation of this inconsistency is that this experimental data are insufficient to identify the parameters in the mathematical model.

We now explore Design 3, where we measure the outer, necrotic, and inhibited radius of multiple tumour spheroids (Figure 6.1j, 6.3b). This design is considered third because measuring the inhibited radius is more difficult and requires substantial additional experimental effort. Using FUCCI transduced cell lines in combination with optical clearing procedures and confocal microscopy powerfully reveals intratumoral spatiotemporal differences with respect to the cycle. This method also requires semi-automated image processing and expert guidance to minimise subjectivity and accurately identify the inhibited region boundary (Supplementary

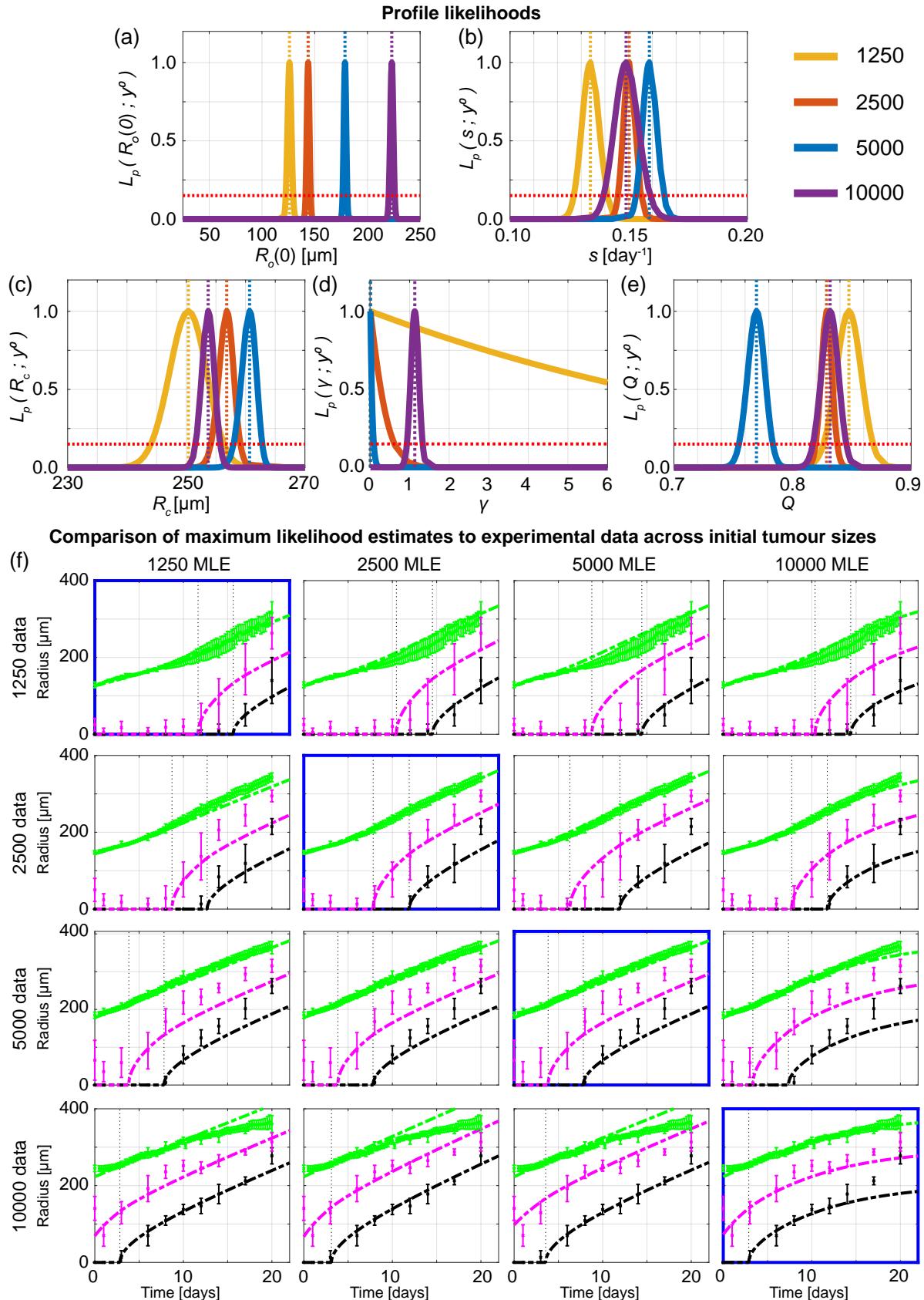
Material 6A.3.2) [218]. Simulating the Greenspan model with the MLE from Design 3 matches the evolution of the outer radius and captures the evolution of the necrotic and inhibited regions very accurately (Figure 6.3d). Furthermore, the profile likelihoods for all parameters are well formed, with a single narrow peak, suggesting that Design 3 identifies the true parameter values (Figure 6.3e-h). Comparing experimental Designs 1, 2, and 3, we observe that the profile likelihoods for  $s$  are consistent across all designs (Figure 6.3e) and the profile likelihoods for  $R_c$  (Figure 6.3f) are consistent for Designs 2 and 3. However, the profile likelihoods for  $\gamma$  (Figure 6.3g) and  $Q$  (Figure 6.3h) emphasise the power of measuring the inhibited radius and using Design 3 in comparison to Designs 1 and 2. These observations are consistent with additional results obtained using synthetic data (Supplementary Material 6A.5), different cell lines (Supplementary Material 6A.6), and initial spheroid sizes (Supplementary Material 6A.4.5). In Supplementary Material 6A.4.2, we also consider Design 3 with different temporal resolutions and experimental durations. Experiments performed for 4 or 10 days after spheroids form do not accurately predict late time behaviour. Designs that use the first days 10 to 20 or days 16 to 19 of measurements do not always accurately predict early time behaviour. Most insight is gained with Resolutions A, B, and C that cover the full experimental duration.



**Figure 6.3:** Measuring the necrotic and inhibited radius provides valuable information. (a)-(b) Experimental data used in Designs 2 and 3. (c)-(d) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set for Designs 2 and 3, where error bars show standard deviation. Profile likelihoods for (e)  $s$ , (f)  $R_c$ , (g)  $\gamma$ , (h)  $Q$ . Yellow, orange, blue lines in (e)-(h) represent profile likelihoods from Designs 1, 2, and 3, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. Results shown for WM793b spheroids formed with 5000 cells per spheroid.

### 6.3.3 Information gained using spheroids of different sizes is consistent

In the literature tumour spheroids are initialised with a wide range of cell numbers, leading to inconsistent results that are difficult to meaningfully compare across different protocols [8, 18, 54, 83, 88, 121, 168, 177]. Furthermore, it is unclear what the impact of this variability is when tumour spheroids are used to study fine-grained molecular-level interventions or potential drug designs. To quantitatively compare how information gained across experimental designs differs with respect to the initial number of cells in a spheroid we consider four initial spheroid sizes: 1250, 2500, 5000, 10000 cells per spheroid (Figure 6.1o). To proceed we use Design 3, and measure outer, necrotic, and inhibited radius, with time resolution A. Profile likelihoods for  $R_o(0)$  show four distinct narrow peaks corresponding to each initial spheroid size as expected (Figure 6.4a). Profile likelihoods for  $s$ ,  $R_c$ , and  $Q$  are consistent across the four initial spheroid sizes, allowing us to compare profile likelihoods on narrower intervals in Figures 6.4b,c,e. The profile likelihoods for  $\gamma$  (Figures 6.4d) are more variable due to the differing number of measurements collected in phase (iii). These results suggest that the initial spheroid size does not play a significant role in determining information from experiments, provided sufficient measurements are obtained in phase (iii). To support these results, we show along the diagonal of Figure 6.4f the solution of the mathematical model evaluated at the MLE associated with each initial spheroid size compared to the experimental measurements. Next, on the off-diagonals of Figure 6.4f, we compare how the Greenspan model simulated with the MLE from one initial spheroid size predicts data from different initial spheroid sizes by only changing the initial radius. For example, in the top right of Figure 6.4f we show that the Greenspan model simulated with the MLE obtained formed with 10000 cells per spheroid agrees well with data from spheroids formed with 1250 cells per spheroid, when the initial radius is set to be the initial radius of the 1250 MLE. Results in Figure 6.4f also show the inhibited and necrotic regions form earlier when considering spheroids formed with more cells, and results for spheroids formed with 10000 cells per spheroid suggest that these spheroids form in phase (ii) rather than phase (i). These observations are consistent with additional results from synthetic data (Supplementary Material 6A.5) and the WM983b cell line (Supplementary Material 6A.6).



**Figure 6.4:** Information gained from experiments across different initial tumour spheroid sizes is mostly consistent. Profile likelihoods for (a)  $R_o(0)$ , (b)  $s$ , (c)  $R_c$ , (d)  $\gamma$ , (e)  $Q$ . Yellow, orange, blue, and purple lines in (a)-(e) represent profile likelihoods from WM793b spheroids formed with 1250, 2500, 5000, 10000 cells per spheroid, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. (f) Comparison of Greenspan model simulated with maximum likelihood estimates compared to full experimental data sets across initial tumour spheroid size, where error bars show standard deviation.

## 6.4 Discussion

In this work we present an objective theoretical framework to quantitatively compare tumour spheroid experiments across a range of experimental designs using the seminal Greenspan mathematical model and statistical profile likelihood analysis. By considering different temporal data resolutions, experiment durations, types of measurements, and initial spheroid sizes we identify the experimental design choices that lead to reliable biological insight. Namely, Design 3 where we obtain outer, necrotic, and inhibited radius measurements requires most effort but is essential to determine the dynamics of tumour spheroid structure and growth. Importantly, we also show that temporal resolution and initial spheroid sizes are less important choices. Therefore, we recommend that for future studies, where tumour spheroid structure is important, that cell cycle data are essential and that some measurements using Design 3 is more valuable than many measurements using Designs 1 or 2. Furthermore, as information from tumour spheroids across varying initial spheroid sizes is relatively consistent, provided sufficient measurements in phase (iii) are obtained, we recommend that performing experiments with larger tumour spheroids can be beneficial to obtain useful information in a shorter experimental duration (Supplementary Material 6A.5.3). However, we also note that this may lead to large tumour spheroids that begin growth in phase (ii) rather than phase (i).

To perform this analysis we use Greenspan's seminal mathematical model, where all parameters have a relatively straightforward biological interpretation. We find that Greenspan's model performs remarkably well across cell lines and initial spheroid sizes, and provides powerful insights into experimental design. Even though Greenspan's model is relatively simple, and may not capture all of the biological details of tumour spheroid growth, the fact that results for experimental data are consistent with those from synthetic data enhances our confidence that key biological features are captured in Greenspan's model (Supplementary Material 6A.5). Future modelling may wish to explore potential model misspecifications, for example WM983b spheroids appear to reduce in size at very late time suggesting a fourth phase in these *in vitro* experiments (Supplementary Material 6A.6); and, WM164 spheroids, possibly due to their lack of spherical symmetry [219], are more challenging to interpret as information gained using spheroids of different sizes is not consistent (Supplementary Material 6A.7).

The general framework presented in this work can be applied to other cell types, for example FUCCI-transduced lung, stomach and breast cancer cells [223, 250, 251, 255], to extract more information from existing experimental data across experimental designs, and is suitable to be extended to consider tumour spheroids grown in different conditions and to more complex mathematical models. Given that cell cycle data is demonstrated to be informative in this study,

we suggest that it may be beneficial for FUCCI technology to be further developed and more widely used, for example in avascular patient-derived organoids [233], and our framework be extended to these heterogeneous populations accordingly. Furthermore, the insights of this study provide a platform for future studies that develop, test, and examine cancer treatments with spheroids. In such future studies cell cycle data will be informative since cytotoxic or cytostatic drugs may result in similar changes in the outer radius but due to different causes, that can be measured by cell death and cell cycle imaging (Haass laboratory unpublished observations).

## 6.5 Methods

### 6.5.1 Mathematical model

Greenspan's mathematical model describes the three phases of avascular tumour spheroid growth [79]. Spherical symmetry is assumed at all times and maintained by adhesion and surface tension. Under these minimal assumptions, the only independent variables are time,  $t$  [days], and radial position,  $r$  [ $\mu\text{m}$ ]. Tumour growth is governed by the evolution of the outer radius,  $R_o(t)$  [ $\mu\text{m}$ ], the inhibited radius,  $R_i(t)$  [ $\mu\text{m}$ ], and the necrotic radius,  $R_n(t)$  [ $\mu\text{m}$ ]. Nutrient diffuses within the spheroid with diffusivity  $k$  [ $\mu\text{m}^2 \text{ day}^{-1}$ ] and is consumed by living cells at a constant rate per unit volume  $\alpha$  [ $\text{mol } \mu\text{m}^{-3} \text{ day}^{-1}$ ]. The external nutrient concentration is  $c_\infty$  [ $\text{mol } \mu\text{m}^{-3}$ ]. The nutrient concentration at a distance  $r$  from the centre of the spheroid and time  $t$ , denoted  $c(r, t)$  [ $\text{mol } \mu\text{m}^{-3}$ ], is assumed to be at diffusive equilibrium. Therefore, at any instant in time we have  $c(r, t) = c(r)$  due to fast diffusion of nutrient. However, as  $R_o(t)$  is growing, nutrient diffusion occurs on a growing domain and we write  $c(r) = c(r(t))$ . The inhibited and necrotic regions form when the nutrient concentration at the centre of the spheroid reaches  $c_i$  [ $\text{mol } \mu\text{m}^{-3}$ ] and  $c_n$  [ $\text{mol } \mu\text{m}^{-3}$ ], respectively. For  $c(r(t)) > c_n$  the rate at which cell volume is produced by mitosis per unit volume of living cells is  $s$  [ $\text{day}^{-1}$ ]. In the necrotic core cellular debris disintegrates into simpler chemical compounds that are freely permeable through cell membranes. The mass lost in the necrotic region is replaced by cells pushed inwards by forces of adhesion and surface tension. The necrotic core loses cell volume at a rate proportional to the necrotic core volume with proportionality constant  $3\lambda$  [ $\text{day}^{-1}$ ], where the three is included for mathematical convenience.

Conservation of mass is written in words as  $A = B + C - D - E$  where  $A$  is the total volume of living cells at any time,  $t$ ;  $B$  is the initial volume of living cells at time  $t = 0$ ;  $C$  is the total volume of cells produced in  $t \geq 0$ ;  $D$  is the total volume of necrotic debris at time  $t$ ;  $E$  is the total volume lost in the necrotic core in  $t \geq 0$ . Writing  $A, B, C, D, E$  in their mathematical form gives the conservation of mass equation and also writing the nutrient diffusion equation gives,

$$R_o^2(t) \frac{dR_o(t)}{dt} = \frac{s}{3} [R_o^3(t) - R_i^3(t)] - \lambda R_n^3(t), \quad (6.5.1.1)$$

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} c(r(t)) \right) = \frac{\alpha}{k} H(r - R_n(t)) H(R_o(t) - r), \quad 0 \leq r \leq R_o(t) \quad (6.5.1.2)$$

where  $R_i(t), R_n(t)$  are the radii implicitly defined by  $c(R_i(t), t) = c_i$ , and  $c(R_n(t), t) = c_n$ , respectively, if the nutrient concentration inside the spheroid is sufficiently small otherwise  $R_i(t) = 0$  or  $R_n(t) = 0$ , and  $H(\cdot)$  is the Heaviside step function. There are eight unknowns:

$\Theta = (s, \lambda, \alpha, k, c_\infty, c_i, c_n, R_o(0))$ . Note this includes  $R_o(0)$  which we treat as a parameter since we also need to estimate this quantity. Rescaling reduces the number of parameters to five:  $\theta = (R_o(0), R_c, s, \gamma, Q)$ . The new dimensionless parameters are: the outer radius when the necrotic region first forms defined as  $R_c^2 = (6k/\alpha)(c_\infty - c_n)$ ;  $Q^2 = (c_\infty - c_i)/(c_\infty - c_n)$ ; and  $\gamma = s/\lambda$ . Further details, and a formal demonstration that this model is equivalent to a model where nutrient determines the necrotic region and waste produced from live cells determines the inhibited region, are provided in Supplementary Material 6A.1.

### 6.5.2 Practical parameter identifiability analysis

To determine the maximum likelihood estimate (MLE) and approximate 95% confidence intervals for the parameters  $\theta = (R_o(0), R_c, s, \gamma, Q)$  we use profile likelihood identifiability analysis [180, 188, 204, 205, 246]. We first choose simple parameter bounds and then compare the width of these simple parameter bounds to realised interval estimates for the parameters. Initial parameter bounds are chosen to be the same across all experimental designs analysed in this study. Outer radius data suggests we choose  $0 < R_o(0) < 350 \text{ [\mu m]}$  and  $0 < R_c < 250 \text{ [\mu m]}$ . Assuming a cell doubling time of at least 12 hours and performing preliminary data exploration, we set  $0 < s < 1 \text{ [day}^{-1}\text{]}$  (Supplementary Material 6A.2.1). Limited information exists for the parameter  $\gamma$  so bounds are determined by preliminary data exploration to be  $0 < \gamma < 6$ . By definition of  $Q$  and experimental results that show the inhibited region forms before the necrotic core, we set  $0 < Q \leq 1$ . Note that the time-evolution of  $R_o(t)$  and  $R_n(t)$  are the same for  $Q = 1$  and  $Q > 1$ . The difference arises for  $R_i(t)$ , where it is equal to  $R_n(t)$  for  $Q = 1$  and equal to zero for  $Q > 1$ .

To determine the interval estimates for the parameters we treat the mathematical model as having two components. The first is the deterministic mathematical model governing the evolution of  $R_o(t)$ ,  $R_n(t)$ , and  $R_i(t)$  and the second is a probabilistic observation model accounting for experimental variability and measurement error. Specifically, we assume that experimental measurements are noisy observations of the deterministic mathematical model [97, 188]. For each of the three measurement types  $R_o(t)$ ,  $R_n(t)$ , and  $R_i(t)$  we assume that the observation error is independent and identically distributed and that the noise is additive and normally distributed with zero mean and variance  $\sigma_o^2$ ,  $\sigma_n^2$ , and  $\sigma_i^2$ , respectively [97, 244] (Supplementary Material 6A.2.3). We approximate  $\sigma_o^2 \approx s_o^2$ ,  $\sigma_n^2 \approx s_n^2$ , and  $\sigma_i^2 \approx s_i^2$  where  $s_o^2$ ,  $s_n^2$ , and  $s_i^2$  are pooled sample variances of the outer, necrotic, and inhibited radius measurements, respectively [31].

The likelihood function  $p(y^o|\theta)$  is the likelihood of the observations  $y^o$  given the parameter  $\theta$ . This corresponds to the probabilistic observation model evaluated at the observed data. The maximum likelihood estimate is  $\hat{\theta} = \operatorname{argmax}_{\theta} p(y^o|\theta)$ . We present results in terms of the normalised likelihood function  $L(\theta; y^o) = p(y^o|\theta) / \max_{\theta} p(y^o|\theta)$  which we consider a function of  $\theta$  for fixed  $y^o$ . Profile likelihoods for each parameter are obtained by assuming the full parameter  $\theta$  can be partitioned into a scalar interest parameter,  $\psi$ , and vector nuisance parameter,  $\phi$ , so that  $\theta = (\psi, \phi)$ . The profile likelihood for  $\psi$  is then  $L_p(\psi; y^o) = \max_{\phi} L(\psi, \phi; y^o)$ . Approximate 95% confidence intervals are then calculated using a profile likelihood threshold value of 0.15 (Supplementary Material 6A.4.1) [180]. Prediction intervals are not shown since confidence intervals are narrow in many cases. Further details, including exact forms of the likelihood

function and the use of log-likelihoods for calculations, and numerical methods are provided in Supplementary Material 6A.2.

### 6.5.3 Experimental methods

*Cell culture.* The human melanoma cell lines WM793b, WM983b, and WM164 were genotypically characterised [87, 100, 213, 214], grown as described in [217], and authenticated by short tandem repeat fingerprinting (QIMR Berghofer Medical Research Institute, Herston, Australia). All cell lines were transduced with fluorescent ubiquitination-based cell cycle indicator (FUCCI) constructs as described in [87, 217].

*Spheroid generation, culture, and experiments.* Spheroids were generated in 96-well cell culture flat-bottomed plates (3599, Corning), with four different seeding densities (1250, 2500, 5000, 10000 total cells/well), using 50 µl total/well non-adherent 1.5% agarose to promote formation of a single spheroid per well [218]. For all spheroid experiments, after a formation phase of 4, 3 and 2 days for WM793b, WM164 and WM983b, respectively (Supplementary Material 6A.3.1), and then every 3-4 days for the duration of the experiment, 50% of the medium in each well was replaced with fresh medium (200 µl total/well). Incubation and culture conditions were as described in *Cell culture*.

To estimate the outer radius, one plate for each cell line, containing 24 spheroids for each initial spheroid size, was placed inside the IncuCyte S3 live cell imaging system (Sartorius, Goettingen, Germany) incubator (37 °C, 5% CO<sub>2</sub>) immediately after seeding the plates. IncuCyte S3 settings were chosen to image every 6 hours for the duration of the experiment with the 4× objective. To estimate the radius of the inhibited and necrotic region and additional outer radius measurements, spheroids maintained in the incubator were harvested, fixed with 4% paraformaldehyde (PFA), and stored in phosphate buffered saline solution, sodium azide (0.02%), Tween-20 (0.1%), and DAPI (1:2500) at 4 °C, on days 3, 4, 5, 7, 10, 12, 14, 16, 18, 21 and 24 after seeding. For necrotic core measurements, 12 hours prior to harvesting 1 µmol total/well DRAQ7™ dye (Abcam, Cambridge, United Kingdom. ab109202) was added to each well [114, 218]. Fixed spheroids were set in place using low melting 2% agarose and optically cleared in 500 µl total/well high refractive index mounting solution (Quadrol 9 % wt/wt, Urea 22 % wt/wt, Sucrose 44 % wt/wt, Triton X-100 0.1 % wt/wt, water) for 2 days in a 24-well glass bottom plate (Cellvis, P24-1.5H-N) before imaging to ensure accurate measurements [47, 222]. Images were then captured using an Olympus FV3000 confocal microscope with the 10× objective focused on the equatorial plane of each spheroid.

*Image processing.* Images captured with the IncuCyte S3 were processed using the accompanying IncuCyte 2020C Rev1 software (spheroid analysis type, red image channel, largest red object area per well). Area masks were visually compared with IncuCyte brightfield images to confirm accuracy. Area was converted to an equivalent radius ( $r^2 = A/\pi$ ). Confocal mi-

croscopy images were converted to TIFF files in ImageJ and then processed with custom MATLAB scripts that use standard MATLAB image processing toolbox functions. These scripts are freely available on Zenodo with DOI:10.5281/zenodo.5121093 [33].



## **Chapter 6A**

# **Supplementary Material for Chapter 6**

<b>Supplementary Material</b>	<b>Page No.</b>
A. Mathematical model	245
A.1 Nutrient only	245
A.2 Nutrient and waste	252
B. Profile likelihood further details	255
B.1 Numerical method	255
B.2 Parameter bounds	259
B.3 Pooled sample variances	260
C. Experimental data	261
C.1 Outer radius experimental measurements and images	261
C.1.1 Spheroid formation duration	265
C.2 Confocal microscopy	269
C.1.1 Measurements	269
C.3 Confocal microscopy supplementary experimental images	270
C.3.1 WM793b	270
C.3.2 WM983b	274
C.3.3 WM164	277
C.3.4 3D renderings	278
D. Additional results for WM793b	279
D.1 Results in tables	279
D.2 Measurement times and experimental duration	281
D.3 Profile likelihoods for $R_o(0)$	286
D.4 Outer radius measurements are not sufficient to predict inhibited and necrotic radii	287
D.5 Cell cycle data are informative	291
E. Synthetic data based on WM793b	295
E.1 Outer radius measurements are not sufficient to predict inhibited and necrotic radii	295
E.2 Cell cycle data are informative	298
E.3 Role of initial spheroid sizes and experimental duration	300
E.4 Increasing number of measurements	302
F. Parameter identifiability for WM983b	304
F.1 Outer radius measurements are not sufficient to predict inhibited and necrotic radii	305
F.2 Cell cycle data are informative	309
F.3 Information gained across spheroid sizes is consistent	312
G. Parameter identifiability for WM164	313

## 6A.1 Mathematical model

### 6A.1.1 Nutrient only

Here we recall Greenspan's mathematical model governing the evolution of  $R_o(t)$ ,  $R_i(t)$ , and  $R_n(t)$  (Figure 6.1f-g). We consider conservation of mass to govern the evolution of  $R_o(t)$ . Assuming: i) all living cells are identical and an incompressible mass of constant volume; ii) cell division occurs instantaneously relative to the growth time of the tumour, and each daughter cell occupies the same volume as any other cell; iii) the proliferation rate is a constant,  $s$ , for cells which have sufficient nutrient; and, iv) the mass density of living cells is constant and equal to density of necrotic debris; then conservation of mass is equivalent to conservation of volume, giving,

$$A = B + C - D - E, \quad (6A.1.1)$$

where  $A$  is the total volume of living cells at any time,  $t$ ;  $B$  is the initial volume of living cells at time  $t = 0$ ;  $C$  is the total volume of cells produced in  $t \geq 0$ ;  $D$  is the total volume of necrotic debris at time  $t$ ; and,  $E$  is the total volume lost in the necrotic core in  $t \geq 0$ .

Writing  $A, B, C, D, E$  in their mathematical forms and recalling that the volume and surface area of a sphere with radius  $r$  is  $4\pi r^3/3$  and  $4\pi r^2$ , respectively, gives

$$A = \frac{4\pi}{3} (R_o^3(t) - R_n^3(t)), \quad (6A.1.2.1)$$

$$B = \frac{4\pi}{3} R_o^3(0), \quad (6A.1.2.2)$$

$$C = 4\pi \int_0^t \int_{R_i(t)}^{R_o(t)} sr^2 dr dt, \quad (6A.1.2.3)$$

$$D = \frac{4\pi}{3} R_n^3(t), \quad (6A.1.2.4)$$

$$E = \frac{4\pi}{3} \int_0^t 3\lambda R_n^3(t) dt, \quad (6A.1.2.5)$$

where the three inside the integral of equation (6A.1.2.5) is included for mathematical convenience. Substituting equations (6A.1.2.1)-(6A.1.2.5) into equation (6A.1.1) and simplifying gives,

$$R_o^3(t) = R_o^3(0) + 3 \int_0^t \int_{R_i(t)}^{R_o(t)} sr^2 dr dt - \int_0^t 3\lambda R_n^3(t) dt. \quad (6A.1.3)$$

Differentiating equation (6A.1.3) with respect to time and simplifying gives the more useful

form,

$$R_o^2(t) \frac{dR_o(t)}{dt} = \underbrace{\frac{s}{3} (R_o^3(t) - R_i^3(t))}_{\text{proliferation of living cells}} - \underbrace{\lambda R_n^3(t)}_{\text{mass lost in necrotic core}}. \quad (6A.1.4)$$

The other important equation concerns the evolution of nutrient within the spheroid. Rewriting equation (6.5.1.2) gives

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} c(r(t)) \right) = \frac{\alpha}{k} H(r - R_n(t)) H(R_o(t) - r), \quad 0 \leq r \leq R_o(t). \quad (6A.1.5)$$

where  $H(\cdot)$  is the Heaviside step function.

To determine the full evolution of the system we solve equations (6A.1.4) and (6A.1.5) together with the nutrient thresholds  $c_i$  and  $c_n$  which implicitly define  $R_i(t)$  and  $R_n(t)$ , respectively, through

$$c(R_i(t), t) = c_i, \quad (6A.1.6.1)$$

$$c(R_n(t), t) = c_n, \quad (6A.1.6.2)$$

if the nutrient concentration inside the spheroid is sufficiently small otherwise  $R_i(t) = 0$  or  $R_n(t) = 0$ . Note that the equation (6A.1.4) for nutrient does not involve any temporal derivative so the only initial condition required to solve the full system of equations (6A.1.4) and (6A.1.5) is the initial outer radius,  $R_o(0)$ .

The solution of equation (6A.1.5) is,

$$c(r(t)) = \begin{cases} c_\infty - \frac{\alpha}{6k} (R_o^2(t) - r^2) + \frac{AR_n^3(t)}{3k} \left( \frac{1}{r} - \frac{1}{R_o(t)} \right), & R_n(t) \leq r \leq R_o(t), \\ c_n, & 0 \leq r \leq R_n(t), \end{cases} \quad (6A.1.7)$$

where

$$c_\infty - c_n = \frac{\alpha}{3k} \left[ \frac{1}{2} (R_o^2(t) - R_n^2(t)) - \frac{R_n^2(t)}{R_o(t)} (R_o(t) - R_n(t)) \right]. \quad (6A.1.8)$$

The necrotic region first forms when the nutrient concentration reaches  $c_n$  at the centre, which occurs when  $R_n(t) = 0$  and  $r = 0$  in equation (6A.1.8), which gives a critical outer radius,

$$R_c^2 = \frac{6k}{\alpha} (c_\infty - c_n). \quad (6A.1.9)$$

Also recall that  $R_i(t)$  corresponds to  $c(R_i(t), t) = c_i$  which we can substitute into equation (6A.1.7) to give,

$$c_\infty - c_i = \frac{\alpha}{3k} \left[ \frac{1}{2} (R_o^2(t) - R_i^2(t)) - R_n^3(t) \left( \frac{1}{R_i(t)} - \frac{1}{R_o(t)} \right) \right]. \quad (6A.1.10)$$

Since the inhibited region first forms when the nutrient concentration reaches  $c_i$  at the centre of the spheroid and the necrotic region forms after the inhibited region, setting  $R_n(t) = 0$  and  $r = 0$  on right-hand-side of equation (6A.1.10) gives the outer radius when the inhibited region first forms

$$R_d^2 = \frac{6k}{\alpha} (c_\infty - c_i). \quad (6A.1.11)$$

We can then define a useful dimensionless quantity,  $Q^2 = R_d^2/R_c^2 = (c_\infty - c_i) / (c_\infty - c_n)$ , which is related to the time when phase (ii) begins.

Equations (6A.1.4), (6A.1.5), and (6A.1.6) can now be solved in each of phase (i), (ii), and (iii). To provide valuable insights into the structure of the solutions to the Greenspan model it helps to consider the non-dimensional form of the equations and their solutions. To non-dimensionalise we rescale time with  $s$  to give  $\tau = st$  and rescale lengths with  $R_c$  via  $\xi_o(t) = R_o(t)/R_c$ ,  $\xi_i(t) = R_i(t)/R_c$ , and  $\xi_n(t) = R_n(t)/R_c$ . Then phase (ii) starts when  $\xi_o(t) = Q$  and phase (iii) starts when  $\xi_o(t) = 1$ . We now consider each phase in turn.

### Phase (i)

In phase (i), all cells are free to proliferate and the nutrient concentration is sufficiently high, i.e.  $c(r, t) > c_i$  for  $0 \leq r \leq R_o(t)$ , such that there is no inhibited or necrotic region (Figure 6.1(a)). Phase (i) ends when the nutrient concentration at the centre of the spheroid equals the inhibited threshold, when  $c(0, t) = c_i$  and  $R_o(t) = R_d$ .

Since  $R_i(t) = 0$  and  $R_n(t) = 0$ , equation (6A.1.4) becomes

$$R_o^2(t) \frac{dR_o(t)}{dt} = \frac{s}{3} R_o^3(t), \quad (6A.1.12)$$

giving,

$$R_o(t) = R_o(0) \exp \left( \frac{st}{3} \right). \quad (6A.1.13)$$

Non-dimensionalising gives,

$$\xi_o(\tau) = \xi_o(0) \exp\left(\frac{\tau}{3}\right), \quad \text{for } 0 \leq \tau \leq 3 \log\left(\frac{Q}{\xi_o(0)}\right) = \tau_1. \quad (6A.1.14)$$

Given the solution in equation (6A.1.14) we determine  $R_o(t)$  by reintroducing  $s$  and  $R_c$ ,

$$R_o(t) = \xi_o(st)R_c, \quad \text{for } 0 \leq t \leq \frac{\tau_1}{s}. \quad (6A.1.15)$$

Note that  $R_i(t) = 0$  and  $R_n(t) = 0$  throughout phase (i). Hence, we have obtained  $R_o(t)$ ,  $R_i(t)$ ,  $R_n(t)$  throughout this phase.

### Phase (ii)

In phase (ii) the spheroid experiences inhibited growth due to a core of inhibited cells and outer region of freely proliferating cells (Figure 6.1(b)). Phase (ii) ends when the necrotic core forms.

Since  $R_i(t) > 0$  and  $R_n(t) = 0$ , equation (6A.1.4) becomes

$$R_o^2(t) \frac{dR_o(t)}{dt} = \frac{s}{3} (R_o^3(t) - R_i^3(t)). \quad (6A.1.16)$$

Non-dimensionalising equation (6A.1.16) gives,

$$\xi_o^2(\tau) \frac{d\xi_o(\tau)}{d\tau} = \frac{1}{3} [\xi_o^3(\tau) - \xi_i^3(\tau)]. \quad (6A.1.17)$$

Equation (6A.1.17) is a function of two variables,  $\xi_o(\tau)$  and  $\xi_i(\tau)$ , which we can simplify to a function of one variable by introducing a change of variables  $y_i(\tau) = \xi_i(\tau)/\xi_o(\tau)$ , and by using the constraint  $Q^2/\xi_o^2(\tau) = 1 - y_i^2(\tau)$ , to give

$$\frac{3y_i(\tau)}{(1 - y_i(\tau)^2)(1 - y_i(\tau)^3)} \frac{dy_i(\tau)}{d\tau} = 1, \quad (6A.1.18)$$

with initial condition  $y_i(\tau) = 0$  at  $\tau = \tau_1$  and terminating condition  $y_i(\tau)^2 = 1 - Q^2$ . The constraint used to derive equation (6A.1.18) and the termination condition for phase (ii) are obtained with the following argument. In phase (ii) equation (6A.1.10) is

$$R_o^2(t) - R_i^2(t) = R_d^2. \quad (6A.1.19)$$

Non-dimensionalising equations (6A.1.8) and (6A.1.19), using definitions of  $\xi_o(\tau)$ ,  $\xi_i(\tau)$ ,  $Q$ , and combining the resulting expressions gives  $Q^2 = \xi_o^2(\tau) - \xi_i^2(\tau)$ . Rewriting in terms of  $y_i(\tau)$  gives

$Q^2/\xi_o^2(\tau) = 1 - y_i^2(\tau)$ , which gives the constraint used to derive equation (6A.1.18). Using the fact that phase (ii) ends when  $\xi_o(\tau) = 1$  and rearranging gives the termination condition for  $y_i(\tau)$ .

Numerically solving equation (6A.1.18), using MATLAB's in-built *ode15s* differential equation solver [145] with absolute and relative tolerances set to  $1 \times 10^{-8}$ , we obtain  $y_i(\tau)$  for phase (ii). To obtain  $R_o(t)$  we use the constraint  $Q^2/\xi_o^2(\tau) = 1 - y_i(\tau)^2$ , and definitions of  $\xi_o(\tau)$  and  $\xi_i(\tau)$  to obtain  $R_o(t) = R_c Q [1 - y_i^2(st)]^{-1/2}$ . Similarly using the constraints we obtain  $R_i(t) = R_c Q [1/(1 - y_i^2(st)) - 1]^{1/2}$ . Recall  $R_n(t) = 0$  throughout phase (ii). Hence, we have obtained  $R_o(t), R_i(t), R_n(t)$  throughout this phase.

### Phase (iii)

In phase (iii) the spheroid experiences inhibited growth due to an outer proliferating region, an intermediate region of inhibited cells, and a necrotic core (Figure 6.1(a)). At steady state there is a balance between the number of cells that are proliferating in the outer region and mass lost from the necrotic core.

Since  $R_i(t) > 0$  and  $R_n(t) > 0$ , all terms in equation (6A.1.4) are non-zero. Non-dimensionalising equation (6A.1.4) gives

$$\xi_o^2(\tau) \frac{d\xi_o(\tau)}{d\tau} = \frac{1}{3} [\xi_o^3(\tau) - \xi_i^3(\tau)] - \gamma \xi_n^3(\tau), \quad (6A.1.20)$$

where  $\gamma = \lambda/s$ . Equation (6A.1.20) is a function of three variables  $\xi_o(\tau), \xi_i(\tau), \xi_n(\tau)$ . Introducing  $y_i(\tau) = \xi_i(\tau)/\xi_o(\tau)$  and  $y_n(\tau) = \xi_n(\tau)/\xi_o(\tau)$  we rewrite equation (6A.1.20) and the non-dimensionalised forms of equations (6A.1.8) and (6A.1.10) as

$$\frac{9y_n(\tau)}{(1 + 2y_n(\tau))(1 - y_n(\tau))} \frac{dy_n(\tau)}{d\tau} = 1 - y_i^3(\tau) - 3\gamma y_n^3(\tau), \quad (6A.1.21.1)$$

$$\xi_o^{-2}(\tau) = (1 - y_n(\tau))^2 (1 + 2y_n(\tau)), \quad (6A.1.21.2)$$

$$\frac{Q^2}{\xi_o^2(\tau)} = 1 - y_i^2(\tau) - 2y_n^3(\tau) \left( \frac{1 - y_i(\tau)}{y_i(\tau)} \right), \quad (6A.1.21.3)$$

noting that equation (6A.1.21.1) is obtained using equation (6A.1.21.2). Then we numerically solve equations (6A.1.21.1)-(6A.1.21.3) to obtain  $y_i(\tau)$  and  $y_n(\tau)$  using the the following approach. First, we substitute equation (6A.1.21.2) into equation (6A.1.21.3) to eliminate  $\xi_o(\tau)$  and rearrange which gives

$$0 = -Q^2 [(1 - y_n(\tau))^2 (1 + 2y_n(\tau))] + 1 - y_i^2(\tau) - 2y_n^3(\tau) \left( \frac{1 - y_i(\tau)}{y_i(\tau)} \right). \quad (6A.1.22)$$

Equations (6A.1.21.1) and (6A.1.22) form a system of differential-algebraic equations which we numerically solve using MATLAB's in-built *ode15s* solver with relative and absolute tolerances set to  $1 \times 10^{-8}$ . Given the solution for  $y_n(\tau)$  and  $y_i(\tau)$  we obtain  $R_o(t)$ ,  $R_i(t)$  and  $R_n(t)$  using the following approach. Given  $y_n(\tau)$  we obtain  $\xi_o(\tau)$  using equation (6A.1.21.2). Then  $R_o(t) = R_c \xi_o(st)$ . Using the definition of  $y_i(\tau)$ ,  $y_n(\tau)$  and  $\xi_o(\tau)$  we obtain  $R_i(t) = R_c y_i(st) \xi_o(st)$  and  $R_n(t) = R_c y_n(st) \xi_o(st)$ . Hence, we have obtained  $R_o(t)$ ,  $R_i(t)$ ,  $R_n(t)$  throughout this phase. Key software for this section is freely available on a GitHub repository.

Greenspan's mathematical model assumes that tumour spheroids experience three phases of growth [79]. While we find experimental evidence confirming that many tumour spheroids experience three phases of growth (Figures 6.1, 6A.9, 6A.10, 6A.11, 6A.13, and 6A.14), we also find experimental evidence suggesting tumour spheroids seeded with a higher number of cells may form in phase (ii) (Figures 6A.12 and 6A.15). Here, we now describe how to initialise Greenspan's mathematical model in phase (ii) and in phase (iii). We consider phase (iii) since calculations used for statistical identifiability analysis may choose parameters such that the likelihood of starting in phase (iii) is evaluated.

To initialise Greenspan's model with a spheroid in phase (ii) we first prescribe  $R_o(0)$  and recall that in phase (ii) there is no necrotic core, so  $R_n(0) = 0$ . Then from equation (6A.1.10) with  $R_n(0) = 0$ , the corresponding inhibited radius is  $R_i(0) = (R_o(0)^2 - Q^2 R_c^2)^{1/2}$ . To initialise Greenspan's model with a spheroid in phase (iii) we first prescribe  $R_o(0)$ . Given  $R_o(0)$  we rewrite equation (6A.1.8) as the following cubic polynomial  $(2/R_o(0))R_n(0)^3 - 3R_n(0)^2 + R_o(0)^2 - R_c^2 = 0$ , where  $R_n(0)$  is the unknown variable. We determine the three solutions of this polynomial using the MATLAB function *roots* [4] and define  $R_n(0)$  as the only physically realistic real-valued solution which satisfies  $0 < R_n(0) < R_o(0)$ . Similarly, to obtain  $R_i(0)$ , we rewrite equation (6A.1.10) as the following cubic polynomial  $R_i(0)^3 + (Q^2 R_c^2 - R_o(0)^2 - 2R_n(0)^3/R_o(0))R_i(0) + 2R_n(0)^3 = 0$ , where  $R_i(0)$  is the unknown variable and  $R_o(0)$  and  $R_n(0)$  are known. We then define  $R_i(0)$  as the only physically realistic real-valued solution which satisfies  $R_n(0) < R_i(0) < R_o(0)$ . For statistical identifiability analysis we assume spheroids may form in phase (i), phase (ii), or phase (iii).

The approach taken in phases (i), (ii), and (iii) means that we do not require knowledge of the values of the parameters  $c_\infty, c_n, c_i, k$  and  $\alpha$  but instead only the value of  $Q = [(c_\infty - c_i)/(c_\infty - c_n)]^{1/2}$ . This reduces the number of parameters describing the evolution of the spheroid from eight to five. The three pieces of information no longer consider regard the nutrient concentration which we do not directly measure in this study and has been explored in other studies [80, 155]. Furthermore, equations (6A.1.8)-(6A.1.10) show that there are con-

straints on the relationships between  $R_o(t)$ ,  $R_i(t)$ ,  $R_n(t)$  which can be explored further.

### 6A.1.2 Nutrient and waste

The model presented in methods section 6.5.1 and supplementary material 6A.1.1 is a special case of Greenspan's model [79]. The general Greenspan model proposes the inhibited region is a result of a build up of waste produced from live or dead cells and the necrotic region forms due to a lack of nutrient. Here, we consider the alternative case where waste is produced from live cells only and show that, for the measurements we obtain, it is equivalent to the nutrient only case we consider in the main manuscript (Figure 6A.1). We do not consider waste produced only from dead cells in this study since with that model the necrotic core must form before the inhibited region which is not what we observe in these experiments (Figure 6.1(b)).

In comparison to the nutrient only model in supplementary material 6A.1.1, the model with nutrient and waste requires an additional equation for the evolution of waste concentration,  $\beta(r, t)$ . The full system of governing equations are,

$$R_o^2(t) \frac{dR_o(t)}{dt} = \frac{s}{3} [R_n^3(t) - \max(R_n^3(t), R_i^3(t))] - \lambda R_n^3(t), \quad (6A.1.23.1)$$

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} c(r(t)) \right) = \frac{\alpha}{k} H(r - R_n(t)) H(R_o(t) - r), \quad 0 \leq r \leq R_o(t), \quad (6A.1.23.2)$$

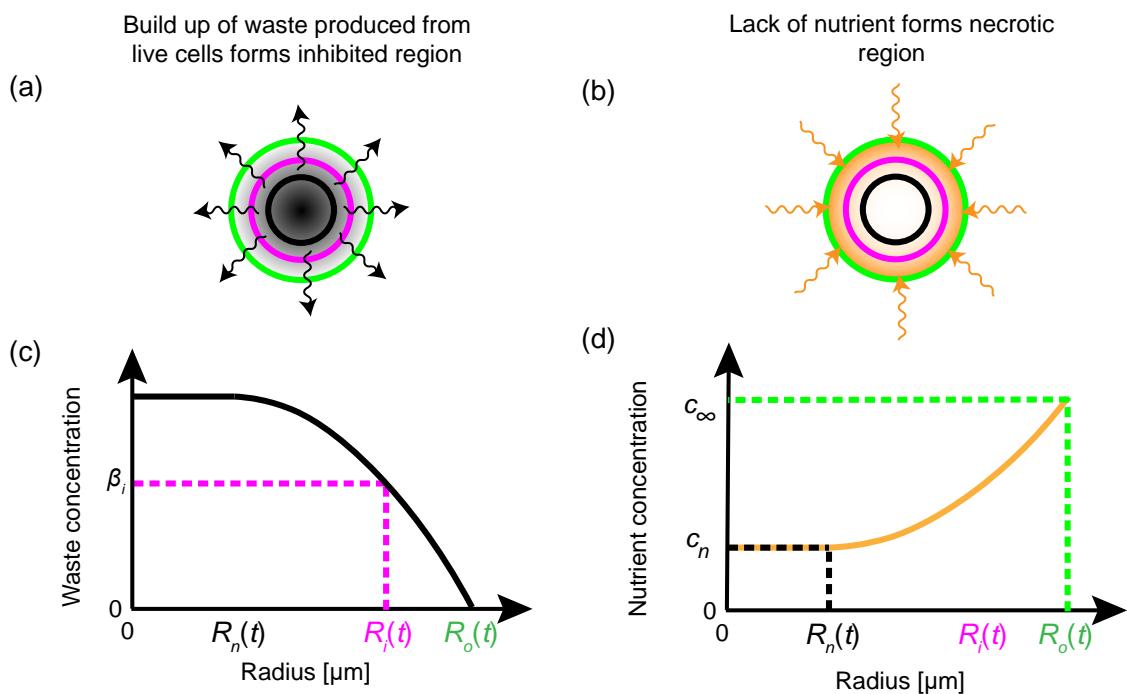
$$\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \beta(r(t)) \right) = \frac{P}{\kappa} H(R_n(t) - r), \quad 0 \leq r \leq R_o(t), \quad (6A.1.23.3)$$

where equations (6A.1.23.1) and (6A.1.23.2) are unchanged, by restricting our attention to the case when the inhibited region forms before the necrotic region, and equation (6A.1.23.3) is new. In equation (6A.1.23.3), the term on the right-hand-side corresponds to production of waste by live cells at a constant rate per unit volume  $P$  that diffuses with diffusivity  $\kappa$ . Furthermore,  $R_i(t)$  is defined as the solution of  $\beta(R_i(t), t) = \beta_i$  if a solution exists and  $R_i(t) = 0$  otherwise.

This model, with nutrient and waste, is equivalent to the nutrient only model when we focus on the five key parameters  $R_o(0), R_c, s, \gamma, Q$  governing the dynamics. The only difference is a new definition of  $Q$ ,

$$Q^2 = \underbrace{\frac{\beta_i \kappa}{P}}_{\text{waste parameters}} \underbrace{\frac{\alpha}{k(c_\infty - c_n)}}_{\text{oxygen parameters}}. \quad (6A.1.24)$$

This new definition of  $Q$  provides a different interpretation of the data since  $Q$  now represents a combination of waste and nutrient parameters. Importantly, with this new definition of  $Q$  there are two cases to consider: i)  $Q \leq 1$ , and ii)  $Q > 1$ . Previously, we only considered case (i). In



**Figure 6A.1:** Greenspan's model with waste and nutrient. (a),(c) Build up of waste from live cells results in the formation of an inhibited region. (a) Schematic of waste produced from live cells and diffusing to the external environment. (c) Snapshot of waste concentration against spheroid radius for a spheroid in phase (iii).  $R_i(t)$  is determined by the inhibited waste threshold  $\beta_i$ . (b),(d) Lack of nutrient forms the necrotic region. (b) Nutrient, shown in orange, diffusing into the spheroid. (d) Snapshot of nutrient concentration against spheroid radius for a spheroid in phase (iii).  $R_n(t)$  is determined by the necrotic nutrient threshold  $c_n$ . External nutrient concentration is  $c_\infty$ .

case (ii) where  $Q > 1$  the necrotic core forms before the inhibited region. We do not observe this scenario in the experiments that we perform and therefore we restrict the attention of this study to case (i) with  $Q \leq 1$ .

## 6A.2 Profile likelihood further details

### 6A.2.1 Numerical method

Parameter identifiability using statistical profile likelihood analysis is outlined in the methods section 6.5.2. We now provide further details.

We partition the full set of observations  $y^o$  into sets of observations  $y_o^o$ ,  $y_n^o$ , and  $y_i^o$  corresponding to experimental measurements of  $R_o(t)$ ,  $R_n(t)$ , and  $R_i(t)$ . For computational accuracy, we perform calculations using the log-likelihood which is, assuming data independence,

$$\begin{aligned} l(\theta; y^o) = & \sum_{j=1}^{N_o} \log [f(y_{o,j}^o; y_{o,j}(\theta), \sigma_o^2)] + \sum_{j=1}^{N_n} \log [f(y_{n,j}^o; y_{n,j}(\theta), \sigma_n^2)] \\ & + \sum_{j=1}^{N_i} \log [f(y_{i,j}^o; y_{i,j}(\theta), \sigma_i^2)], \end{aligned} \quad (6A.2.1)$$

where  $y_{o,j}(\theta)$ ,  $y_{n,j}(\theta)$ , and  $y_{i,j}(\theta)$  are values of  $R_o(t)$ ,  $R_n(t)$ , and  $R_i(t)$  generated from Greenspan's deterministic mathematical model and evaluated at time points corresponding to the experimental observations  $y_{o,j}^o$ ,  $y_{n,j}^o$ , and  $y_{i,j}^o$ , respectively;  $f(x; \mu, \sigma^2)$  denotes a Gaussian probability density function with mean  $\mu$  and variance  $\sigma^2$ , calculated using MATLAB's `normpdf` function [3];  $N_o$ ,  $N_n$ , and  $N_i$  denote the total number of experimental observations of  $R_o(t)$ ,  $R_n(t)$ , and  $R_i(t)$ , respectively; and,  $\sigma_o^2$ ,  $\sigma_n^2$ , and  $\sigma_i^2$  correspond to pooled variances of the three measurement types  $R_o(t)$ ,  $R_n(t)$ , and  $R_i(t)$ , respectively [97, 244]. We approximate  $\sigma_o^2 \approx s_o^2$ ,  $\sigma_n^2 \approx s_n^2$ , and  $\sigma_i^2 \approx s_i^2$ , where  $s_o^2$ ,  $s_n^2$ ,  $s_i^2$  are pooled sample variances of the outer, necrotic, and inhibited radius measurements, respectively [31]. The pooled sample variance for the outer radius is defined as

$$s_o^2 = \frac{1}{N_o - 1} \sum_{j=1}^{N_o} \left( y_{o,j}^o - \bar{y}_{o,j}^o \right)^2, \quad (6A.2.2)$$

where  $y_{o,j}^o$  is the  $j^{\text{th}}$  observation in  $y_o^o$  and  $\bar{y}_{o,j}^o$  is the sample mean of  $y_o^o$  corresponding to the time at which the  $j^{\text{th}}$  measurement was observed. We define  $s_n^2$  and  $s_i^2$  similarly.

The maximum likelihood estimate (MLE),  $\hat{\theta}$ , is defined as,

$$\hat{\theta} = \operatorname{argmax}_{\theta} [l(\theta; y^o)], \quad (6A.2.3)$$

which we determine by numerically solving the equivalent minimisation problem,

$$\hat{\theta} = \operatorname{argmin}_{\theta} [-(\theta; y^0)]. \quad (6A.2.4)$$

By assuming the full parameter  $\theta$  can be partitioned into an interest scalar parameter,  $\psi$ , and a nuisance vector parameter,  $\phi$ , the profile log-likelihood is

$$\begin{aligned} l_p(\psi; y^0) = \max_{\phi} & \left[ \sum_{j=1}^{N_o} \log [f(y_{o,j}^0; y_{o,j}(\psi, \phi), \sigma_o^2)] + \sum_{j=1}^{N_n} \log [f(y_{n,j}^0; y_{n,j}(\psi, \phi), \sigma_n^2)] \right. \\ & \left. + \sum_{j=1}^{N_i} \log [f(y_{i,j}^0; y_{i,j}(\psi, \phi), \sigma_i^2)] \right]. \end{aligned} \quad (6A.2.5)$$

Given the five-dimensional parameter space that we are searching to find the maximum likelihood estimate and the four-dimensional parameter space we search to find profile likelihoods, we sequentially determine the maximum likelihood estimate (MLE) and profile likelihoods. All subsequent minimisation optimisations are performed using functions in MATLABs global optimisation toolbox. Specifically, we use the *GlobalSearch* function [146] where we create the following optimisation problem structure. We set the local solver to be the *fmincon* function using the sequential quadratic programming (sqp) algorithm, `MaxIterations` = 2500 and `MaxFunctionEvaluations` = 5000. The objective function is defined as the argument of the minimisation of the right hand side of equation (6A.2.4). Other non-default settings that we vary, include `NumTrialPoints`, `MaxTime`, `FirstGuess`, `lowerbounds`, `upperbounds`, along with the method we use to find the MLE and approximate 95% confidence intervals are now discussed.

1. Firstly, we search for MLE. We set the `lowerbounds` and `upperbounds` in agreement with the simple parameter bounds defined in the methods section 6.5.2. By setting `NumTrialPoints` = 5000 and `MaxTime` = 7200 [seconds], we search for the maximum likelihood estimate for 2 hours with the `FirstGuess` as  $(Q, \gamma, s, R_c, R_o(0)) = (0.9, 3, 0.5, 175, 125)$ . This gives the first estimate for the maximum likelihood estimate  $\hat{\theta}_1$ . However, numerical experimentation indicates this first estimate is not always an accurate estimate of the true MLE.
2. Secondly, we partition the simple parameter bounds into two sets:  $[\text{lowerbounds}, \hat{\theta}_1]$  which we refer to as the lower set, and  $[\hat{\theta}_1, \text{upperbounds}]$  which we refer to as the upper set. We then discretise each lower and upper set uniformly using 20 grid points, including the end points. Starting at the grid point associated with  $\hat{\theta}_1$  we set `FirstGuess`

$= \hat{\theta}_1$ , `NumTrialPoints = 2000` and `MaxTime = 900` [seconds] in the `GlobalSearch` function. We then move to the next closest grid point and adjust `FirstGuess`. If we are at the closest grid point to  $\hat{\theta}_1$  we set `FirstGuess` to be the solution at the previous gridpoint. If we are at any other grid point we make a first order approximation of the first guess by linear extrapolation of the values obtained from the two previous grid points. Before using the first order approximation as a first guess we also check that the value remains within the parameter bounds and if it does not we set `FirstGuess` to be the solution at the previous gridpoint. After calculating the likelihood at each point in the lower and upper set we combine these together to form the first approximation for the profile likelihood.

3. Thirdly, we calculate an estimate of the confidence intervals using a profile likelihood threshold value of 0.15, which can be approximately calibrated via simulation or the  $\chi^2$  – distribution [180]. Specifically, we start at either end of the simple parameter bounds until we determine the first grid point where the normalised profile likelihood,  $L_p(\psi; y^0) = \exp\left(l_p(\psi; y^0) - \max_{\theta}[(\theta; y^0)]\right)$ , is greater than 0.15. We then set new lower and upper bounds as being two grid points to the left or right of that location, respectively. Note that a more sophisticated approach to determine the approximate 95% confidence intervals is applied in step seven to compute the results shown in Table 6A.3, which is not required here.
4. Fourth, we repeat the search for the maximum likelihood estimate using the new lower and upper bounds with the same settings as we first used.
5. Fifth, we repeat the calculations for the profile likelihoods using the new lower and upper bounds.
6. Sixth, we determine the maximum likelihood estimate to be the value across all calculations which maximises the likelihood. We form the final profile likelihood from steps two and four and present in figures the normalised likelihood function,  $L_p(\psi; y^0) = \exp\left(l_p(\psi; y^0) - \max_{\theta}[(\theta; y^0)]\right)$ .
7. To compute approximate 95% confidence intervals for each parameter, as shown in Table 6A.3, we form the profile likelihood from steps two and four. Next, we start at either end of the simple parameter bounds until we determine the first grid point where the normalised profile likelihood,  $L_p(\psi; y^0) = \exp\left(l_p(\psi; y^0) - \max_{\theta}[(\theta; y^0)]\right)$ , is greater than profile likelihood threshold value of 0.15 [180]. These two grid points are a first approximation of the lower and upper 95% confidence interval boundaries. Finally, to obtain a more

accurate estimate of the approximate 95% confidence interval boundaries, we consider each of these two grid points in turn as the `FirstGuess` for the MATLAB function `fsove` [1], and use linear interpolation, with the MATLAB function `interp1` [2].

### 6A.2.2 Parameter bounds

To interpret  $s$  we consider the evolution of the tumour spheroid in phase (i). Equation (6A.1.13) can be written in terms of volume  $V(t)$ , recalling that the volume of a sphere is  $4\pi R_o^3(t)/3$ , as

$$V(t) = V(0) \exp(st), \quad (6A.2.6)$$

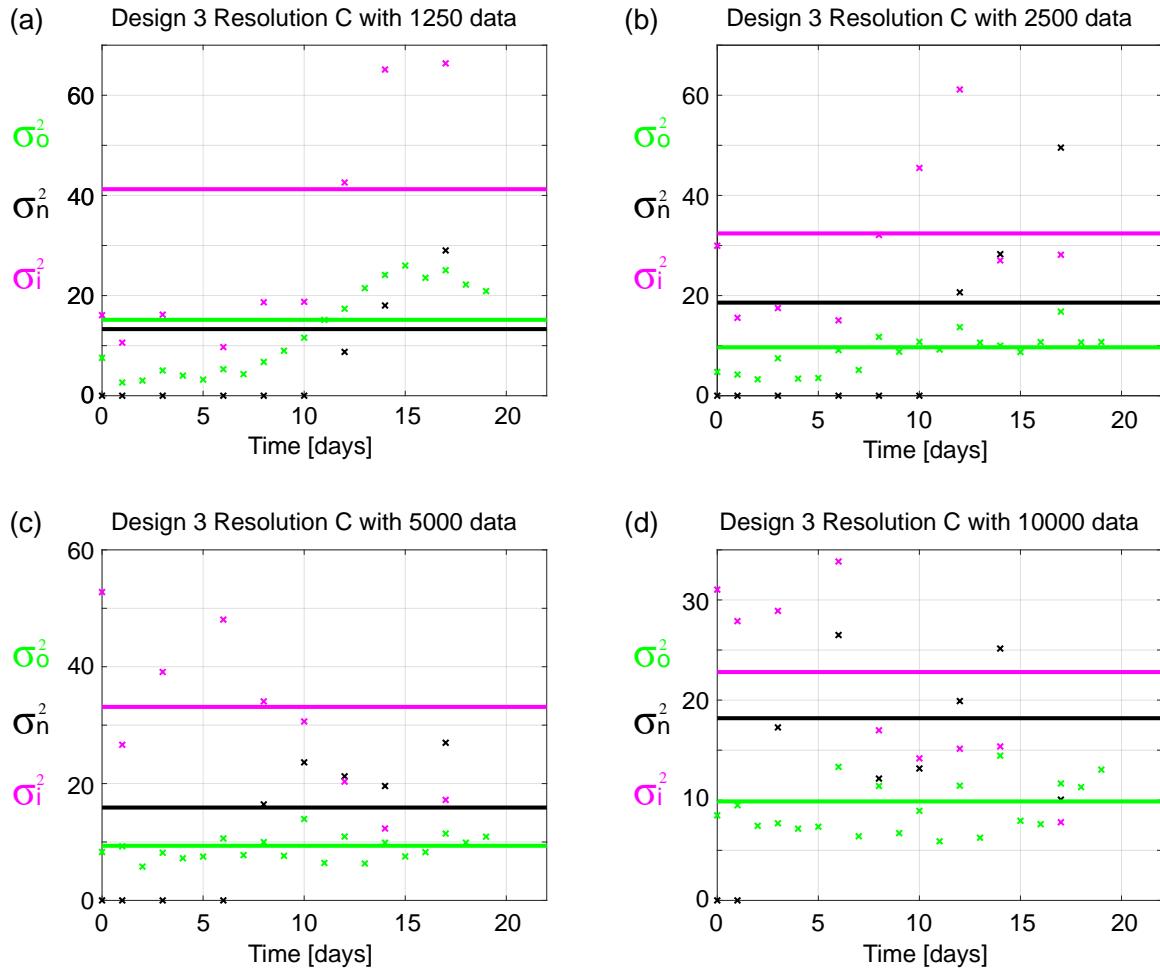
where  $V(0) = 4\pi R_o^3(0)/3$ . Then by letting  $T$  define the time when  $V(T) = 2V(0)$ , we relate  $s$  to the doubling time of the cells through

$$s = \frac{1}{T} \log_e(2). \quad (6A.2.7)$$

Then assuming that the doubling time is greater than 12 hours ( $= 1/2$  day) we obtain an upper bound of  $s$  as  $2 \log_e(2)$  [day $^{-1}$ ]  $\approx 1.39$  [day $^{-1}$ ]. Preliminary exploration confirms this estimate is very conservative so we set the upper bound to unity.

### 6A.2.3 Pooled sample variances

To identify parameters we assume distinct pooled sample variances for the outer, necrotic, and inhibited radius measurements, as opposed to a single pooled variance for all measurements. In Figure 6A.2 we plot the pooled sample variances for different experimental designs which justify the use of a variance for each measurement type.



**Figure 6A.2:** Variances  $\sigma_o^2$ ,  $\sigma_n^2$ , and  $\sigma_i^2$ , for outer, necrotic and inhibited radii, respectively. Results shown for WM793b spheroids, using Design 3 and Temporal Resolution C, formed with (a) 1250, (b) 2500, (c) 5000, (d) 10000 cells per spheroid, respectively.

## 6A.3 Experimental data

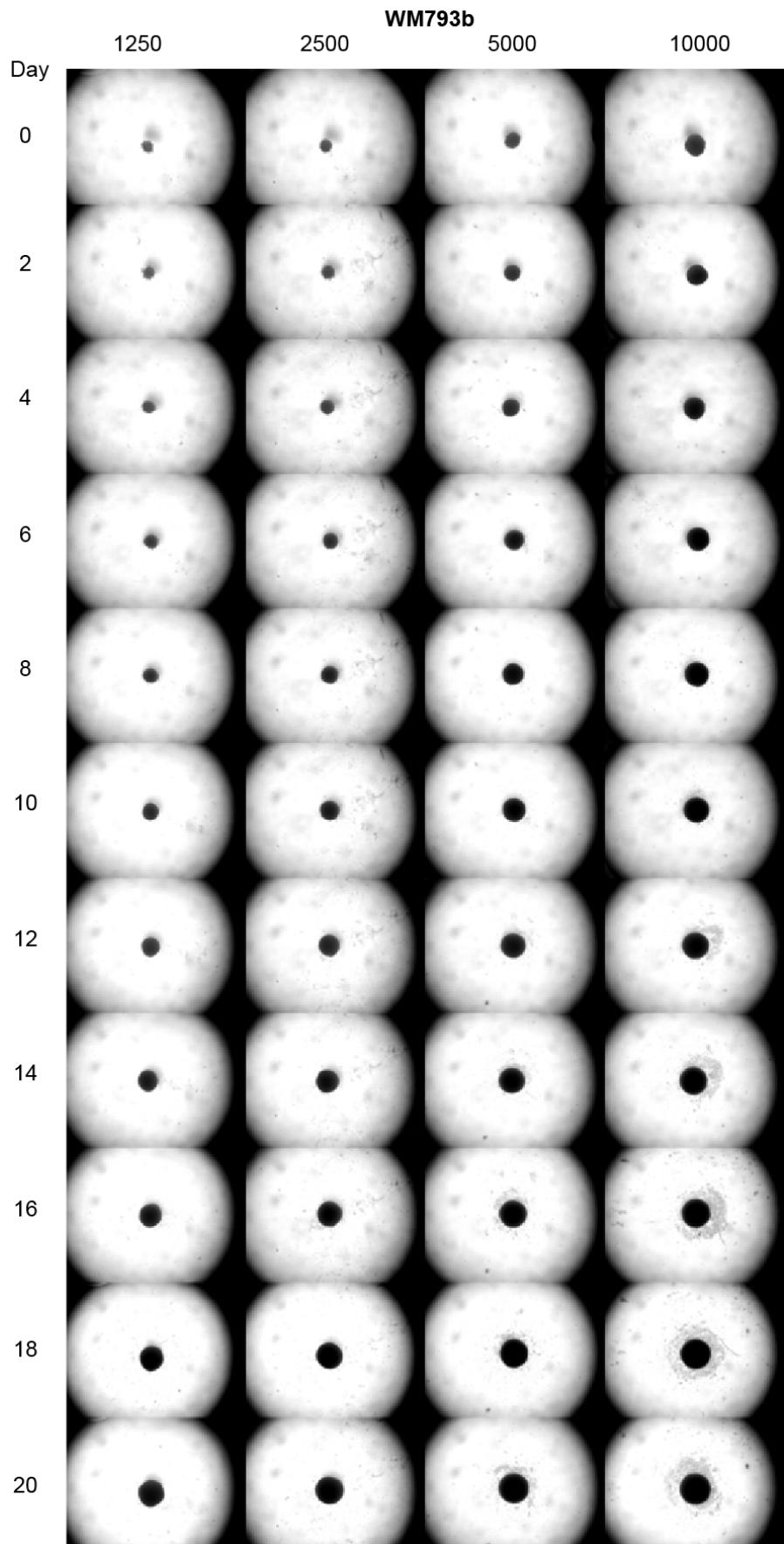
### 6A.3.1 Outer radius experimental measurements and images

The IncuCyte S3 live cell imaging system is a useful tool that we use to obtain many outer radius measurements. The other outer radius measurements are obtained from confocal microscopy.

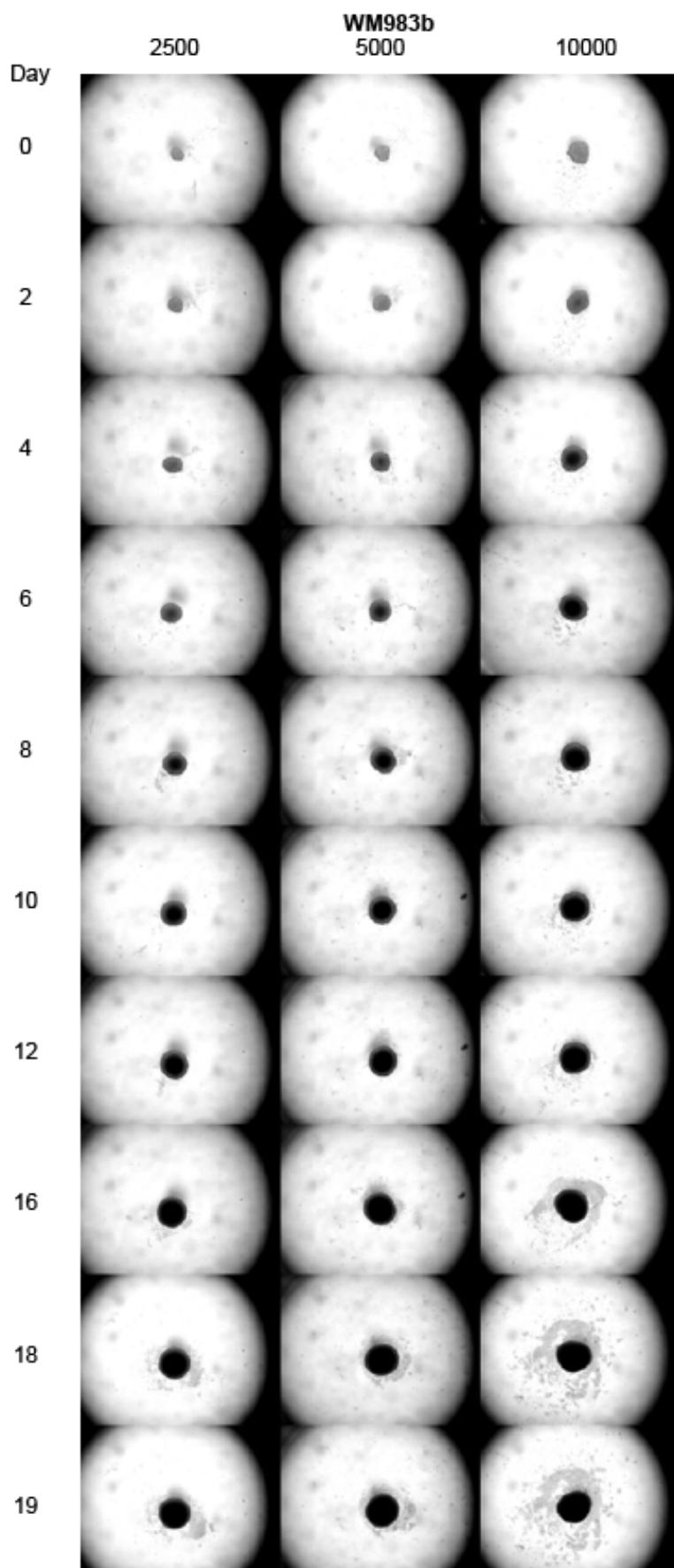
We start with 24 spheroids in the IncuCyte S3 live cell imaging system for each cell line and initial spheroid size and image every 6 hours for the duration of the experiment. However, some measurements could not be obtained primarily due to blurring of the automated imaging, spheroids not forming properly, or spheroids losing their structural integrity at very late time. In Table 6A.1 we show the total number of measurements obtained at 24 hour intervals starting from Day 0 which corresponds to the time that we determined as when spheroid formation ends and growth begins (Supplementary Material 6A.3.1). In Figures 6A.3-6A.5 we present representative experimental images obtained from IncuCyte S3 live cell imaging system for different days and WM793b, WM983b, and WM164 cell lines, respectively.

	WM793				WM983b			WM164			
Day	1250	2500	5000	10000	2500	5000	10000	1250	2500	5000	10000
0	20	24	24	23	24	22	23	23	22	23	21
1	20	23	24	23	24	22	23	24	22	24	21
2	20	23	24	23	24	22	23	18	22	24	20
3	21	23	24	23	24	22	23	19	23	24	20
4	21	24	24	23	24	22	23	18	23	24	20
5	21	23	24	23	24	22	22	19	20	19	-
6	21	24	24	23	24	22	21	19	19	19	-
7	21	24	24	23	23	22	20	-	-	-	-
8	21	24	24	23	23	22	20	-	-	-	-
9	21	24	24	23	23	22	20	-	-	-	-
10	21	24	24	23	23	22	20	-	-	-	-
11	21	24	24	23	23	22	20	-	-	-	-
12	21	24	24	23	22	22	20	-	-	-	-
13	21	24	24	23	22	22	20	-	-	-	-
14	21	24	24	23	-	-	-	-	-	-	-
15	20	24	22	23	22	22	20	-	-	-	-
16	18	24	22	23	22	22	20	-	-	-	-
17	18	24	22	23	22	22	20	-	-	-	-
18	19	24	22	23	22	22	20	-	-	-	-
19	19	24	22	23	22	22	20	-	-	-	-
20	13	24	13	11	-	-	-	-	-	-	-

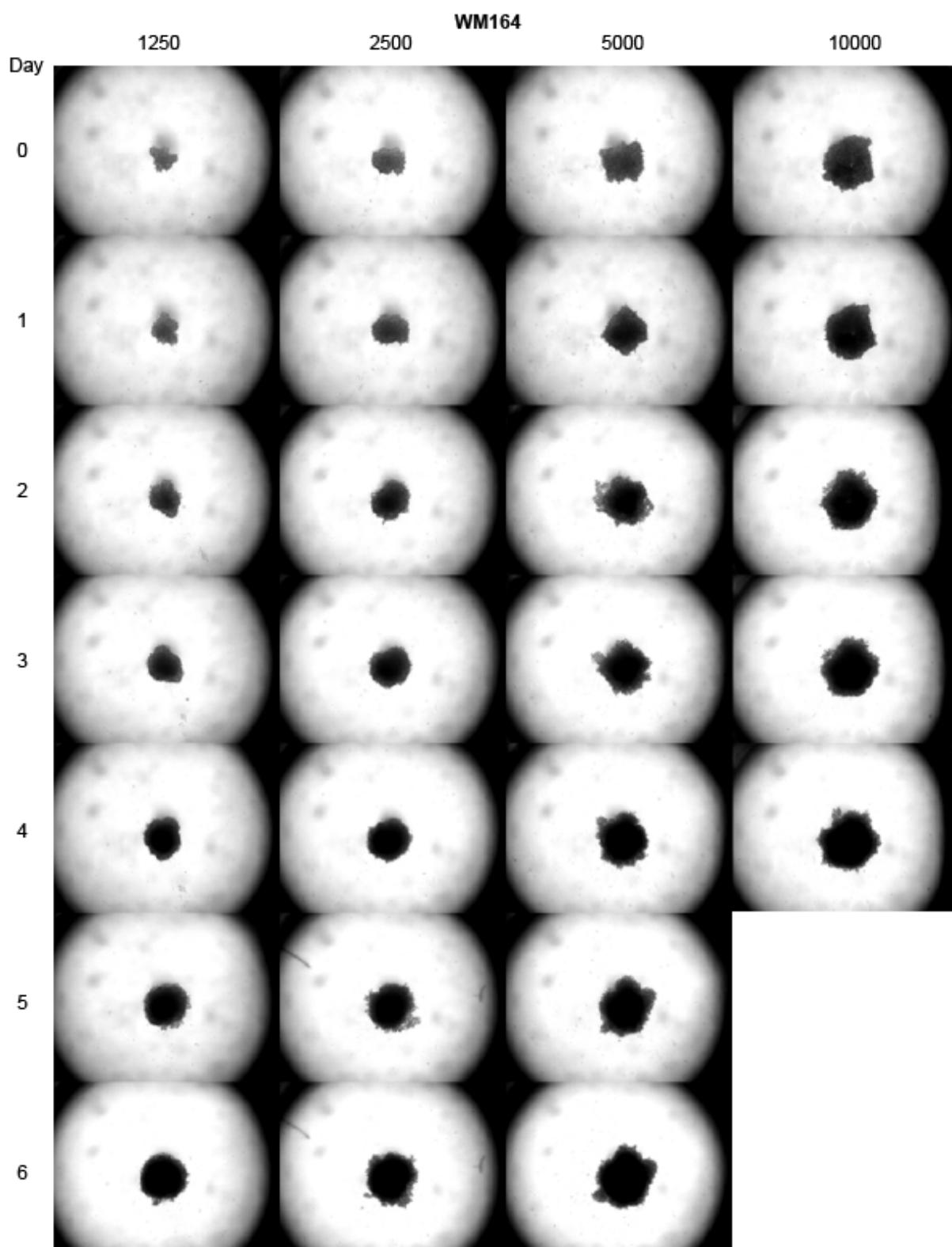
**Table 6A.1:** Number of outer radius measurements obtained from the IncuCyte S3 live cell imaging system for the cell lines WM793b, WM983b, and WM164. Day 0 corresponds to the time that we determined as when spheroid formation ends and growth begins, see supplementary material 6A.3.1.



**Figure 6A.3:** Snapshots of WM793b tumour spheroids from IncuCyte S3 live cell imaging system at 0, 2, 4, 6, 8, 10, 12, 14, 16, 18, and 20 days after formation for tumour spheroids formed with 1250, 2500, 5000, and 10000 cells per spheroid. Each image shows a  $4.34 \times 3.25$  mm field of view.



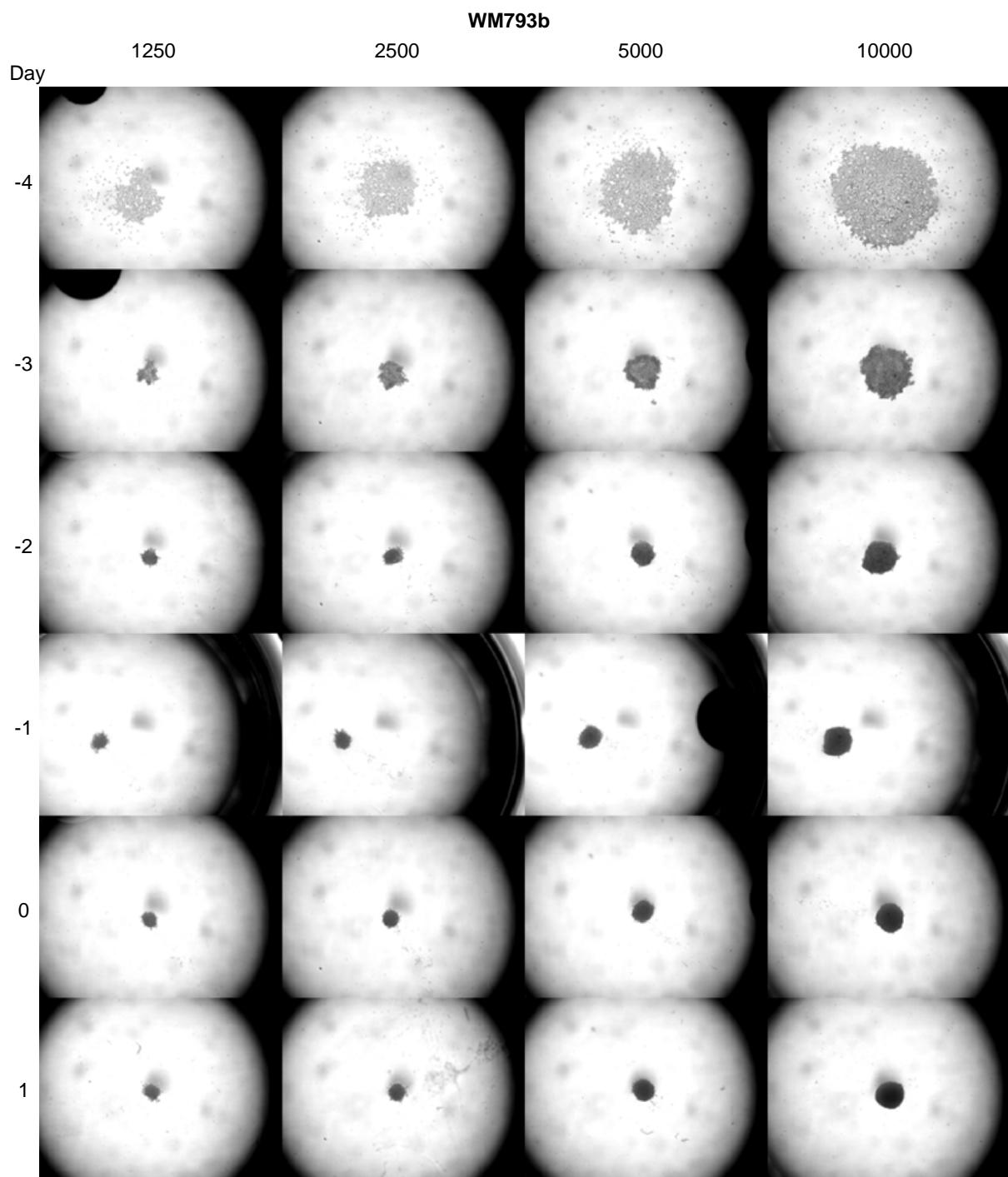
**Figure 6A.4:** Snapshots of WM983b tumour spheroids from IncuCyte S3 live cell imaging system at 0, 2, 4, 6, 8, 10, 12, 16, 18, and 19 days after formation for tumour spheroids formed with 2500, 5000, and 10000 cells per spheroid. Each image shows a 4.34 × 3.25 mm field of view.



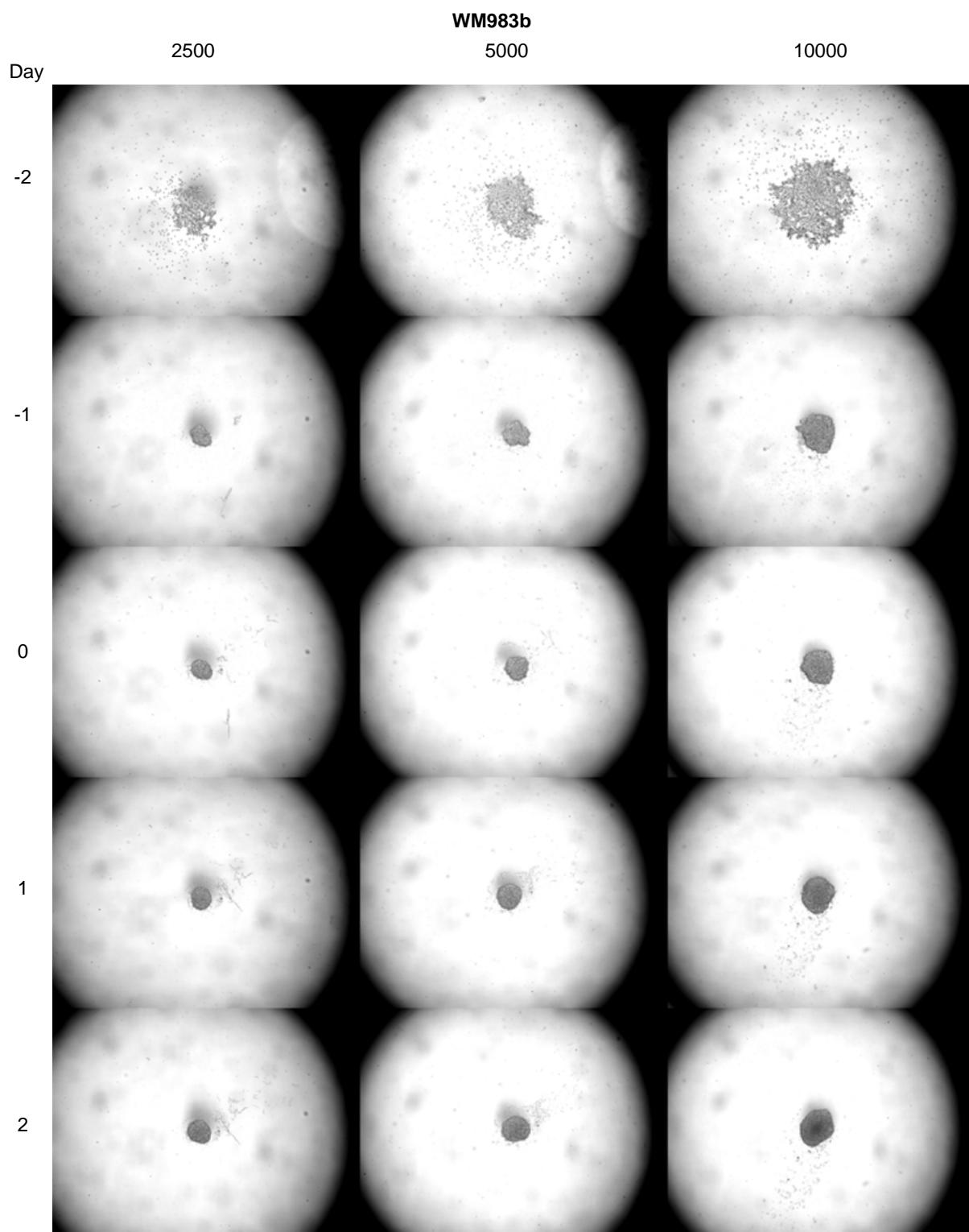
**Figure 6A.5:** Snapshots of WM164 tumour spheroids from IncuCyte S3 live cell imaging system at 0, 1, 2, 3, 4, 5, and 6 days after formation for tumour spheroids formed with 1250, 2500, 5000, and 10000 cells per spheroid. Each image shows a  $4.34 \times 3.25$  mm field of view.

### Spheroid formation duration

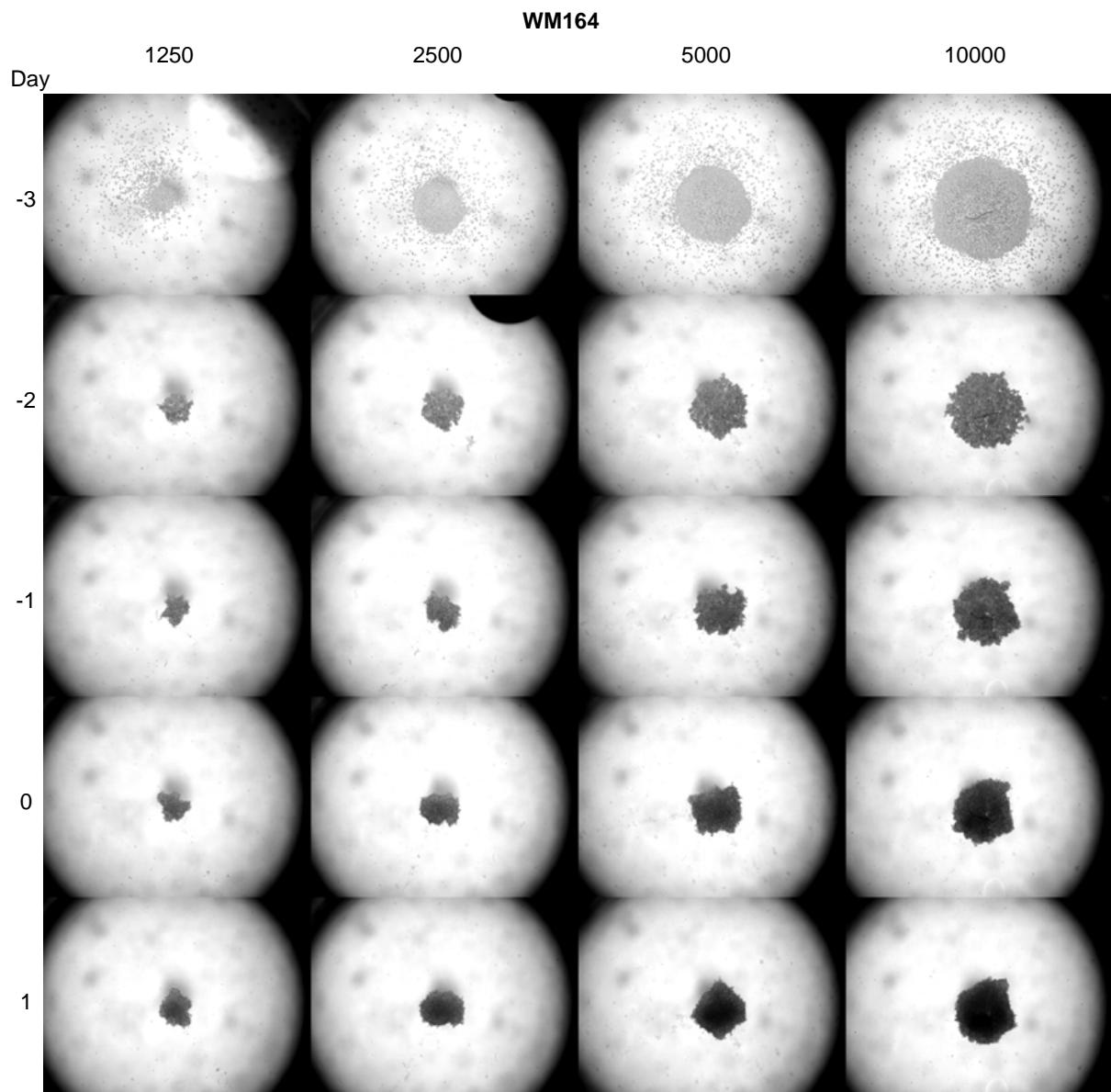
In Figure 6A.6 we show time snapshots of forming tumour spheroids obtained in the IncuCyte S3 live cell imaging system. These snapshots, alongside monitoring the evolution of the outer radius obtained from image processing, validate the assumption that the tumour spheroids have formed 4 days after seeding for WM793b. This method was also used to determine the duration of spheroid formation for the WM983b (Figure 6A.7) and WM164 (Figure 6A.8) cell lines.



**Figure 6A.6:** Spheroids are formed 4 days after seeding for WM793b. Snapshots from IncuCyte S3 live cell imaging system at -4, -3, -2, -1, 0, and 1 days after formation for tumour spheroids formed 1250, 2500, 5000, and 10000 cells per spheroid. Each image shows a 4.34 × 3.25 mm field of view.



**Figure 6A.7:** Spheroids are formed at 2 days after seeding for WM983b. Snapshots from IncuCyte S3 live cell imaging system at -2, -1, 0, 1, and 2 days after formation for tumour spheroids formed with 2500, 5000, and 10000 cells per spheroid. Each image shows a  $4.34 \times 3.25$  mm field of view.



**Figure 6A.8:** Spheroids are formed at 3 days after seeding for WM164. Snapshots from IncuCyte S3 live cell imaging system at -3, -2, -1, 0 and 1 days after formation for tumour spheroids formed with 1250, 2500, 5000, and 10000 cells per spheroid. Each image shows a  $4.34 \times 3.25$  mm field of view.

### 6A.3.2 Confocal microscopy

#### Measurements

In Table 6A.2 we show the number of confocal measurements obtained. Spheroids damaged during harvesting and fixing procedures are not included.

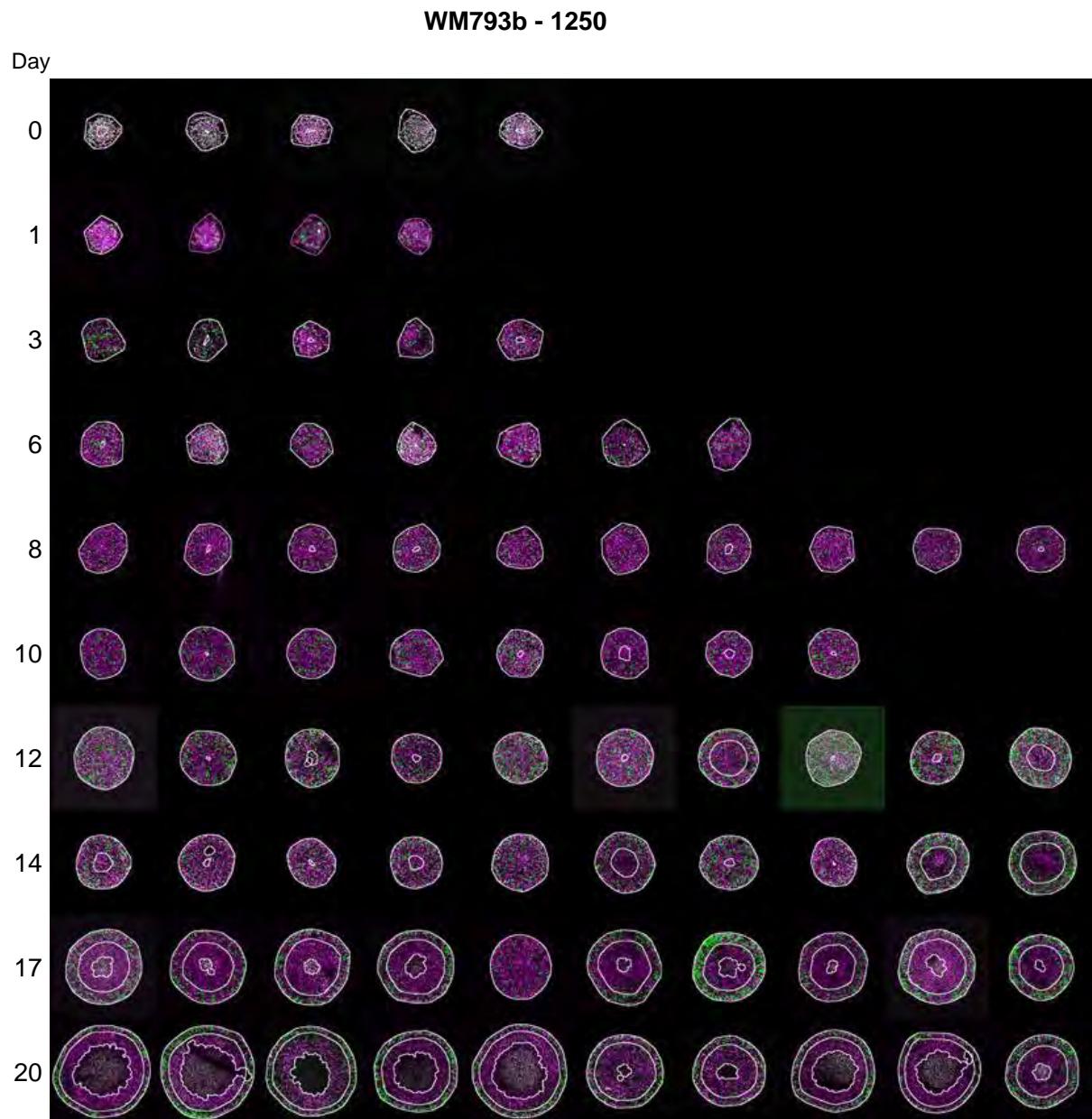
Day	WM793				WM983b			WM164			
	1250	2500	5000	10000	2500	5000	10000	1250	2500	5000	10000
0	5	5	12	7	-	-	-	-	-	-	-
1	4	10	11	12	6	9	6	6	4	14	6
2	-	-	-	-	12	9	10	13	10	10	9
3	5	22	23	18	12	10	9	-	-	-	-
4	-	-	-	-	-	-	-	13	8	-	-
5	-	-	-	-	20	15	18	-	-	-	-
6	7	28	25	25	-	-	-	-	-	-	-
7	-	-	-	-	-	-	-	-	-	-	-
8	12	27	20	23	16	13	15	-	-	-	-
9	-	-	-	-	-	-	-	-	-	-	-
10	8	19	21	15	16	17	21	-	-	-	-
11	-	-	-	-	-	-	-	-	-	-	-
12	12	18	19	17	13	14	13	-	-	-	-
13	-	-	-	-	-	-	-	-	-	-	-
14	15	19	22	21	17	21	19	-	-	-	-
15	-	-	-	-	-	-	-	-	-	-	-
16	-	-	-	-	11	20	19	-	-	-	-
17	11	15	14	5	-	-	-	-	-	-	-
18	-	-	-	-	-	-	-	-	-	-	-
19	-	-	-	-	25	31	16	-	-	-	-
20	22	23	21	20	-	-	-	-	-	-	-

**Table 6A.2:** Number of spheroids imaged using confocal microscopy for the cell lines WM793b, WM983b, and WM164. For each imaged spheroid we obtain a measurement of the outer radius, inhibited radius, and necrotic radius. Day 0 corresponds to the time for the each cell line that we determined as when spheroid formation ends and growth begins. Measurements were taken on days days 3, 4, 5, 7, 10, 12, 14, 16, 18, 21 and 24 after seeding, and appear on different days in the table due to the different formation times.

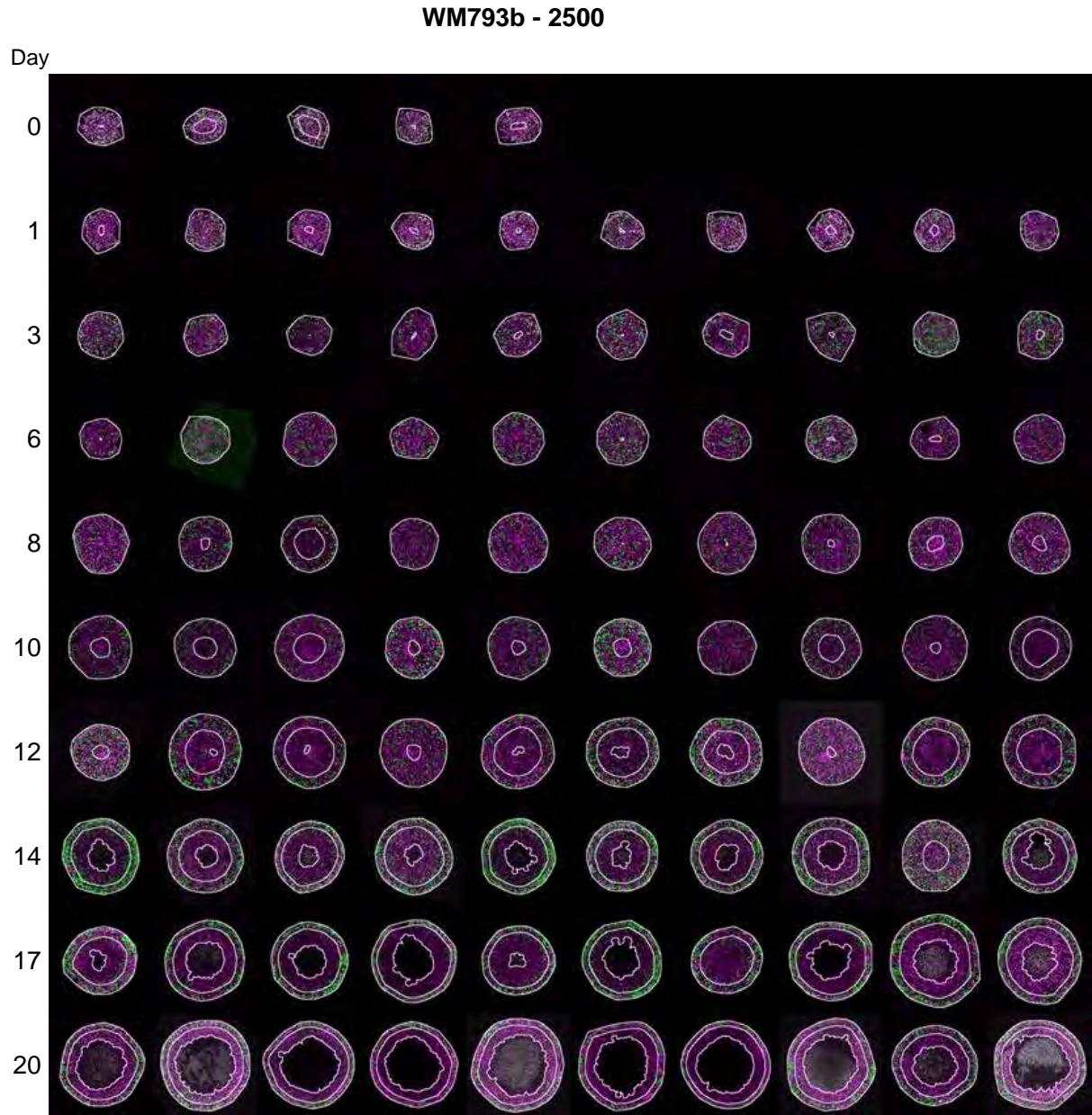
### 6A.3.3 Confocal microscopy supplementary experimental images

Here we present confocal microscopy images of spheroids formed with the WM793b, WM983b, and WM164 cell lines. In the images we outline each spheroids outer boundary, inhibited region, and necrotic region.

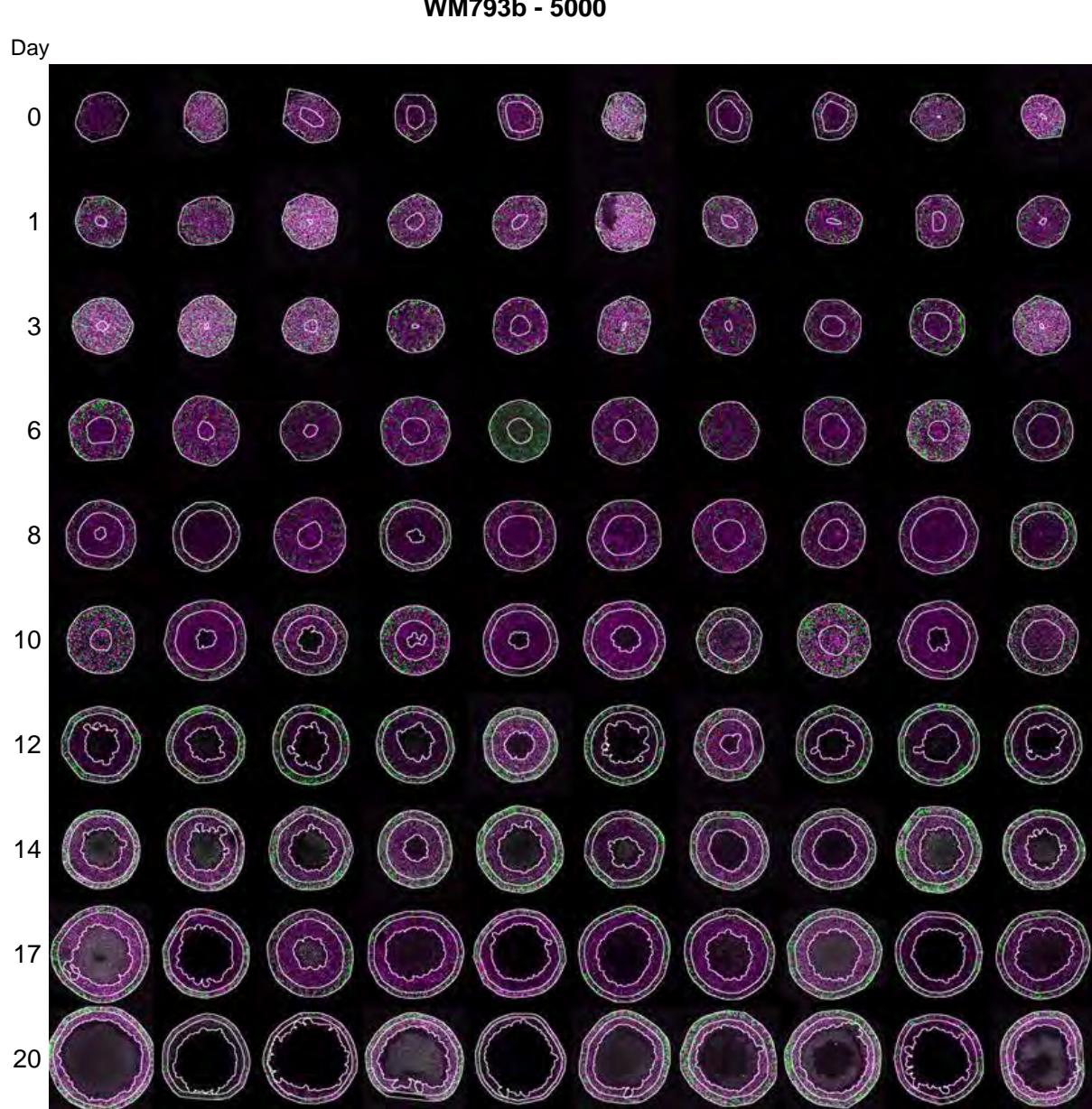
#### WM793b



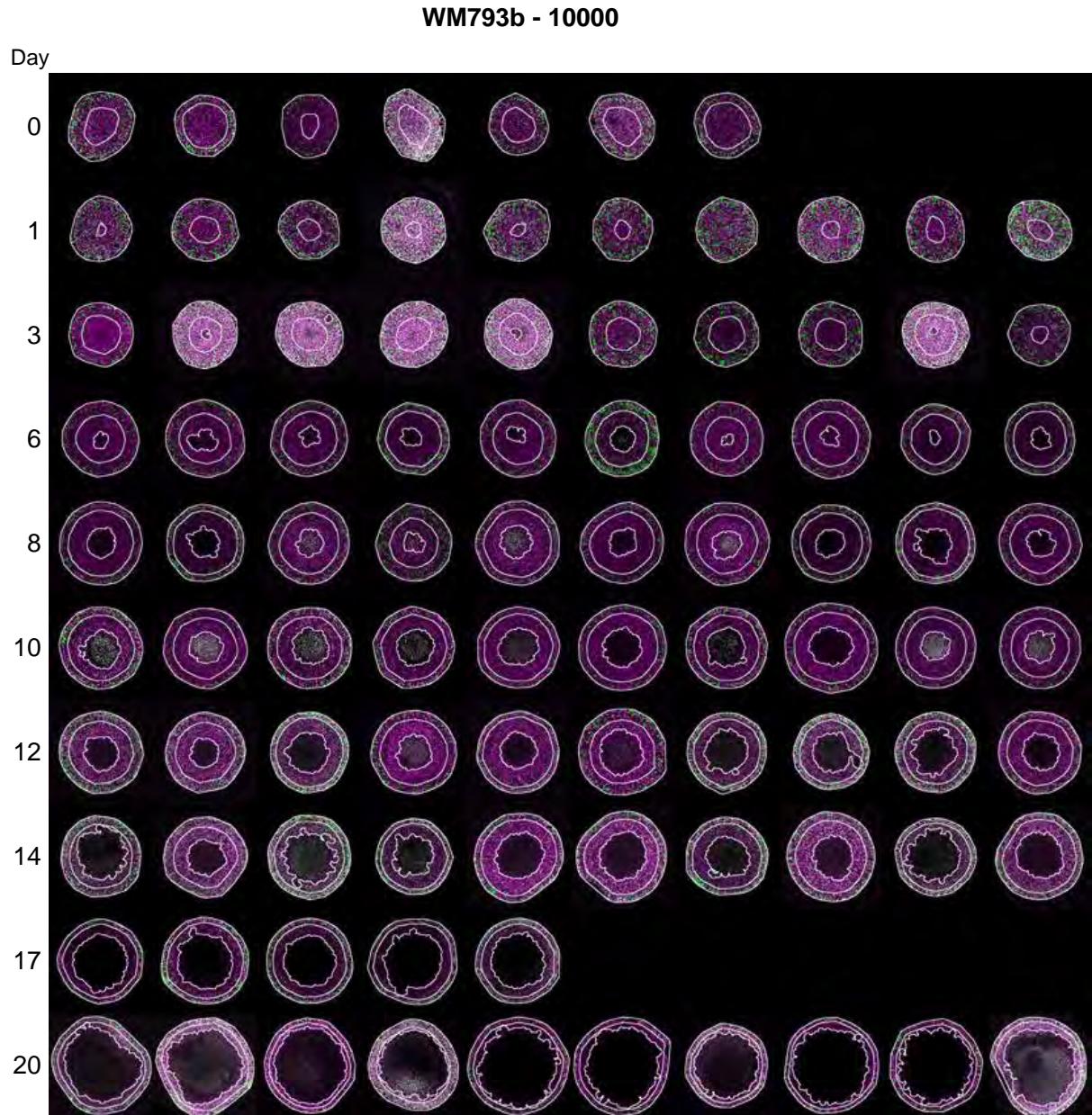
**Figure 6A.9:** Experimental images of WM793b tumour spheroids formed with 1250 cells per spheroid. Each image shows a 800 $\times$ 800  $\mu\text{m}$  field of view.



**Figure 6A.10:** Experimental images of WM793b tumour spheroids formed with 2500 cells per spheroid. Each image shows a 800×800  $\mu\text{m}$  field of view.



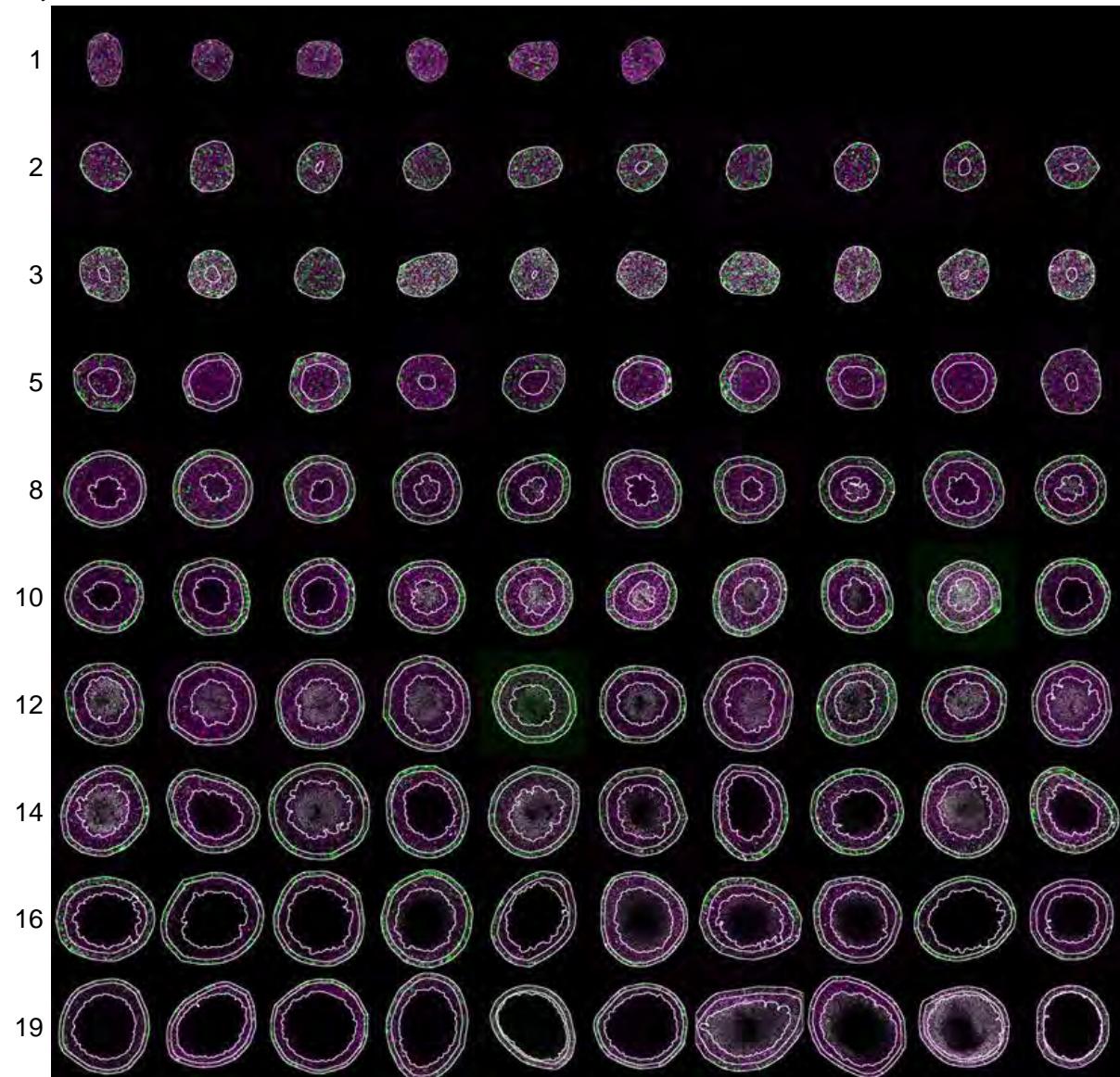
**Figure 6A.11:** Experimental images of WM793b tumour spheroids formed with 5000 cells per spheroid. Each image shows a  $800 \times 800 \mu\text{m}$  field of view.



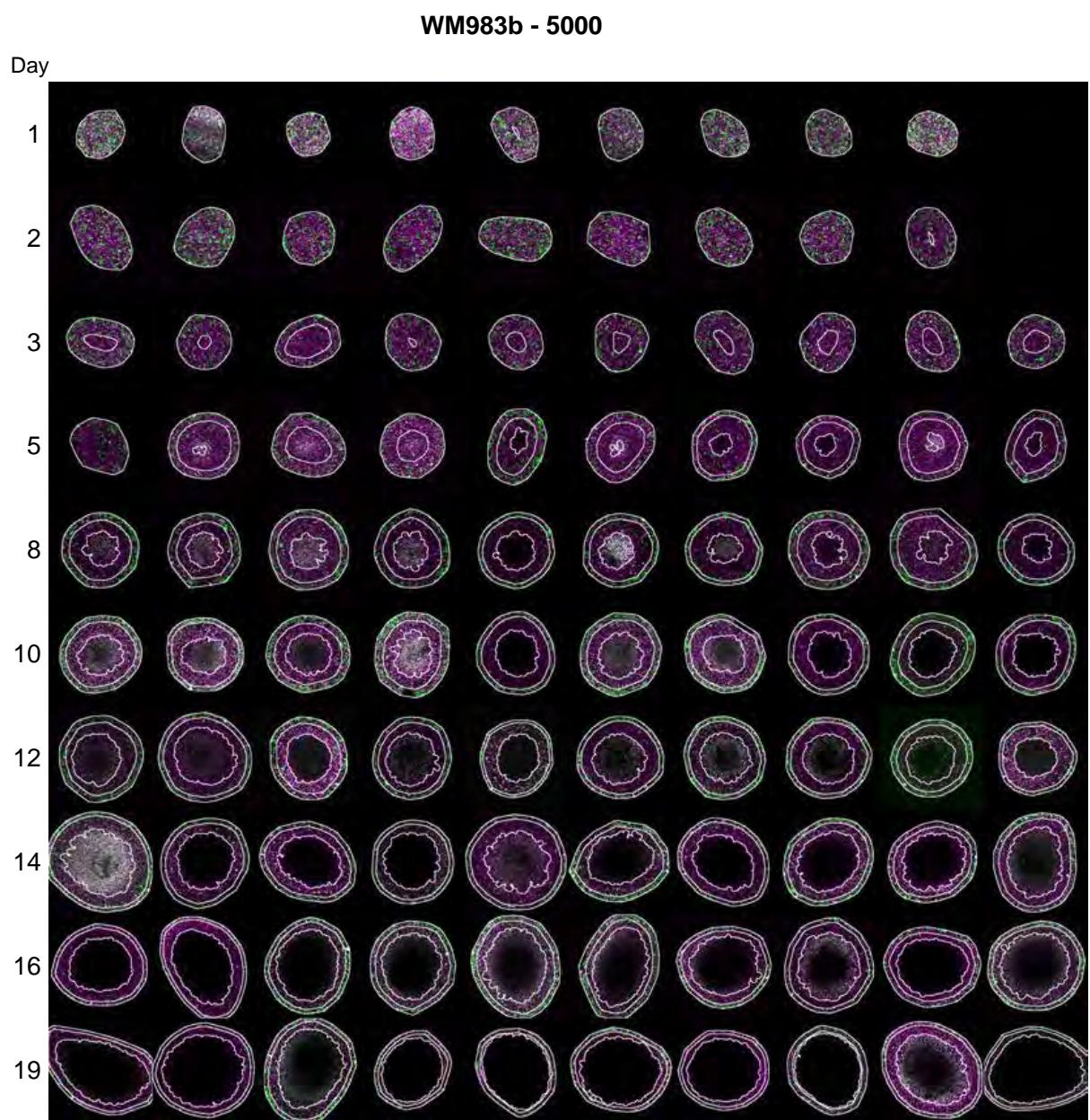
**Figure 6A.12:** Experimental images of WM793b tumour spheroids formed with 10000 cells per spheroid. Each image shows a  $800 \times 800 \mu\text{m}$  field of view.

**WM983b****WM983b - 2500**

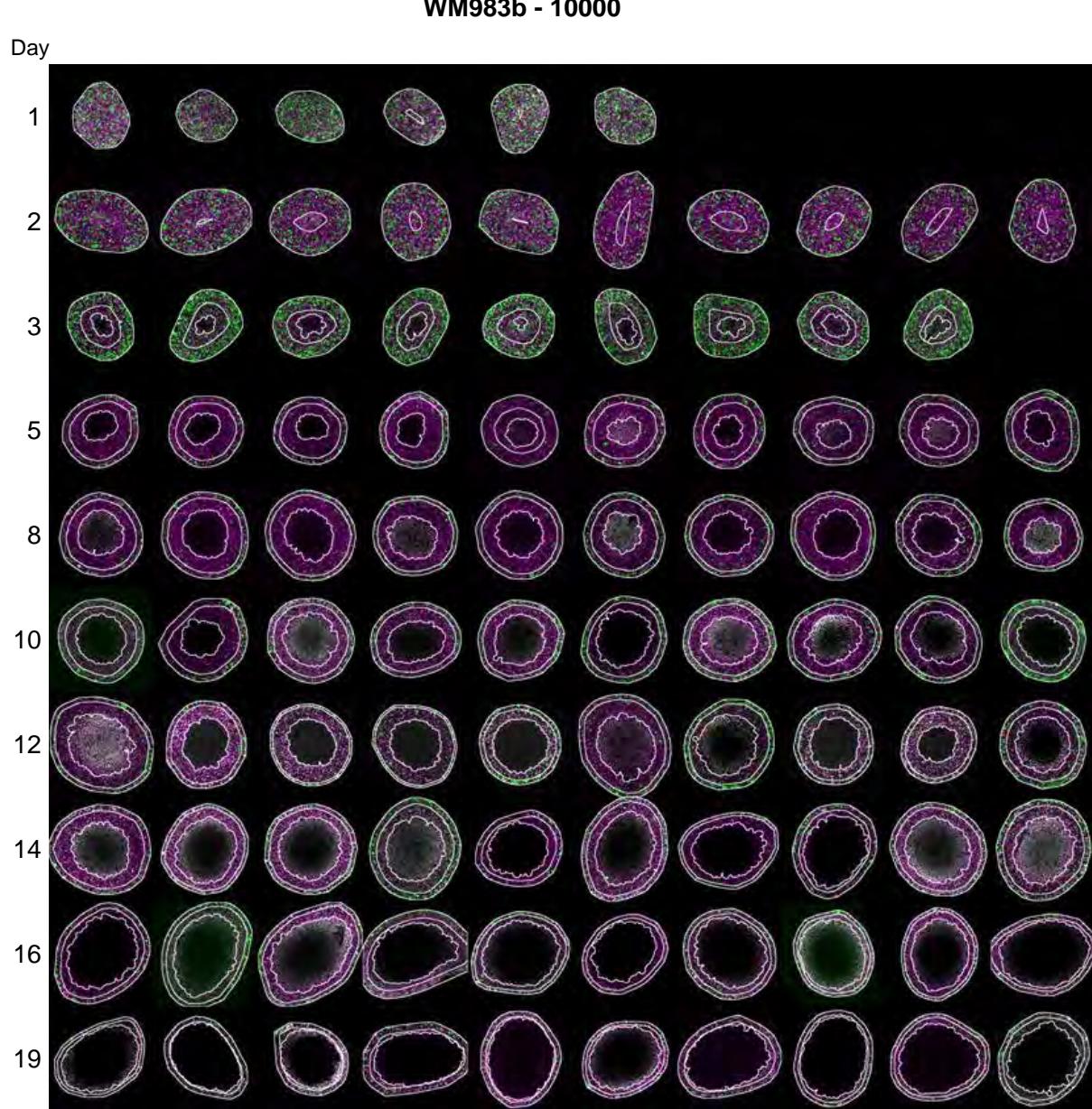
Day



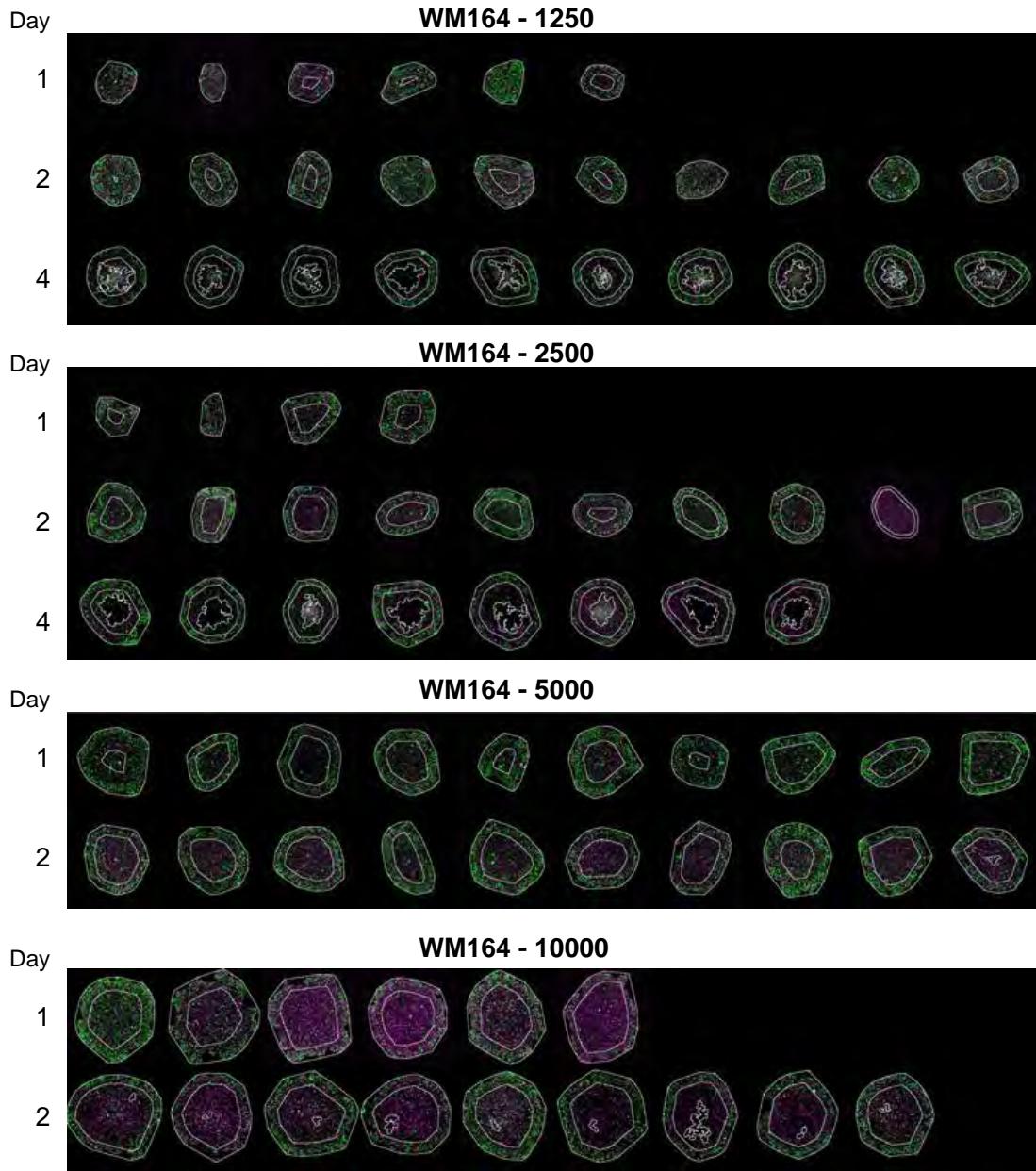
**Figure 6A.13:** Experimental images of WM983b tumour spheroids formed with 2500 cells per spheroid. Each image shows a  $800 \times 800 \mu\text{m}$  field of view.



**Figure 6A.14:** Experimental images of WM983b tumour spheroids formed with 5000 cells per spheroid. Each image shows a  $800 \times 800 \mu\text{m}$  field of view.



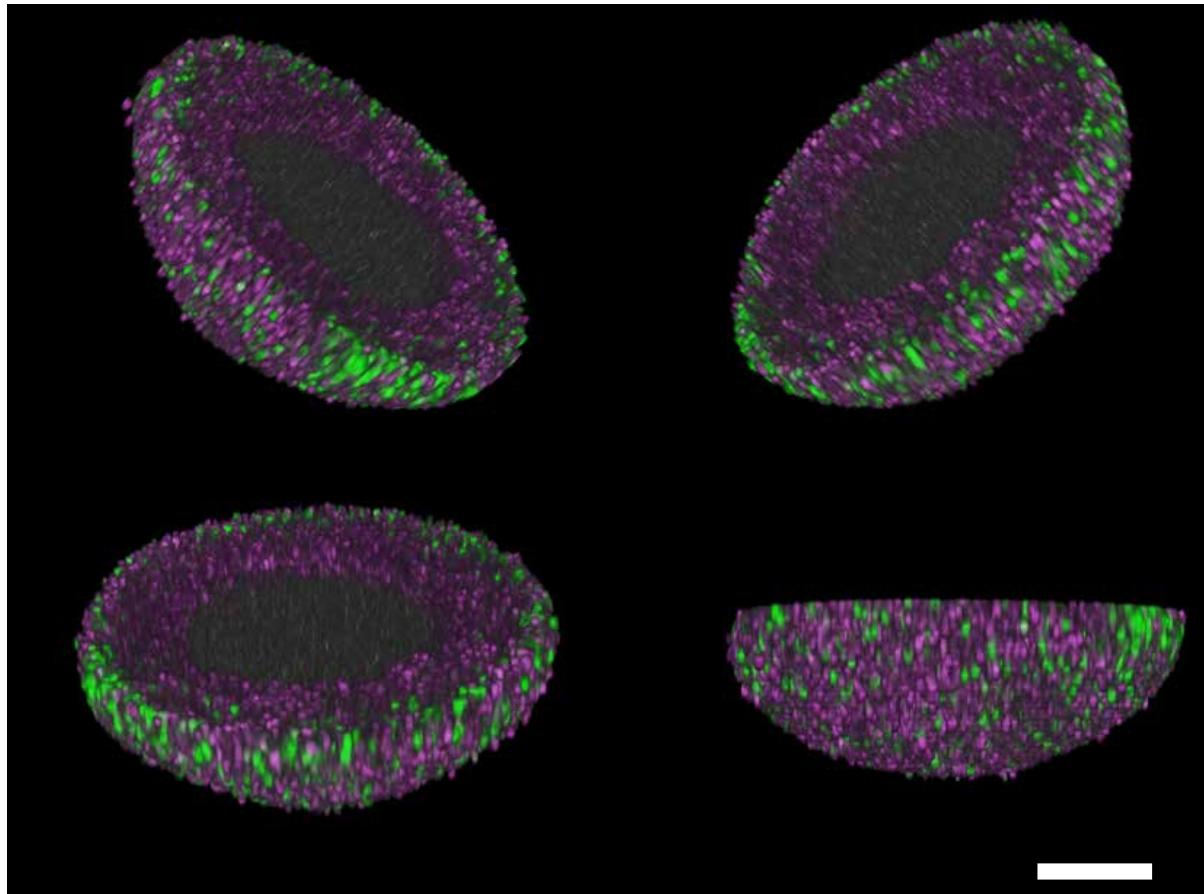
**Figure 6A.15:** Experimental images of WM983b tumour spheroids formed with 10000 cells per spheroid. Each image shows a  $800 \times 800 \mu\text{m}$  field of view.

**WM164**

**Figure 6A.16:** Experimental images of WM164 tumour spheroids formed with 1250, 2500, 5000, and 10000 cells per spheroid. Each image shows a 800×800 µm field of view.

**3D rendering**

Here we present a 3D rendering of a confocal microscopy image z-stack of half of a FUCCI-melanoma WM793b spheroid 17 days after formation with 5000 cells.



**Figure 6A.17:** 3D rendering of half of a FUCCI-melanoma WM793b spheroid 17 days after formation with 5000 cells. Scale bar 200  $\mu\text{m}$ .

## 6A.4 WM793b additional results

### 6A.4.1 Results in tables

In all figures with profile likelihoods we include a red-dashed horizontal line at 0.15 indicating the 95% confidence interval threshold value [180]. Here, in Table 6A.3 we present the corresponding MLE's and approximate 95% confidence intervals for a range of experimental designs.

Cell line	Figure	Experimental Design	$Q$	$\gamma$	$s [\text{day}^{-1}]$	$R_c [\mu\text{m}]$	$R_o(0) [\mu\text{m}]$
WM793b	2	1 Res A	0.711 (0.479, 1.000)	0.020 (0.010, 6.000)	0.141 (0.127, 1.000)	299.980 (25.000, 350.000)	183.509 (179.118, 186.825)
	1	Res B	0.675 (0.550, 1.000)	0.010 (0.010, 6.000)	0.143 (0.128, 0.324)	312.725 (85.574, 350.000)	182.940 (178.565, 186.470)
	1	Res C	0.885 (0.609, 1.000)	0.484 (0.010, 6.000)	0.134 (0.128, 0.144)	258.100 (203.186, 350.000)	183.415 (181.221, 185.260)
WM793b	3	2 Res A	1.000 (0.940, 1.000)	0.924 (0.299, 1.262)	0.126 (0.121, 0.133)	257.056 (254.484, 259.373)	184.543 (181.892, 187.082)
	3	3 Res A	0.770 (0.753, 0.784)	0.010 (0.010, 0.110)	0.159 (0.153, 0.166)	260.672 (257.408, 262.899)	178.733 (176.178, 181.102)
	4	3 Res A 1250 MLE	0.847 (0.824, 0.869)	0.010 (0.010, 6.000)	0.134 (0.126, 0.141)	250.352 (243.485, 256.425)	126.100 (122.459, 129.875)
WM793b	4	3 Res A 2500 MLE	0.829 (0.817, 0.838)	0.010 (0.010, 0.641)	0.150 (0.145, 0.157)	256.818 (253.607, 259.141)	143.797 (141.164, 146.444)
	4	3 Res A 5000 MLE	0.770 (0.753, 0.784)	0.010 (0.010, 0.110)	0.159 (0.153, 0.166)	260.672 (257.408, 262.899)	178.733 (176.178, 181.102)
	4	3 Res A 10000 MLE	0.832 (0.816, 0.846)	1.130 (0.919, 1.333)	0.149 (0.139, 0.161)	253.692 (251.158, 255.826)	223.118 (219.825, 226.245)
WM983b	S39	3 Res A 2500 MLE	0.801 (0.779, 0.822)	0.395 (0.334, 0.457)	0.319 (0.304, 0.336)	209.774 (205.679, 214.047)	128.274 (125.128, 131.309)
	S39	3 Res A 5000 MLE	0.824 (0.815, 0.832)	0.398 (0.366, 0.445)	0.316 (0.300, 0.333)	210.220 (208.168, 212.156)	155.949 (152.684, 159.075)
	S39	3 Res A 10000 MLE	0.841 (0.818, 0.862)	0.517 (0.475, 0.562)	0.368 (0.341, 0.398)	209.470 (205.893, 212.974)	187.514 (183.013, 191.669)
WM164	S40	3 Res A 1250 MLE	0.891 (0.856, 0.923)	0.010 (0.010, 0.408)	0.430 (0.398, 0.465)	325.842 (315.384, 335.715)	230.322 (222.762, 238.128)
	S40	3 Res A 2500 MLE	0.813 (0.779, 0.842)	0.010 (0.010, 1.006)	0.419 (0.383, 0.487)	357.305 (351.641, 363.356)	275.862 (263.881, 285.260)
	S40	3 Res A 5000 MLE	0.735 (0.530, 0.770)	1.365 (0.010, 6.000)	0.380 (0.298, 0.478)	434.753 (428.901, 600.000)	355.357 (339.002, 370.630)
WM164	S40	3 Res A 10000 MLE	0.701 (0.665, 0.736)	0.127 (0.010, 6.000)	0.287 (0.226, 0.394)	528.532 (519.431, 538.034)	471.672 (450.601, 488.117)

**Table 6A.3:** Most likely estimates and approximate 95% confidence intervals for a range of experimental designs. Results shown to three decimal places.

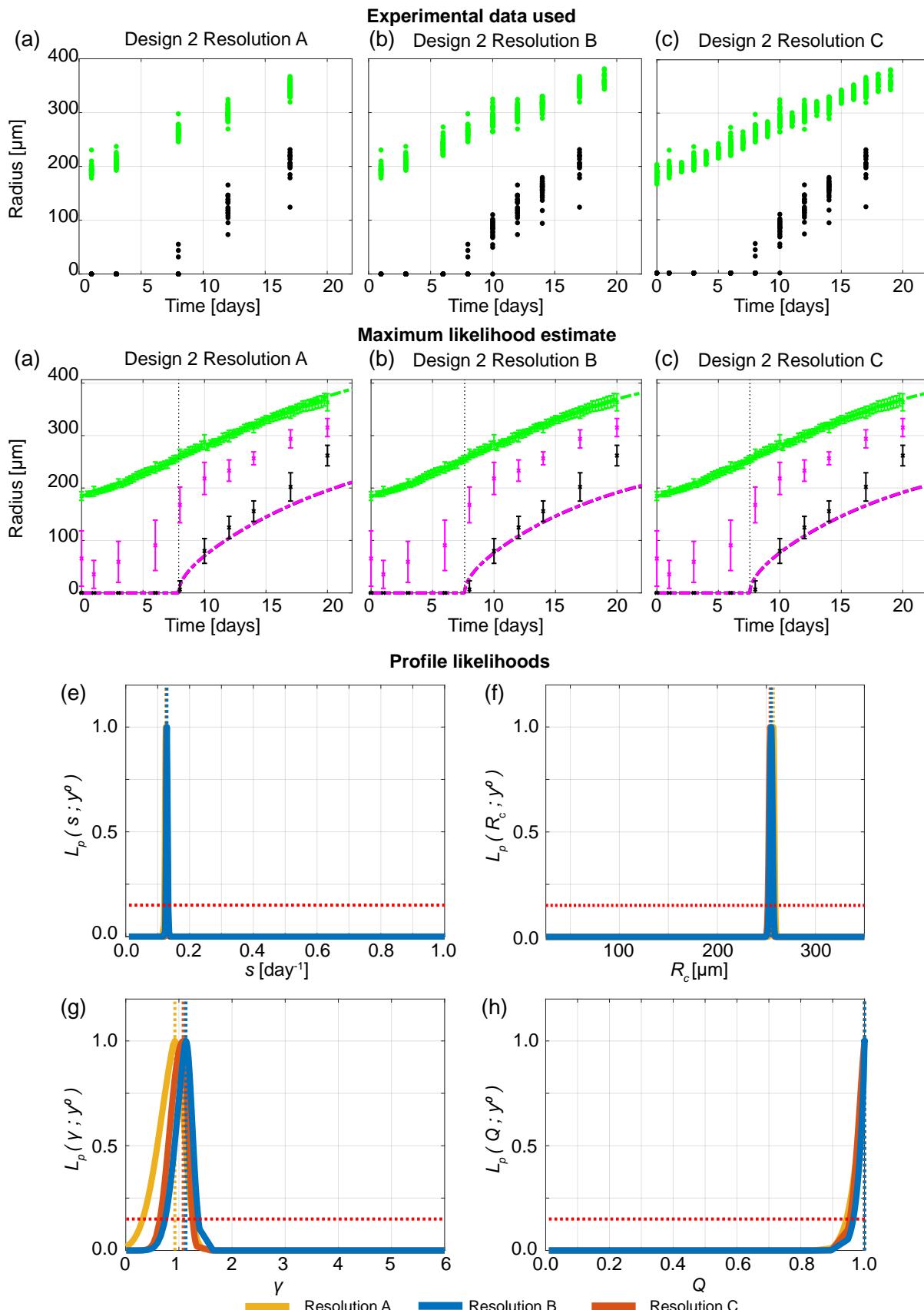
### 6A.4.2 Measurement times and experimental duration

Figure 6.1 shows that varying the temporal resolution in Design 1 is not sufficient to predict necrotic and inhibited radii. Here, in Figures 6A.18 and 6A.19, we show that varying the temporal resolution using Designs 2 and 3, respectively, gives consistent results across temporal resolutions A, B, and C.

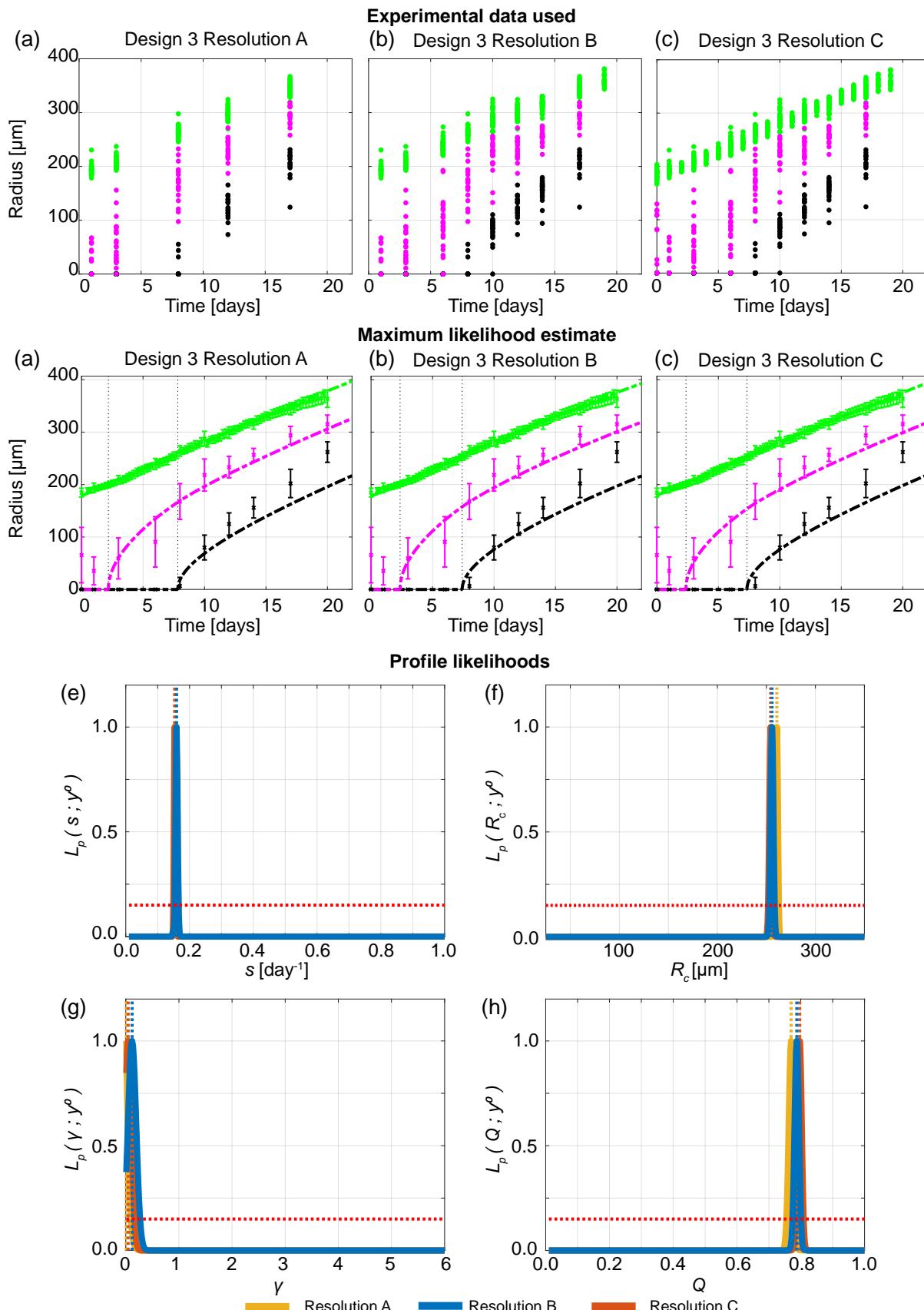
Next we consider four additional experimental designs that use different temporal measurements

- Temporal Resolution D: the first 4 days (Day 1, 2, 3),
- Temporal Resolution E: the first 10 days (Day 1, 2, 3, 4, 5, 6, 7, 8, 9, 10),
- Temporal Resolution F: the last 10 days (Day 10, 11, 12, 13, 14, 15, 16, 17, 18, 19),
- Temporal Resolution G: the last 4 days (Day 16, 17, 18, 19),

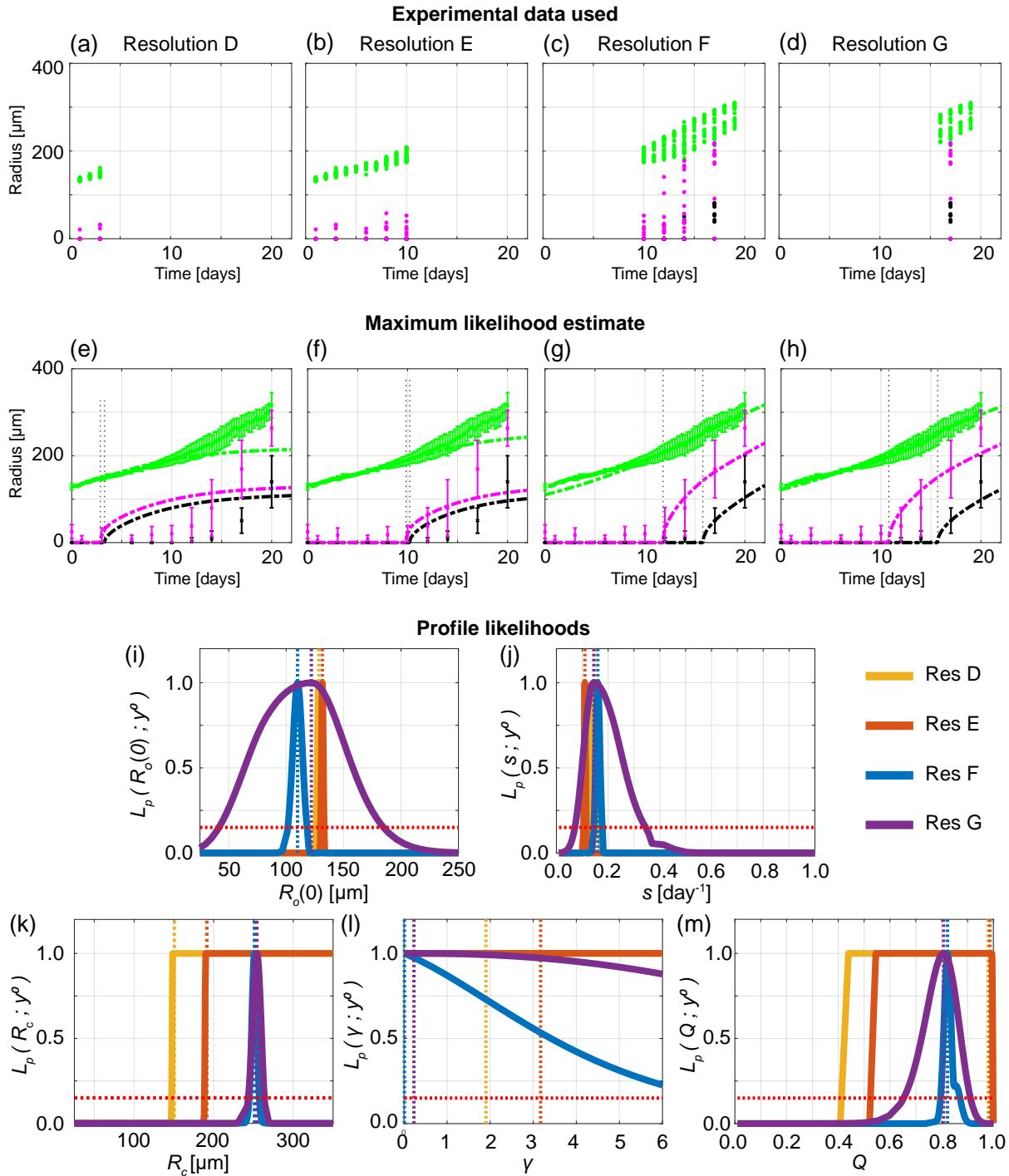
In Figures 6A.20 and 6A.21, we present results for the WM793b cell line for spheroids formed with 1250 and 5000 cells, respectively. These results show that using Temporal Resolution D is not sufficient to predict late time behaviour (Figure 6A.20e, Figure 6A.21e) and Temporal Resolution E can also not be sufficient to predict late time behaviour (Figure 6A.20f). Similarly, using late time experimental measurements, as in Temporal Resolutions F and G, is insufficient to determine tumour spheroid structure at early times (Figure 6A.21g-h).



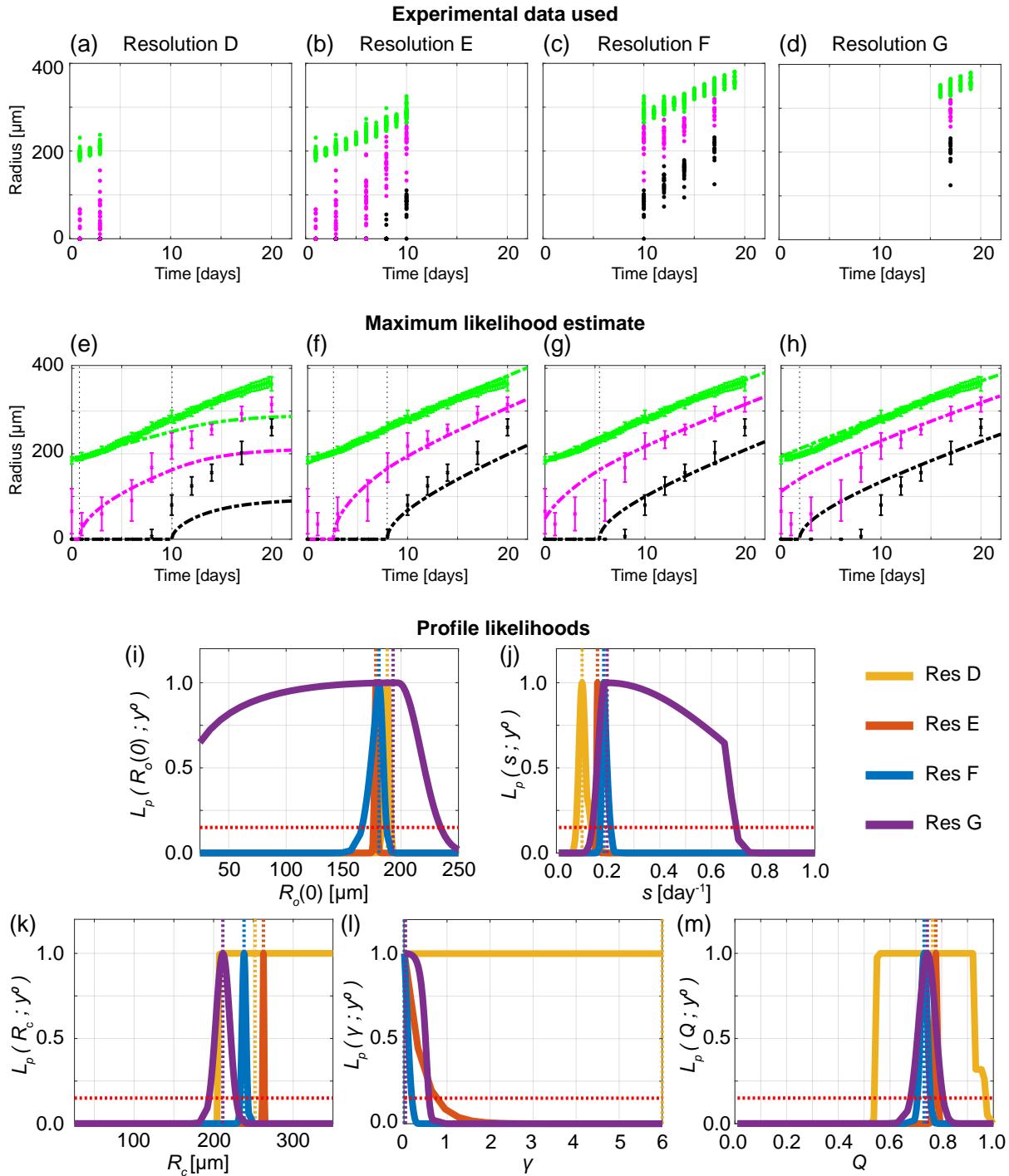
**Figure 6A.18:** Increasing times when outer and necrotic radius is measured gives consistent information. (a)-(c) Experimental data used in Design 2 with Temporal Resolutions A, B, and C. Profile likelihoods for (e)  $s$ , (f)  $R_c$ , (g)  $\gamma$ , (h)  $Q$ . Yellow, blue, and orange lines in (e)-(h) represent profile likelihoods from Design 2 with Temporal Resolutions A, B, and C, respectively. (i)-(k) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set for Design 2 with Temporal Resolutions A, B, and C, where error bars show standard deviation.



**Figure 6A.19:** Increasing times when outer, necrotic and inhibited radius and necrotic is measured gives consistent information. (a)-(c) Experimental data used in Design 3 with Temporal Resolutions A, B, and C. Profile likelihoods for (e)  $s$ , (f)  $R_c$ , (g)  $\gamma$ , (h)  $Q$ . Yellow, blue, and orange lines in (e)-(h) represent profile likelihoods from Design 3 with Temporal Resolutions A, B, and C, respectively. (i)-(k) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set for Design 3 with Temporal Resolutions A, B, and C, where error bars show standard deviation.



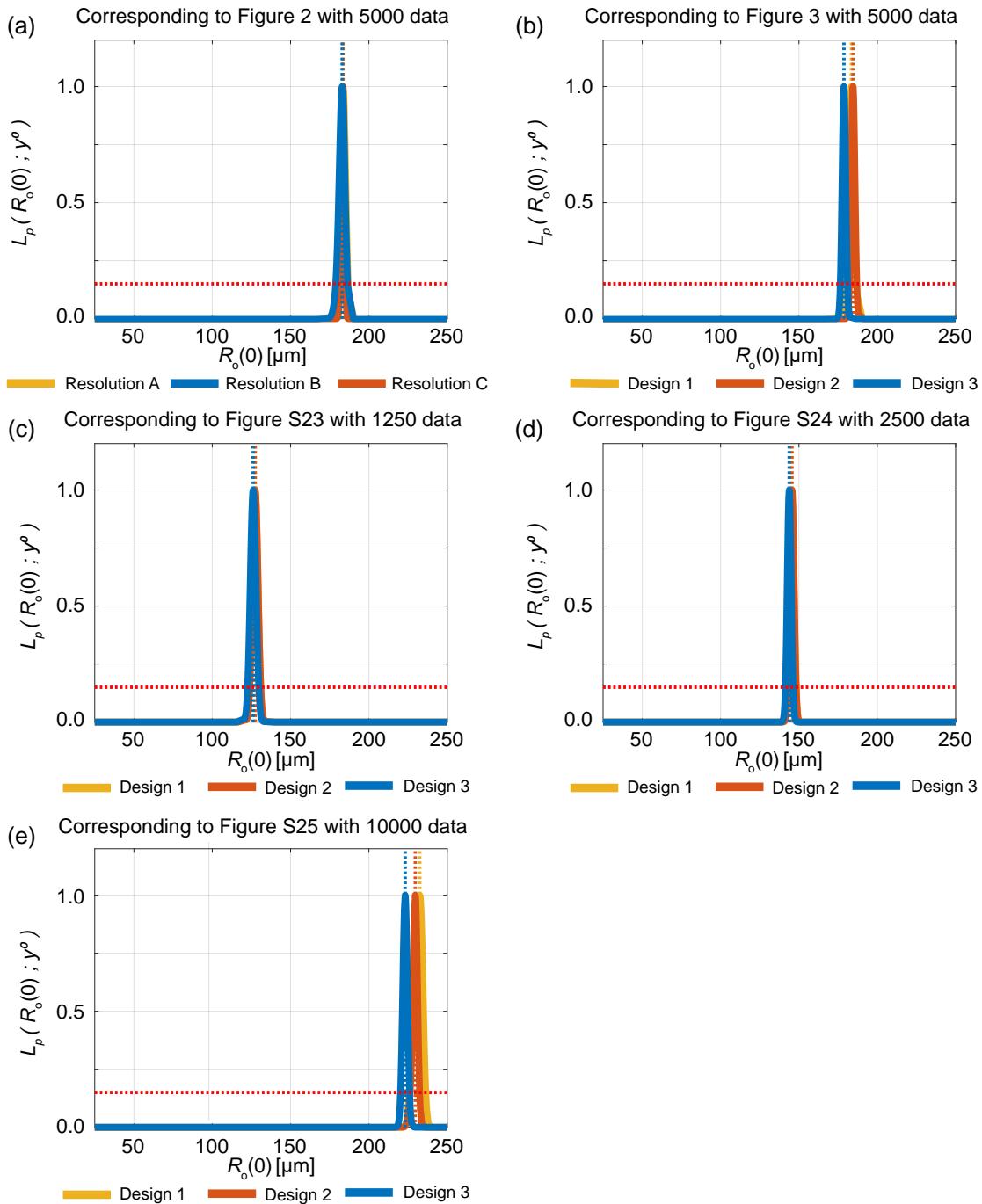
**Figure 6A.20:** Design 3 with temporal Resolutions D, E, F, and G for WM793b tumour spheroids formed with 1250 cells per spheroid. (a)-(d) Experimental data used in Design 3 Temporal Resolutions D, E, F, and G, respectively. (e)-(h) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set for Design 3 Temporal Resolutions D, E, F, and G, respectively, where error bars show standard deviation. Profile likelihoods for (i)  $R_o(0)$ , (j)  $s$ , (k)  $R_c$ , (l)  $\gamma$ , (m)  $Q$ . Yellow, orange, blue, and purple lines in (e)-(h) represent profile likelihoods from Designs 3 Temporal Resolutions D, E, F, and G, respectively.



**Figure 6A.21:** Design 3 with temporal Resolutions D, E, F, and G for WM793b tumour spheroids formed with 5000 cells per spheroid. (a)-(d) Experimental data used in Design 3 Temporal Resolutions D, E, F, and G, respectively. (e)-(h) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set for Design 3 Temporal Resolutions D, E, F, and G, respectively, where error bars show standard deviation. Profile likelihoods for (i)  $s$ , (j)  $R_o(0)$ , (k)  $R_c$ , (l)  $\gamma$ , (m)  $Q$ . Yellow, orange, blue, and purple lines in (e)-(h) represent profile likelihoods from Designs 3 Temporal Resolutions D, E, F, and G, respectively.

### 6A.4.3 Profile likelihoods for $R_o(0)$

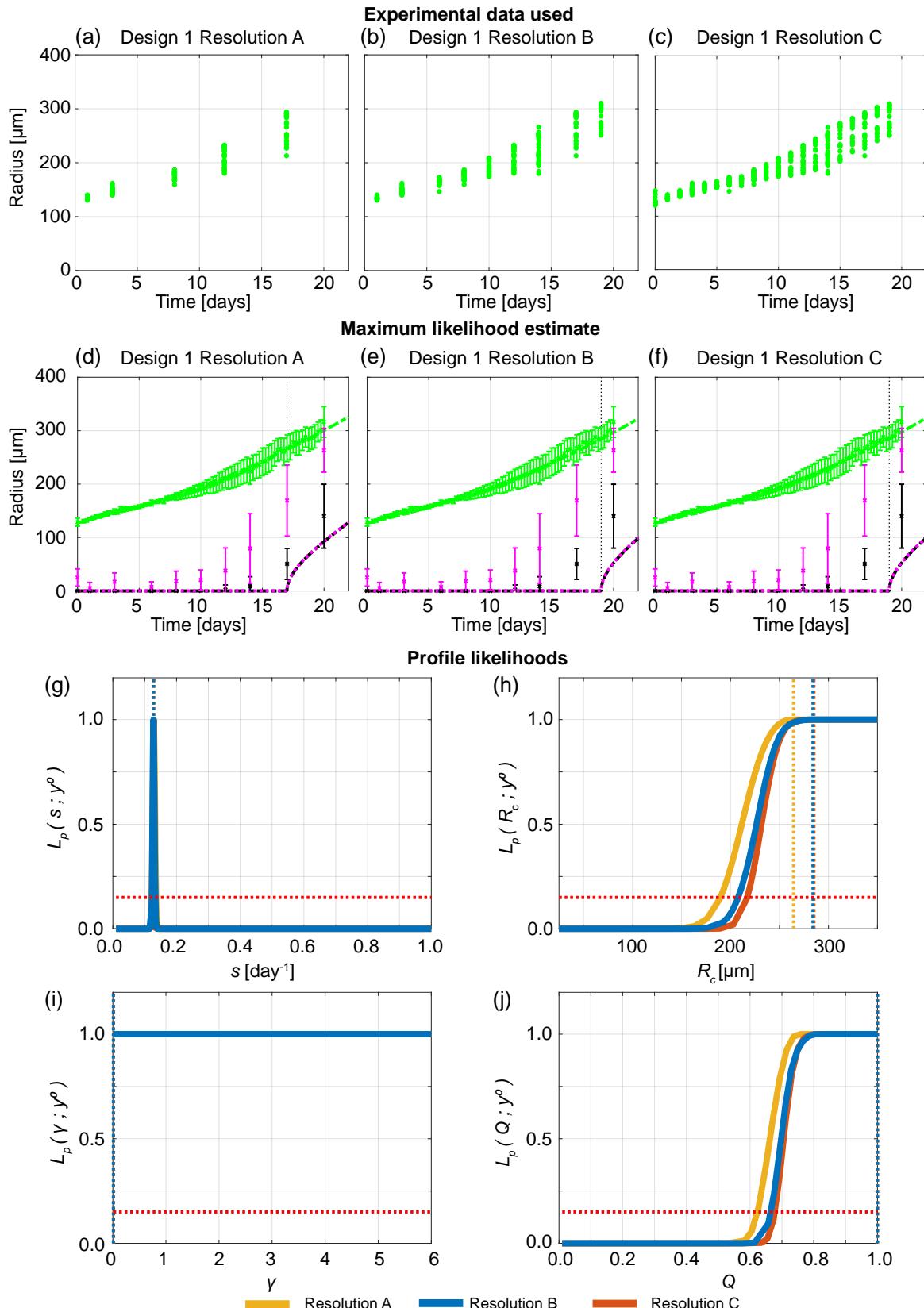
To perform statistical identifiability analysis we treat the initial outer radius,  $R_o(0)$ , as a parameter. Here, in Figure 6A.22 we show that profile likelihoods for  $R_o(0)$  are consistent across temporal resolutions and experimental designs.



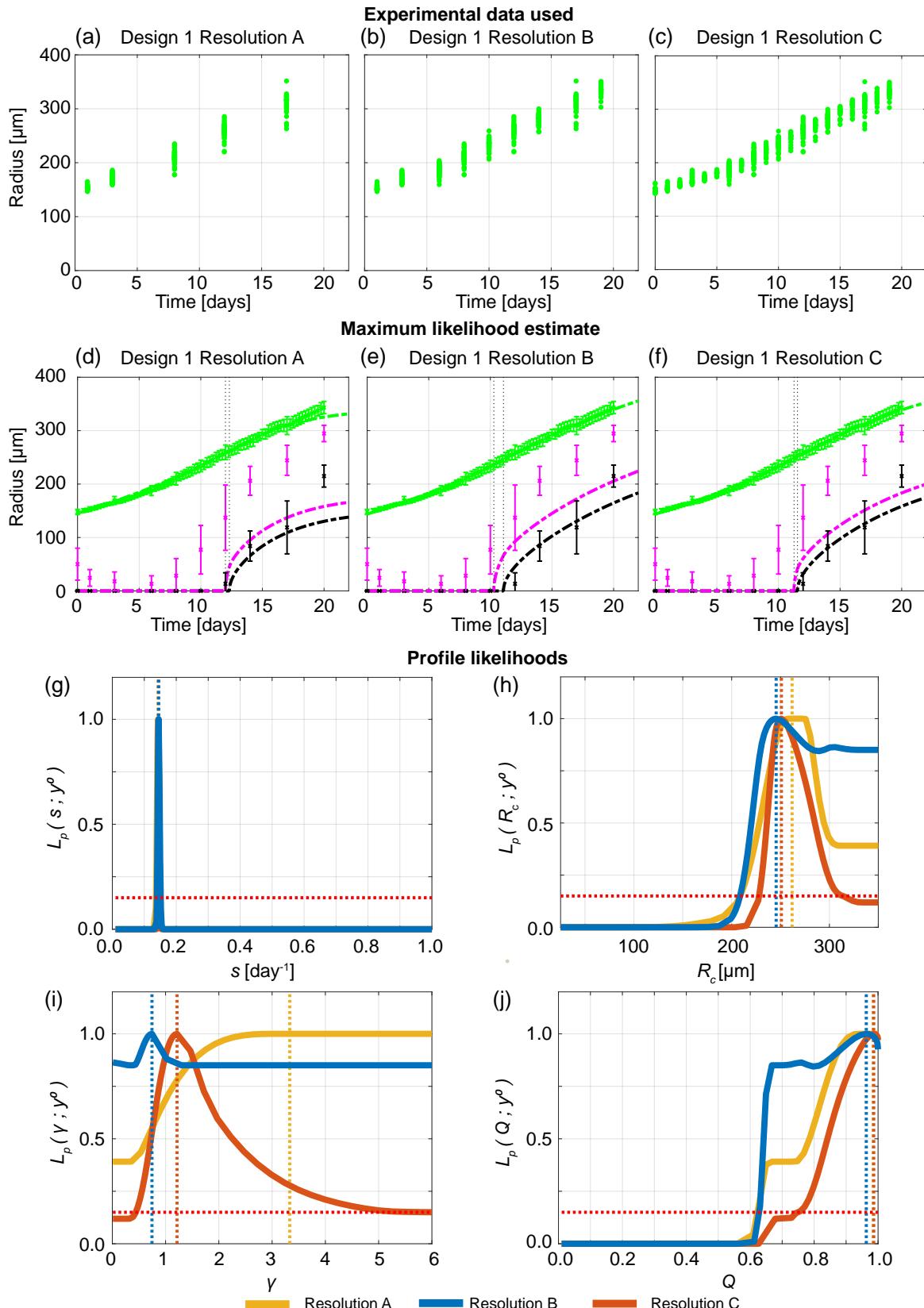
**Figure 6A.22:** Profile likelihoods for  $R_o(0)$

#### **6A.4.4 Outer radius measurements are not sufficient to predict inhibited and necrotic radii**

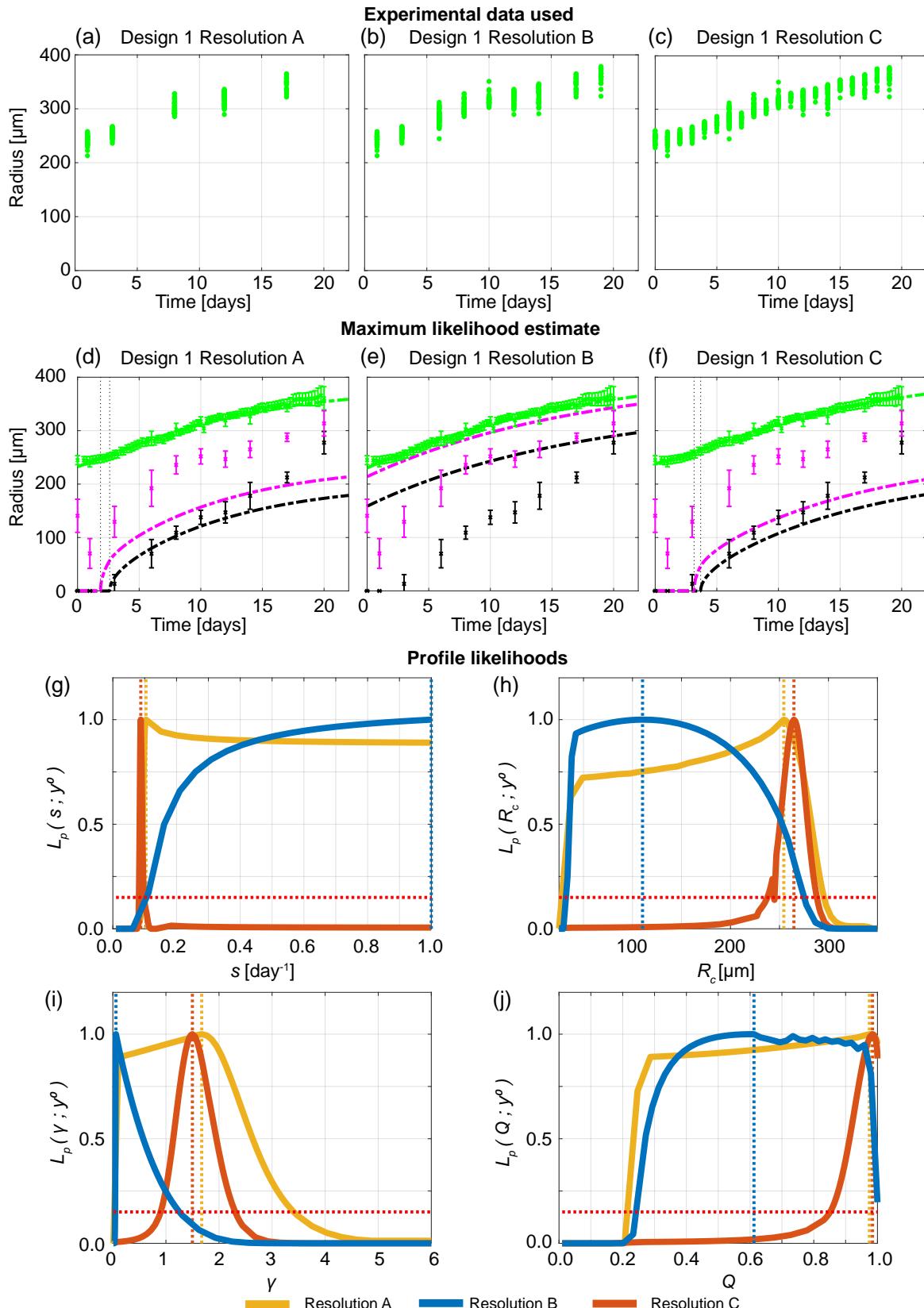
In Figure 6.2 we compare Design 1 with Temporal Resolutions A, B, and C for the WM793b cell line formed with 5000 cells. Here, in Figures 6A.23, 6A.24, and 6A.25, we compare Design 1 with Temporal Resolutions A, B, and C for the WM793b spheroids formed with 1250, 2500, and 10000 cells, respectively. These results show that Design 1 is not a reliable design and that outer radius measurements are not sufficient to predict inhibited and necrotic radii.



**Figure 6A.23:** Increasing the temporal resolution when the outer radius is measured does not provide accurate information on internal structure. (a)-(c) Experimental data used in Design 1 with temporal resolutions A, B, and C. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set, where error bars show standard deviation. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, blue, orange lines in (g)-(j) represent profile likelihoods from Design 1 with temporal resolutions A, B, and C, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. Results shown for WM793b spheroids formed with 1250 cells per spheroid.



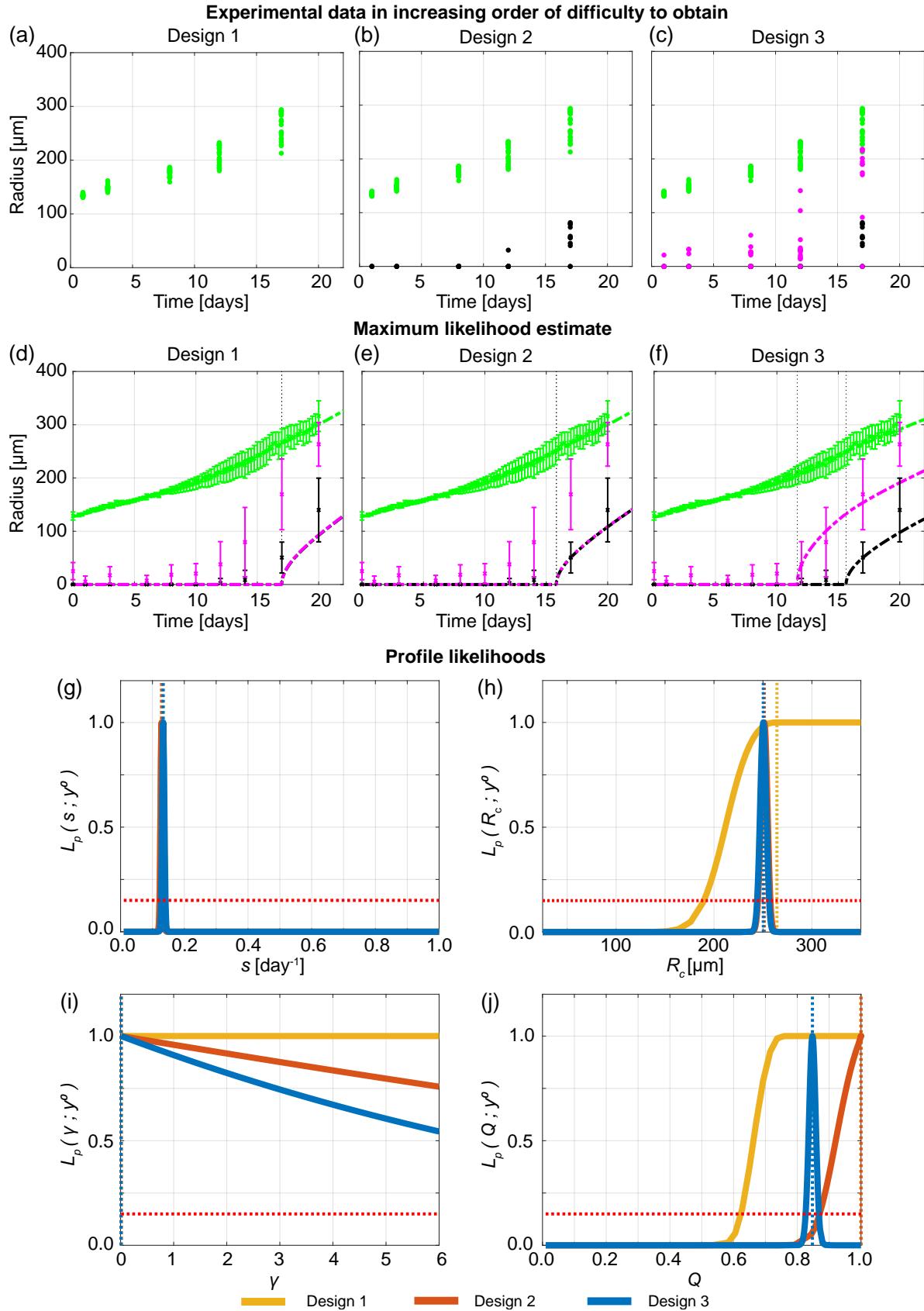
**Figure 6A.24:** Increasing the temporal resolution when the outer radius is measured does not provide accurate information on internal structure. (a)-(c) Experimental data used in Design 1 with temporal resolutions A, B, and C. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set, where error bars show standard deviation. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, blue, orange lines in (g)-(j) represent profile likelihoods from Design 1 with temporal resolutions A, B, and C, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. Results shown for WM793b spheroids formed with 2500 cells per spheroid.



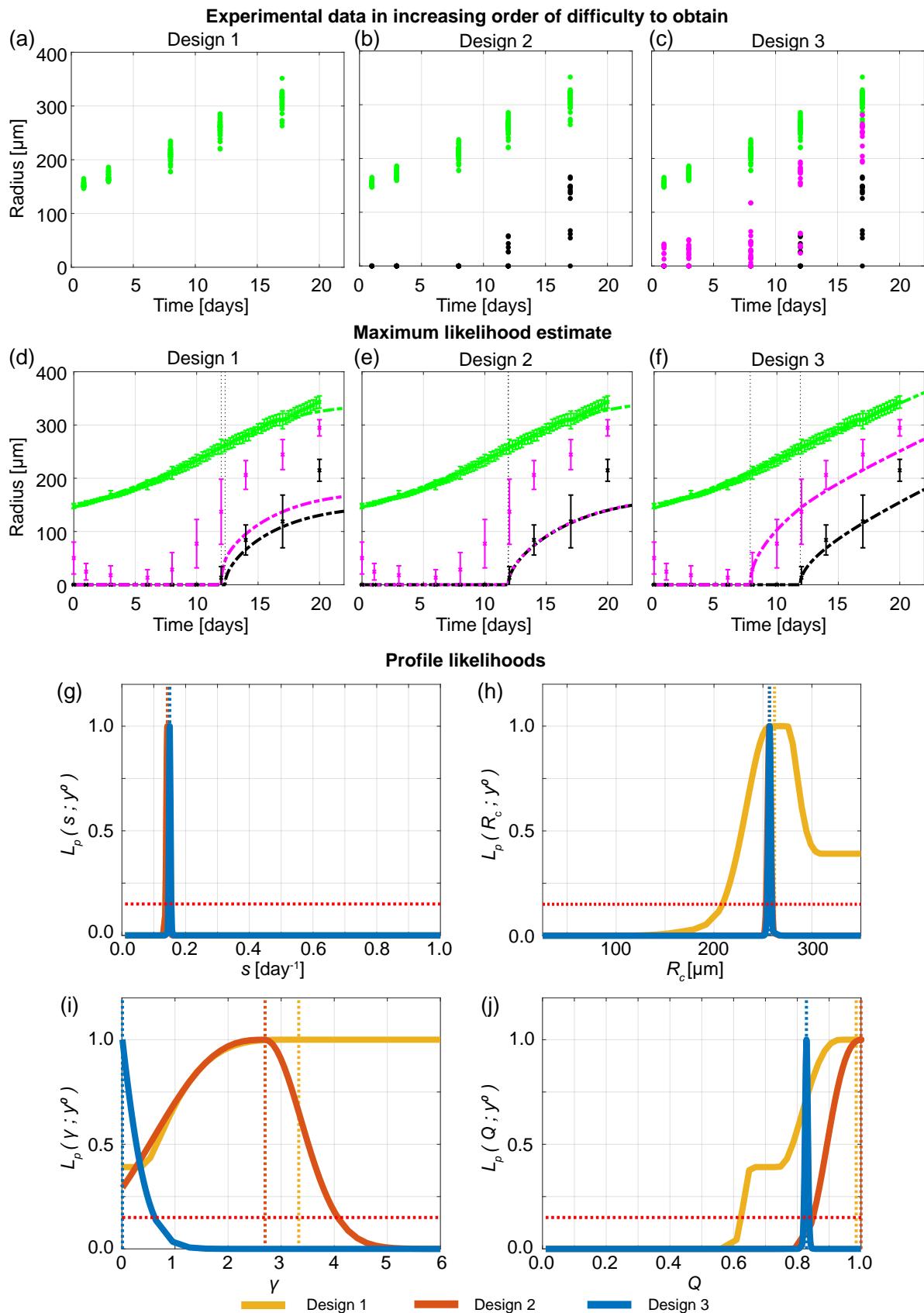
**Figure 6A.25:** Increasing the temporal resolution when the outer radius is measured does not provide accurate information on internal structure. (a)-(c) Experimental data used in Design 1 with temporal resolutions A, B, and C. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set, where error bars show standard deviation. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, blue, orange lines in (g)-(j) represent profile likelihoods from Design 1 with temporal resolutions A, B, and C, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. Results shown for WM793b spheroids formed with 10000 cells per spheroid.

#### 6A.4.5 Cell cycle data are informative

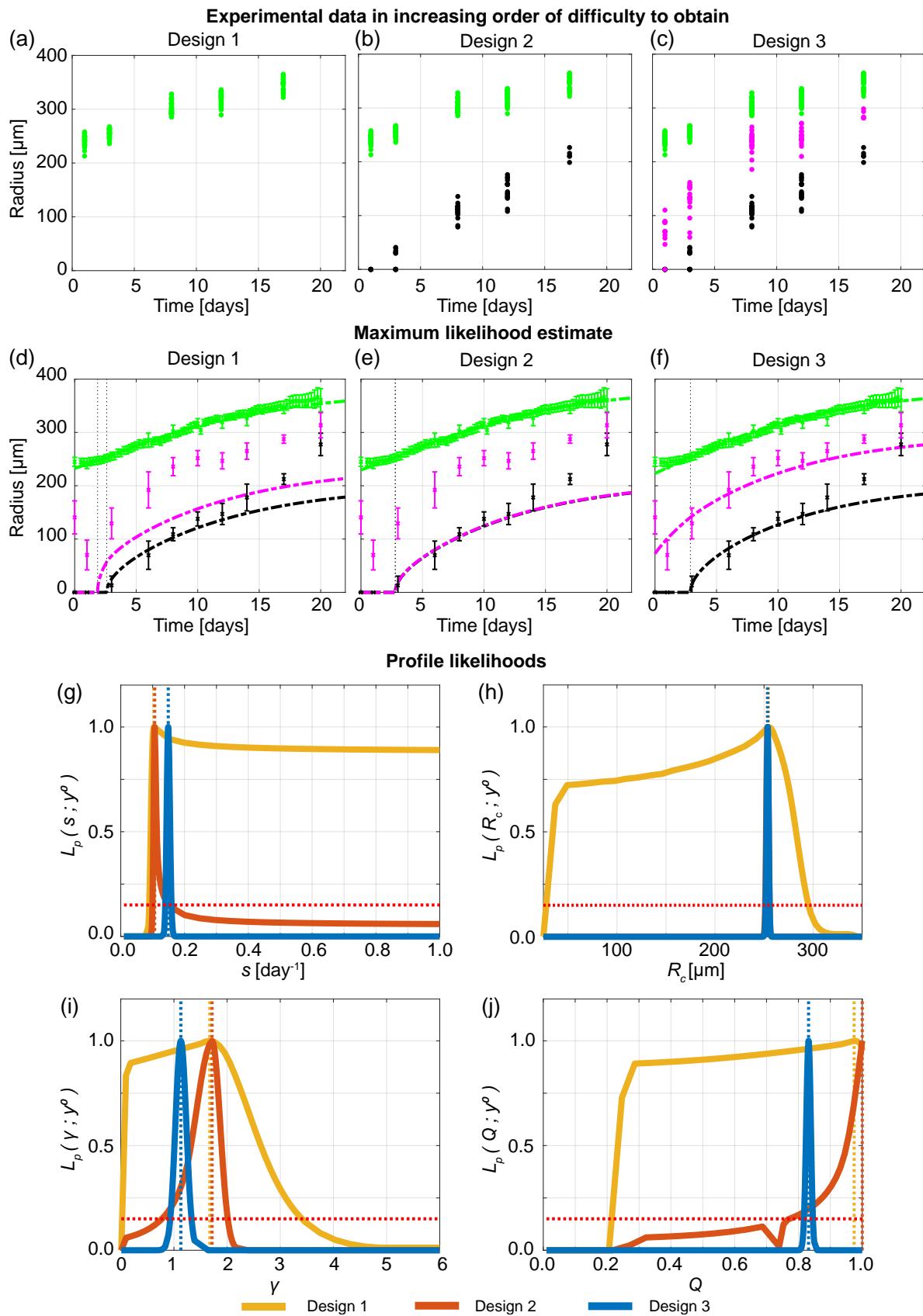
In Figure 6.3 we compare Designs 1, 2, and 3 for the WM793b cell line formed with 5000 cells. Here, in Figures 6A.26, 6A.27, and 6A.28, we compare Designs 1, 2, and 3 for the WM793b spheroids formed with 1250, 2500, and 10000 cells, respectively. These results also show that Design 3 provides most insight.



**Figure 6A.26:** Measuring the necrotic and inhibited radius provides valuable information. (a)-(c) Experimental data used in Designs 1, 2 and 3. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set for Designs 1, 2 and 3, where error bars show standard deviation. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, orange, blue lines in (g)-(j) represent profile likelihoods from Designs 1, 2, and 3, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. Results shown for WM793b spheroids formed with 1250 cells per spheroid.



**Figure 6A.27:** Measuring the necrotic and inhibited radius provides valuable information. (a)-(c) Experimental data used in Designs 1, 2 and 3. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set for Designs 1, 2 and 3, where error bars show standard deviation. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, orange, blue lines in (g)-(j) represent profile likelihoods from Designs 1, 2, and 3, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. Results shown for WM793b spheroids formed with 2500 cells per spheroid.



**Figure 6A.28:** Measuring the necrotic and inhibited radius provides valuable information. (a)-(c) Experimental data used in Designs 1, 2 and 3. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set for Designs 1, 2 and 3, where error bars show standard deviation. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, orange, blue lines in (g)-(j) represent profile likelihoods from Designs 1, 2, and 3, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. Results shown for WM793b spheroids formed with 10000 cells per spheroid.

## 6A.5 Synthetic data: WM793b

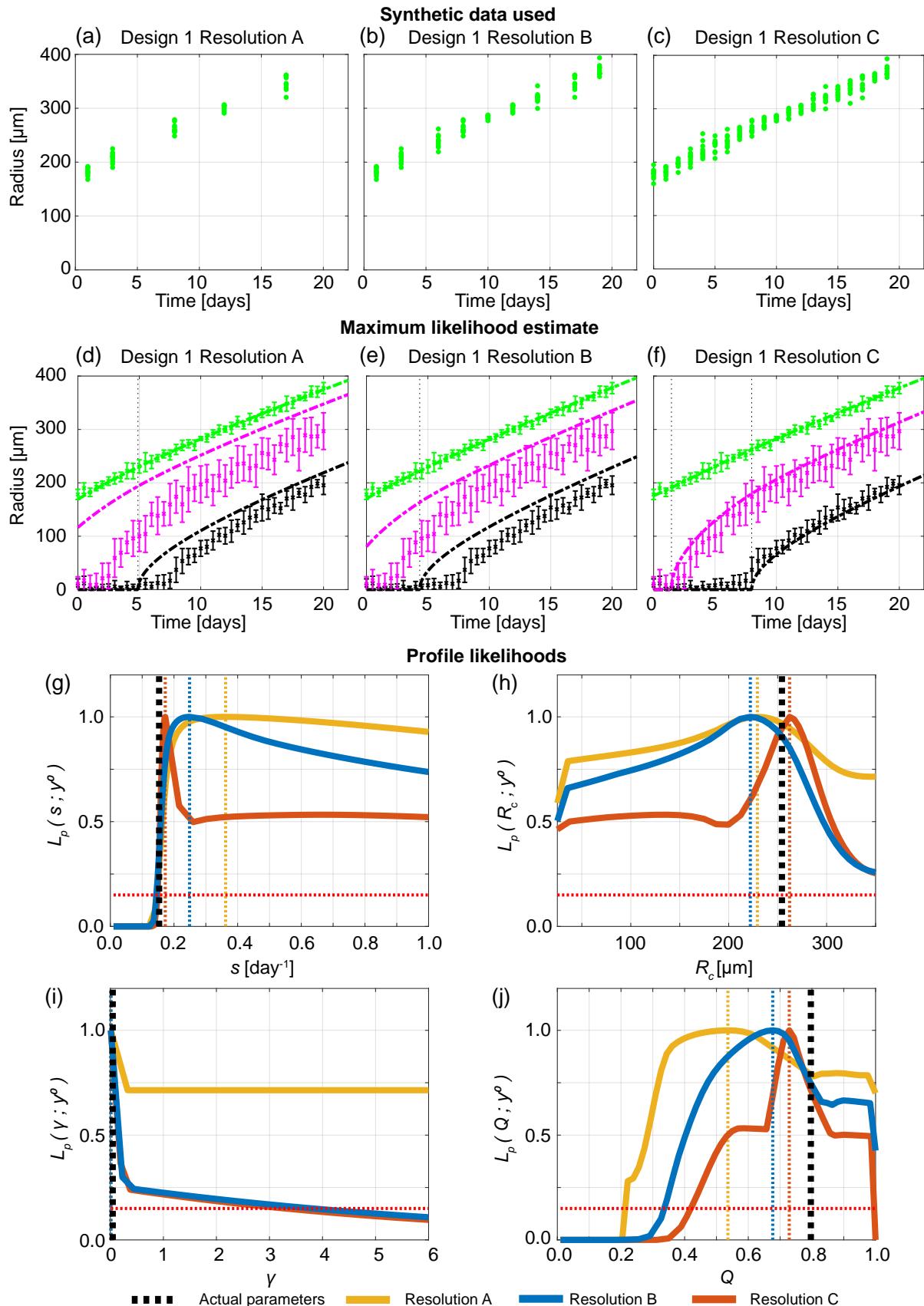
To confirm that profile likelihood analysis works as expected, we generate synthetic data from Greenspan's mathematical model using known parameters. We then explore when these known parameters are recovered using the varying experimental designs considered in the main manuscript: Design 1 with varying temporal resolutions (Supplementary Material 6A.5.1); comparing Design 1, Design 2 and Design 3 (Supplementary Material 6A.5.2), and exploring the role of initial spheroid size and also here experimental duration (Supplementary Material 6A.5.3). Since Greenspan's model may be misspecified, and may not capture all of the biological details of tumour spheroid growth, the fact that these results for synthetic data are consistent with those from experimental data enhances our confidence that key biological features are captured in Greenspan's model. Furthermore, when generating synthetic data there is additional flexibility so we also explore what may happen if we were to spend significantly more time collecting measurements (Supplementary Material 6A.5.4).

To generate synthetic data, we use the MLE from Design 3 Resolution C applied to experimental data obtained from WM793b spheroids each formed with 5000 cells:  $(R_c, s, \gamma, Q, R_o(0)) = (254.366, 0.1532, 0.045, 0.797, 179.550)$ . First, we simulate Greenspan's deterministic mathematical model with these known parameters. Next, to obtain one noisy synthetic outer radius measurement we record the outer radius from Greenspan's model generated from the known parameters at one time point. Next, we sample a normal distribution with zero mean and variance given by experimentally obtained outer radius pooled sample variance  $s_o^2 = 9.35$ . We add this sampled noise to the recorded outer radius measurement. We repeat this process to obtain additional outer radius measurements. Similarly, we repeat this process to obtain necrotic and inhibited radius measurements, using experimentally obtained pooled sample variances  $s_n^2 = 15.89$ , and  $s_i^2 = 33.12$ , respectively. We generate 10 measurements, or 48 when exploring the role of additional measurements in supplementary material 6A.5.4, of the outer radius, inhibited radius and necrotic radius every half day from day 0 to day 20.

### 6A.5.1 Outer radius measurements are not sufficient to predict inhibited and necrotic radii

Similarly to Figure 6.2, we observe in Figure 6A.29 that outer radius measurements are not sufficient to predict inhibited and necrotic radii. Simulating Greenspan's model with the MLE from Design 1 Time Resolution A (Figure 6A.29d) and with Design 1 Time Resolution B (Figure 6A.29e) shows the time evolution of the outer radius is captured but the time evolution

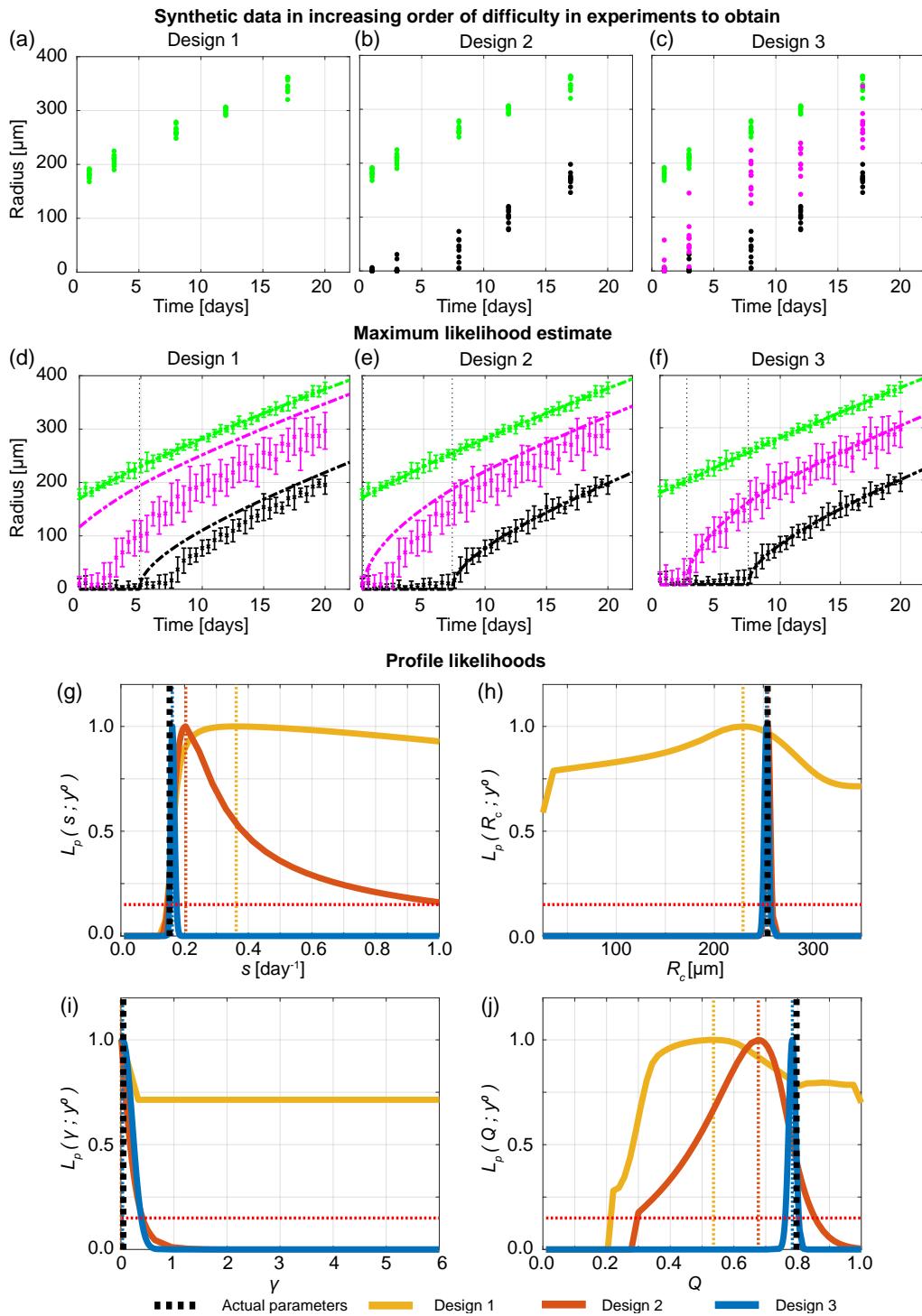
of the inhibited and necrotic region are not. However, simulating Greenspan's model with the MLE from Design 1 Time Resolution C (Figure 6A.29f) appears to capture the time evolution of the outer, inhibited, and necrotic radii. However, inspecting the profile likelihoods in Figures 6A.29g-j shows that, while the known parameters are captured, the profiles are wide suggesting that parameters are non-identifiable. This means that many parameter values give a similar match to the outer radius experimental data and these parameters do not necessarily agree with the inhibited and necrotic radii measurements.



**Figure 6A.29:** Synthetic data shows that outer radius measurements are not sufficient to predict inhibited and necrotic radii. (a)-(c) Synthetic data used in Designs 1 with Temporal Resolutions A, B, and C. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full synthetic data set. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, blue, and orange lines in (g)-(j) represent profile likelihoods from Design 1 with Temporal Resolutions A, B, and C, respectively. Black dashed lines in (g)-(j) show known parameters used to generate the synthetic data.

### 6A.5.2 Cell cycle data are informative

Similarly to Figure 6.3, Design 3 provides most insight and best captures the known parameter values used to generate the synthetic data (Figure 6A.30).



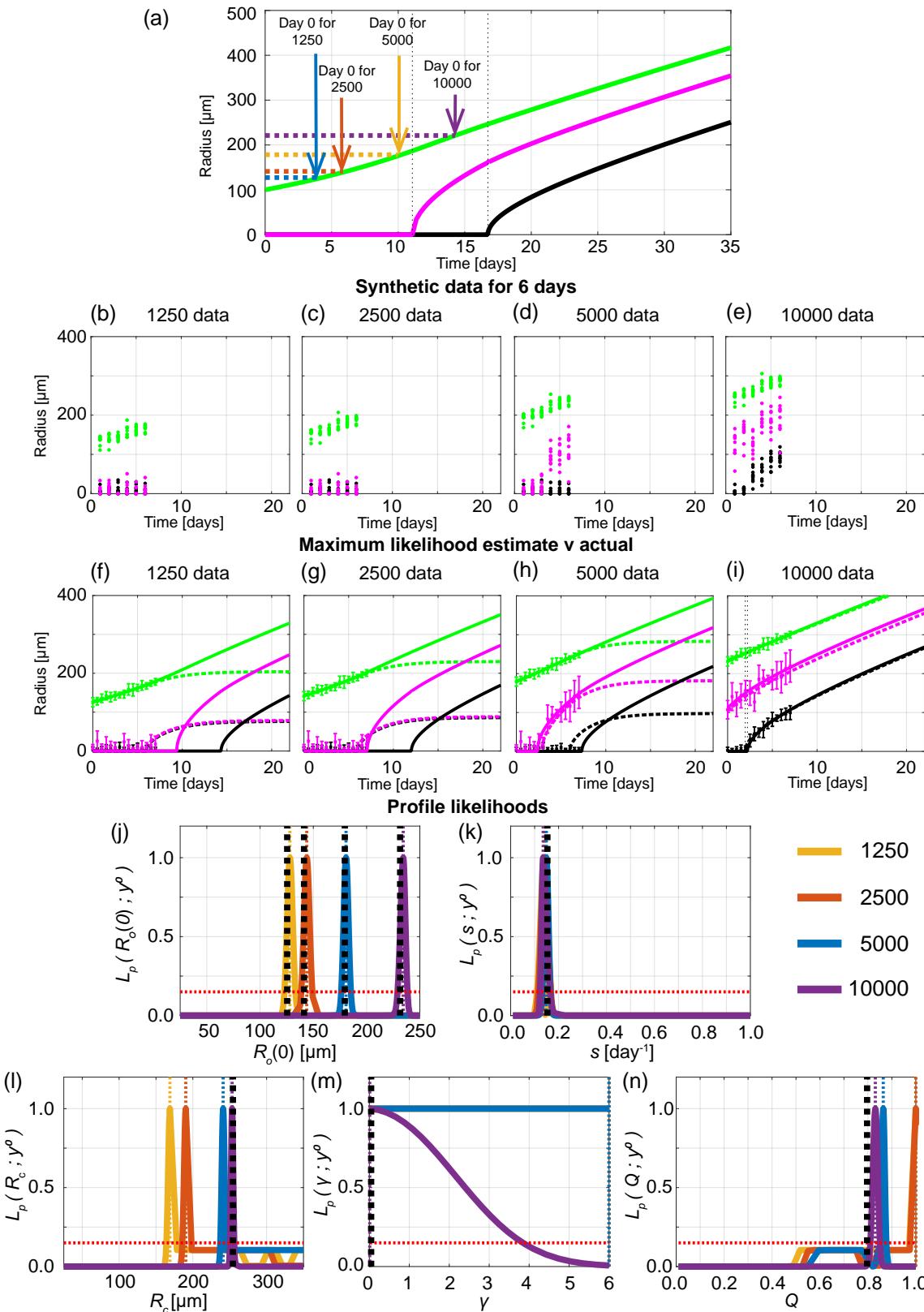
**Figure 6A.30:** Synthetic data shows that measuring the necrotic and inhibited radius provides valuable information. (a)-(c) Synthetic data used in Designs 1, 2 and 3. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full synthetic data set for Designs 1, 2 and 3. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, orange, blue lines in (g)-(j) represent profile likelihoods from Designs 1 low temporal resolution, 2, and 3, respectively. Black dashed lines in (e)-(h) show known parameters used to generate the synthetic data.

### 6A.5.3 Role of initial spheroid size and experiment duration

In Greenspan's model a change in the initial radius,  $R_o(0)$ , corresponds to a shift in time (Figure 6A.31a). We now consider the role of initial spheroid size and experiment duration. As before, we use the MLE from spheroids formed with 5000 cells per spheroid to generate synthetic data. To generate synthetic data for spheroids formed with 1250, 2500, and 10000 cells per spheroid we use the MLE obtained from spheroids formed with 5000 cells per spheroid and only update  $R_o(0)$ . To update  $R_o(0)$  we use the MLE from Design 3 applied to experimental data obtained from WM793b spheroids formed with 1250, 2500, and 10000 cells per spheroid, respectively.

We assume that each experiment is performed to Day 6 after formation, and use Design 3 with 10 measurements obtained on Day 1, 2, 3, 4, 5, and 6 (Figure 6A.31b-e). Note that during this experimental duration only spheroids formed with 10000 cells per spheroid form a necrotic core with the known parameters, while only spheroids formed with 5000 and 10000 cells per spheroid form an inhibited region with the known parameters. Therefore, we expect that most insight will be gained from the experiment formed with spheroids formed with 10000 cells per spheroid.

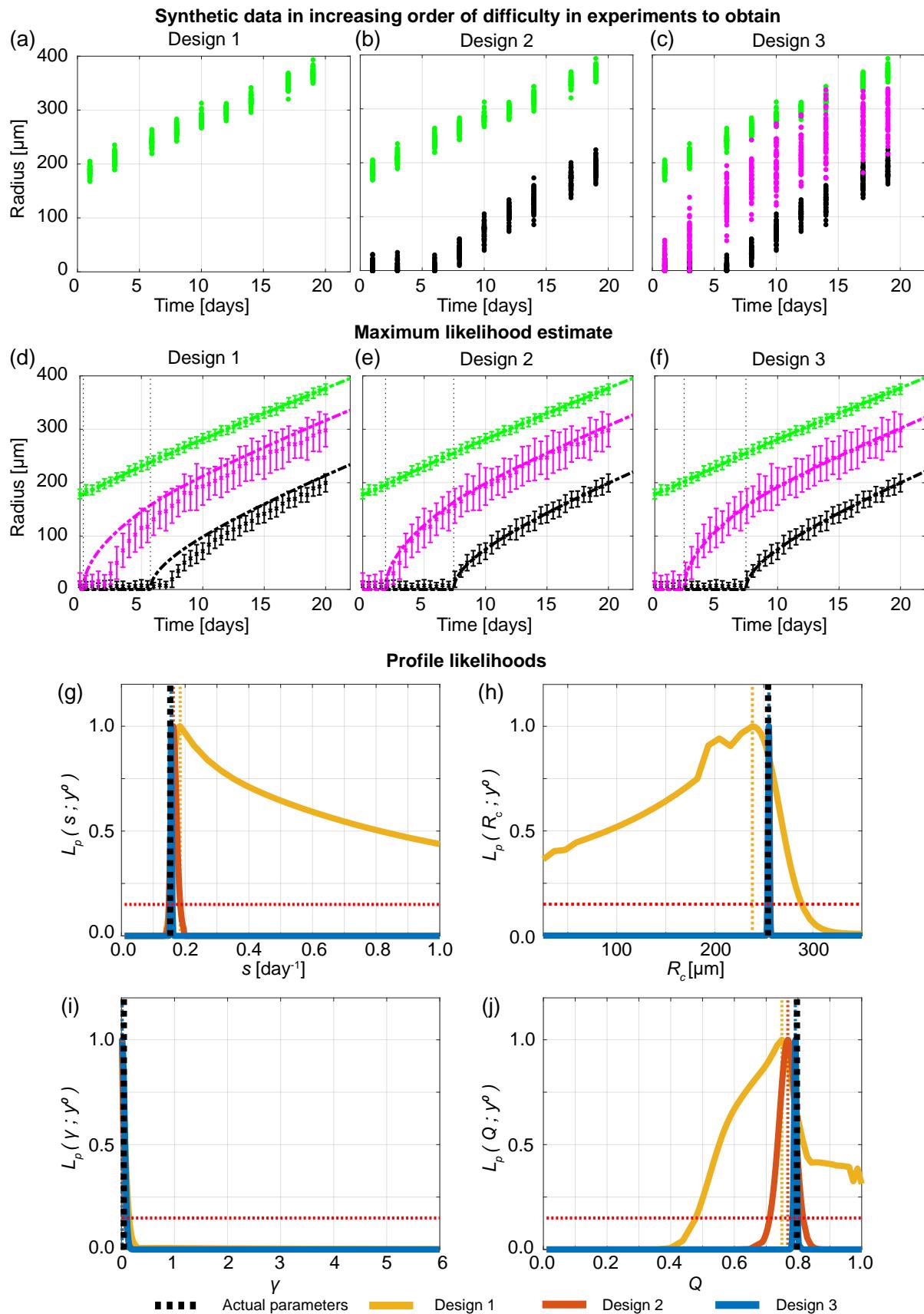
Simulating Greenspan's model with the MLE obtained from each of those data sets (dashed lines in Figures 6A.31f-i) we observe good agreement to the first six days of synthetic data for each initial spheroid size. However, simulating Greenspan's model with the MLE obtained from each of those data sets (dashed lines in Figures 6A.31f-i) and comparing to Greenspan's model simulated over 20 days with the known parameters used to generate the synthetic data (solid lines in Figures 6A.31f-i)) this is not the case. We only observe excellent agreement for the experiment with spheroids formed with 10000 cells, since this experiment has measurements in phase (iii). Profile likelihoods for the parameters also show that only the experiment performed with spheroids formed with 10000 cells accurately captures the known parameters (Figure 6A.31j-n).



**Figure 6A.31:** Synthetic data exploring role of initial spheroid size and experimental duration. (a) In Greenspan's model a change in  $R_o(0)$  corresponds to a shift in time. (b)-(e) Synthetic generated for the first six days after formation for spheroids formed with (b) 1250, (c) 2500, (d) 5000, (e) 10000 cells per spheroid. (f)-(i) Comparison of Greenspan model simulated with maximum likelihood estimate (dashed lines) compared to synthetic data for first 6 days compared to Greenspan model simulated with known parameters used to generate the synthetic data (solid lines). Profile likelihoods for (j)  $R_o(0)$ , (k)  $s$ , (l)  $R_c$ , (m)  $\gamma$ , (n)  $Q$ . Yellow, orange, blue, and purple lines in (j)-(n) represent profile likelihoods from spheroids formed with 1250, 2500, 5000, and 10000 cells per spheroid, respectively. Black dashed lines in (j)-(n) show known parameters used to generate the synthetic data.

#### 6A.5.4 Increasing number of measurements

In biological experiments it is time consuming and expensive to increase the number of measurements obtained. However, by generating synthetic data we can easily simulate additional measurements. We generate 48 measurements of the outer radius, inhibited radius, and necrotic radius every half day from Day 0 to Day 20. We choose 48 measurements since this corresponds to half of a 96-well plate and is extremely large in comparison to typical experiments. These results show that many measurements of Design 2 may provide good insight in this extreme scenario but that Design 3 still provides most insight.



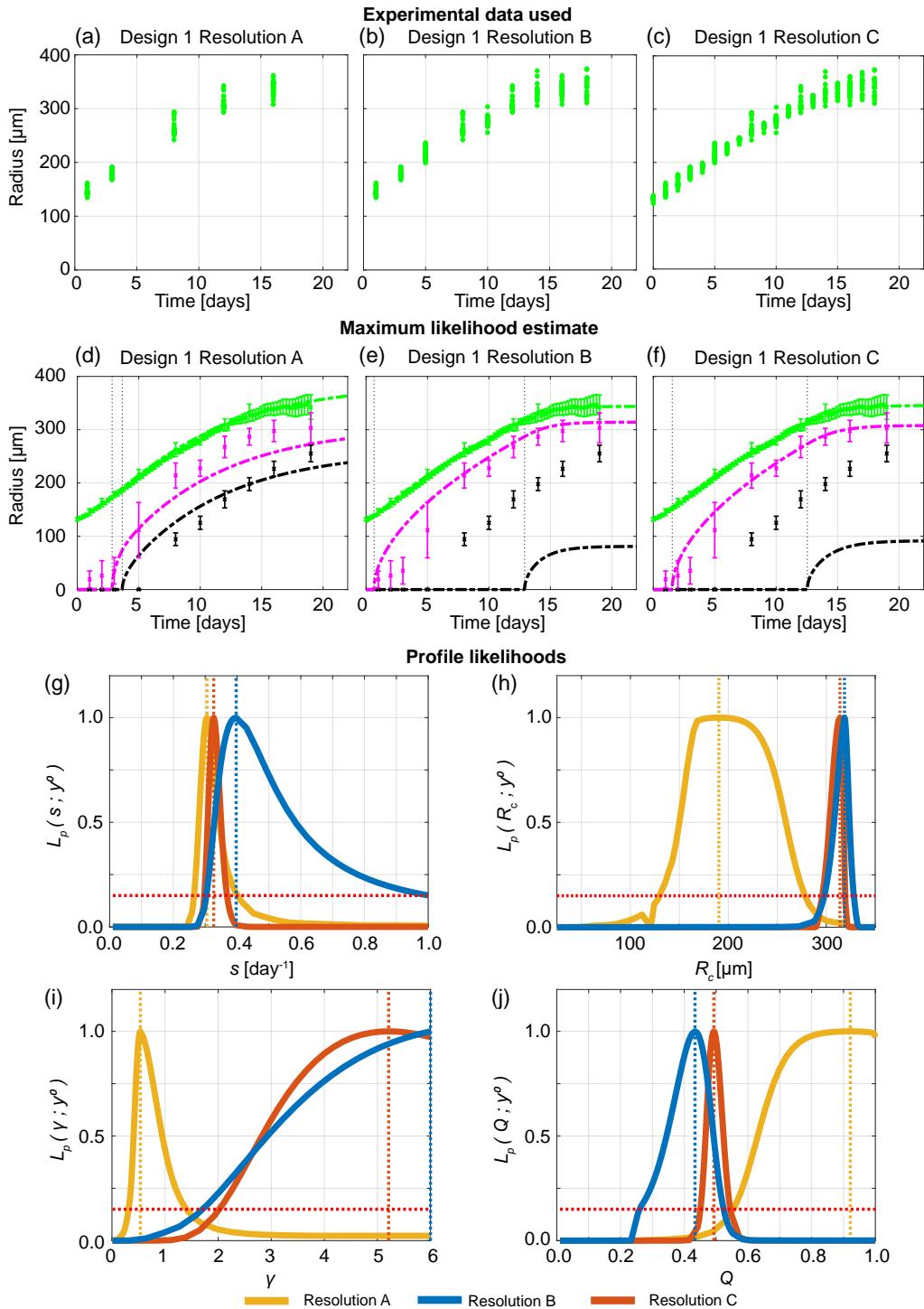
**Figure 6A.32:** Synthetic data shows that more measurements of the necrotic and inhibited radius provides valuable information. (a)-(c) Synthetic data used in Designs 1, 2 and 3. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full synthetic data set for Designs 1, 2 and 3. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, orange, blue lines in (e)-(h) represent profile likelihoods from Designs 1, 2, and 3, with low temporal resolution, respectively. Black dashed lines in (g)-(j) show known parameters used to generate the synthetic data.

## 6A.6 Parameter identifiability analysis for WM983b

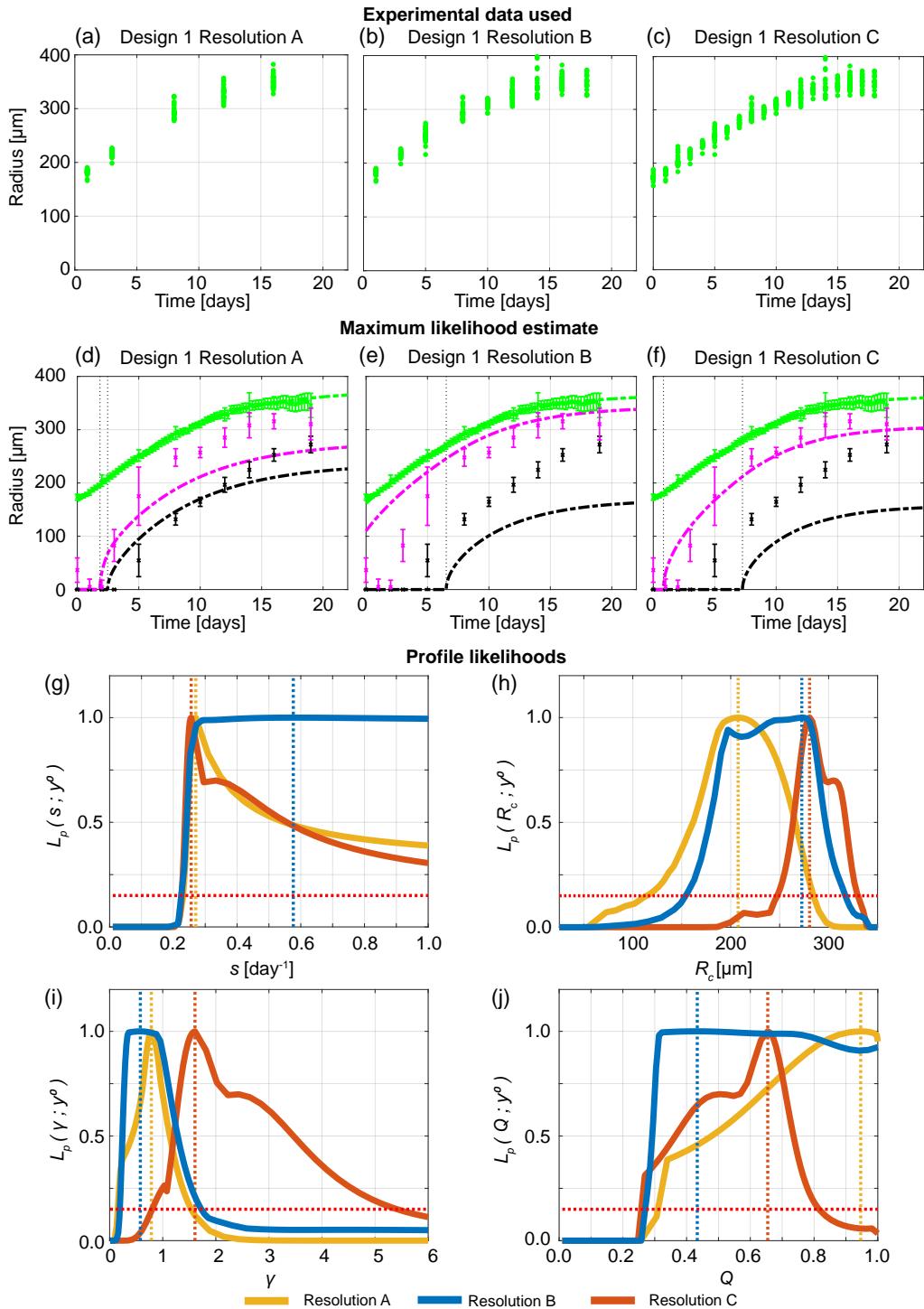
The main manuscript focuses on results for the human melanoma WM793b cell line. Here, we present the corresponding results for the human melanoma WM983b spheroids formed with 2500, 5000, and 10000 cells.

All key observations made in reference to the WM793b cell line hold for the WM983b cell line. Specifically, in Figures 6A.33, 6A.34 and 6A.35, for spheroids formed with 2500, 5000, and 10000 cells, respectively, we show that varying the temporal resolution using only Design 1 is insufficient to determine necrotic and inhibited radii. In Figures 6A.36, 6A.37, and 6A.38 for spheroids formed with 2500, 5000, and 10000 cells, respectively, we show that Design 3 provides most insight. In Figure 6A.39 we show that information gained across experiments with different initial spheroid sizes is consistent. Minor modifications were applied to the experimental designs as the WM983b tumour spheroids form after 3 days, which is 1 day earlier than the WM793b tumour spheroids, and the experiment was terminated after 19 days so the updated temporal resolutions are for this cell line are chosen as: Resolution A using Days 1, 3, 8, 12, 16; Resolution B using Days 1, 3, 5, 8, 10, 12, 14, 16, 18; Resolution C using Days 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18; where Day 0 corresponds to the time that we determined as when spheroid formation ends and growth begins.

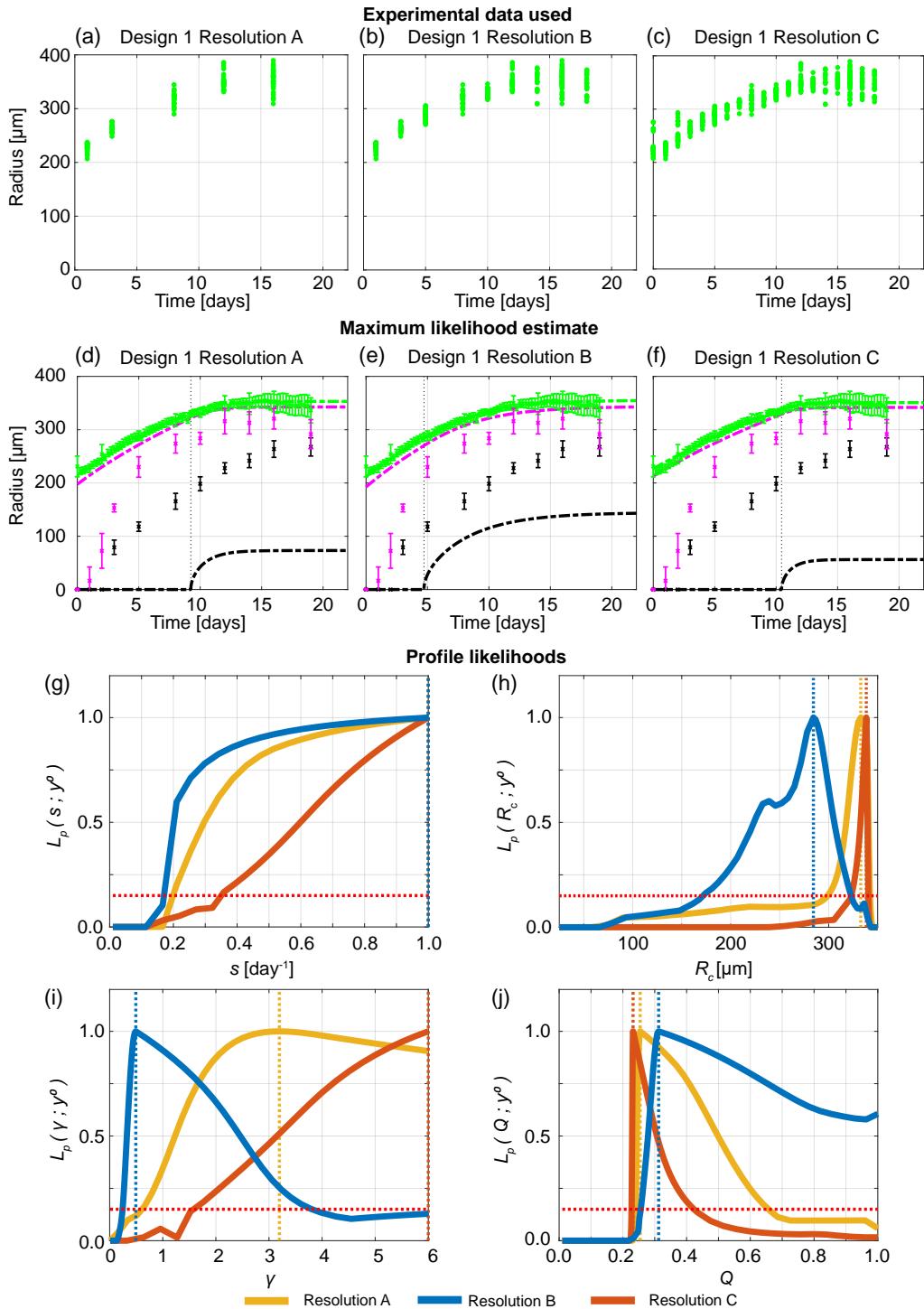
**6A.6.1 Outer radius measurements are not sufficient to predict inhibited and necrotic radii**



**Figure 6A.33:** Increasing the temporal resolution when the outer radius is measured does not provide accurate information on internal structure. (a)-(c) Experimental data used in Design 1 with temporal resolutions A, B, and C. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set, where error bars show standard deviation. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, blue, orange lines in (g)-(j) represent profile likelihoods from Design 1 with temporal resolutions A, B, and C, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. Results shown for WM983b spheroids formed with 2500 cells per spheroid.

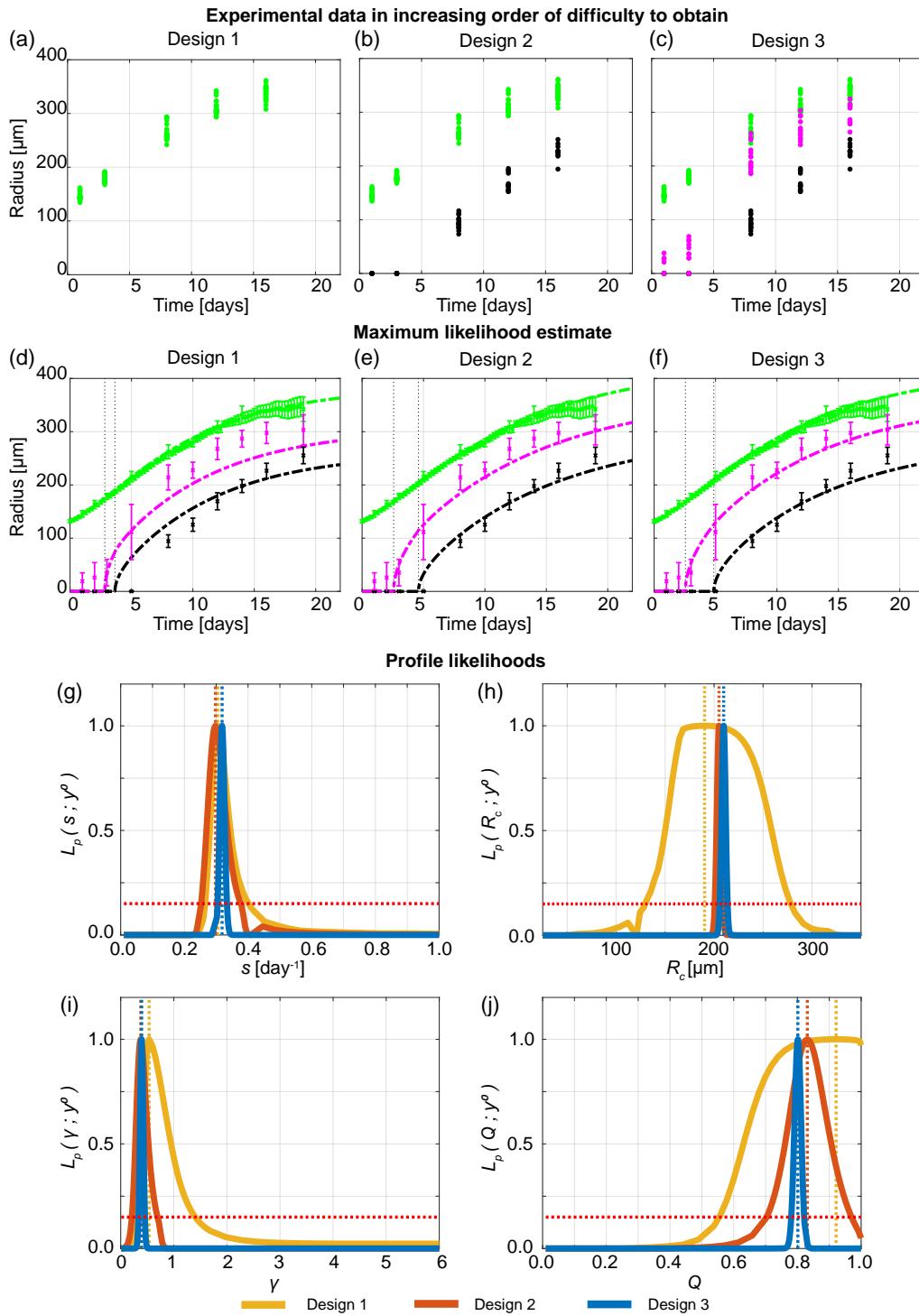


**Figure 6A.34:** Increasing the temporal resolution when the outer radius is measured does not provide accurate information on internal structure. (a)-(c) Experimental data used in Design 1 with temporal resolutions A, B, and C. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set, where error bars show standard deviation. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, blue, orange lines in (g)-(j) represent profile likelihoods from Design 1 with temporal resolutions A, B, and C, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. Results shown for WM983b spheroids formed with 5000 cells per spheroid.

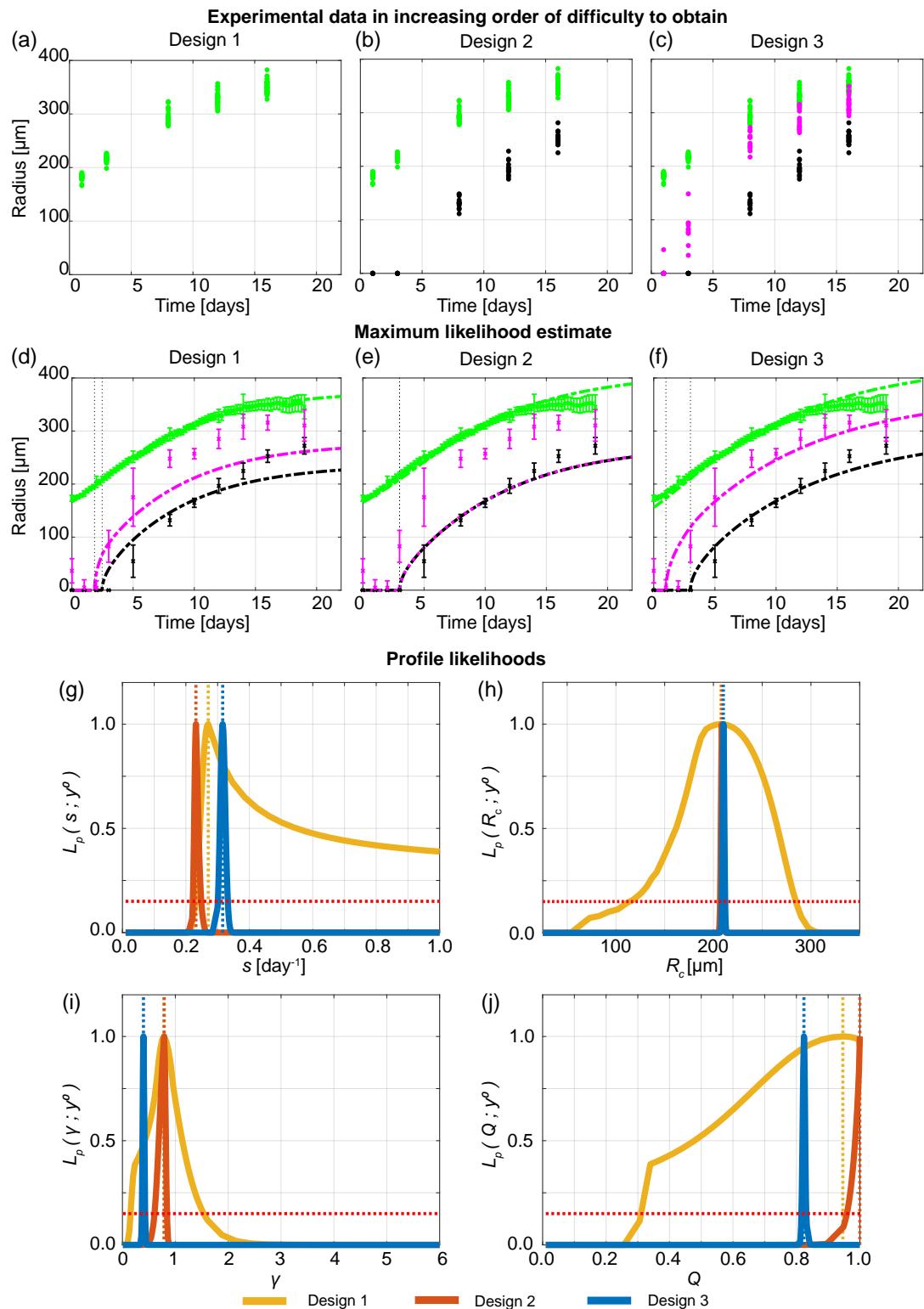


**Figure 6A.35:** Increasing the temporal resolution when the outer radius is measured does not provide accurate information on internal structure. (a)-(c) Experimental data used in Design 1 with temporal resolutions A, B, and C. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set, where error bars show standard deviation. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, blue, orange lines in (g)-(j) represent profile likelihoods from Design 1 with temporal resolutions A, B, and C, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. Results shown for WM983b spheroids formed with 10000 cells per spheroid.

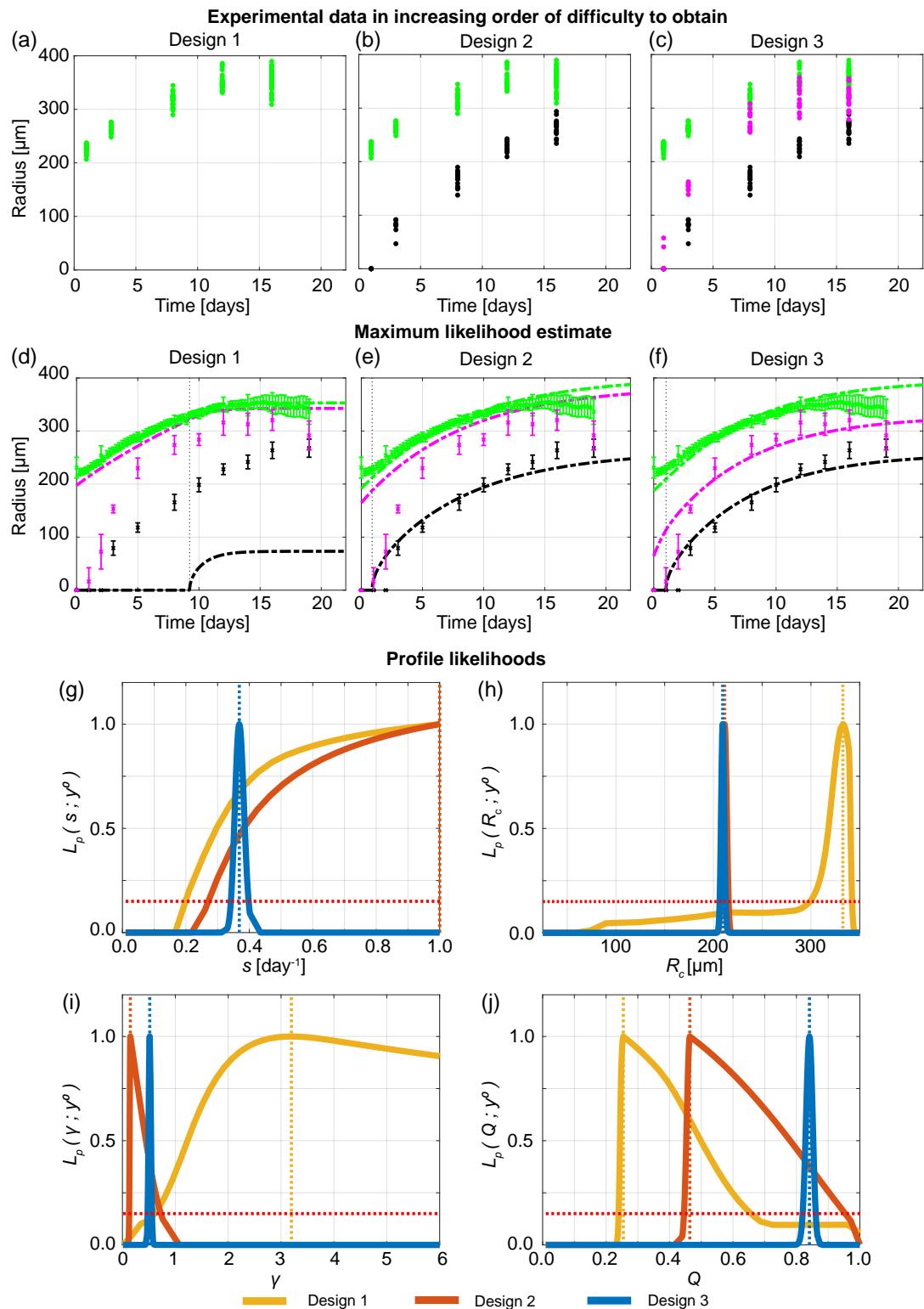
### 6A.6.2 Cell cycle data are informative



**Figure 6A.36:** Measuring the necrotic and inhibited radius provides valuable information. (a)-(c) Experimental data used in Designs 1, 2 and 3. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set for Designs 1, 2 and 3, where error bars show standard deviation. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, orange, blue lines in (g)-(j) represent profile likelihoods from Designs 1, 2, and 3, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. Results shown for WM983b spheroids formed with 2500 cells per spheroid.

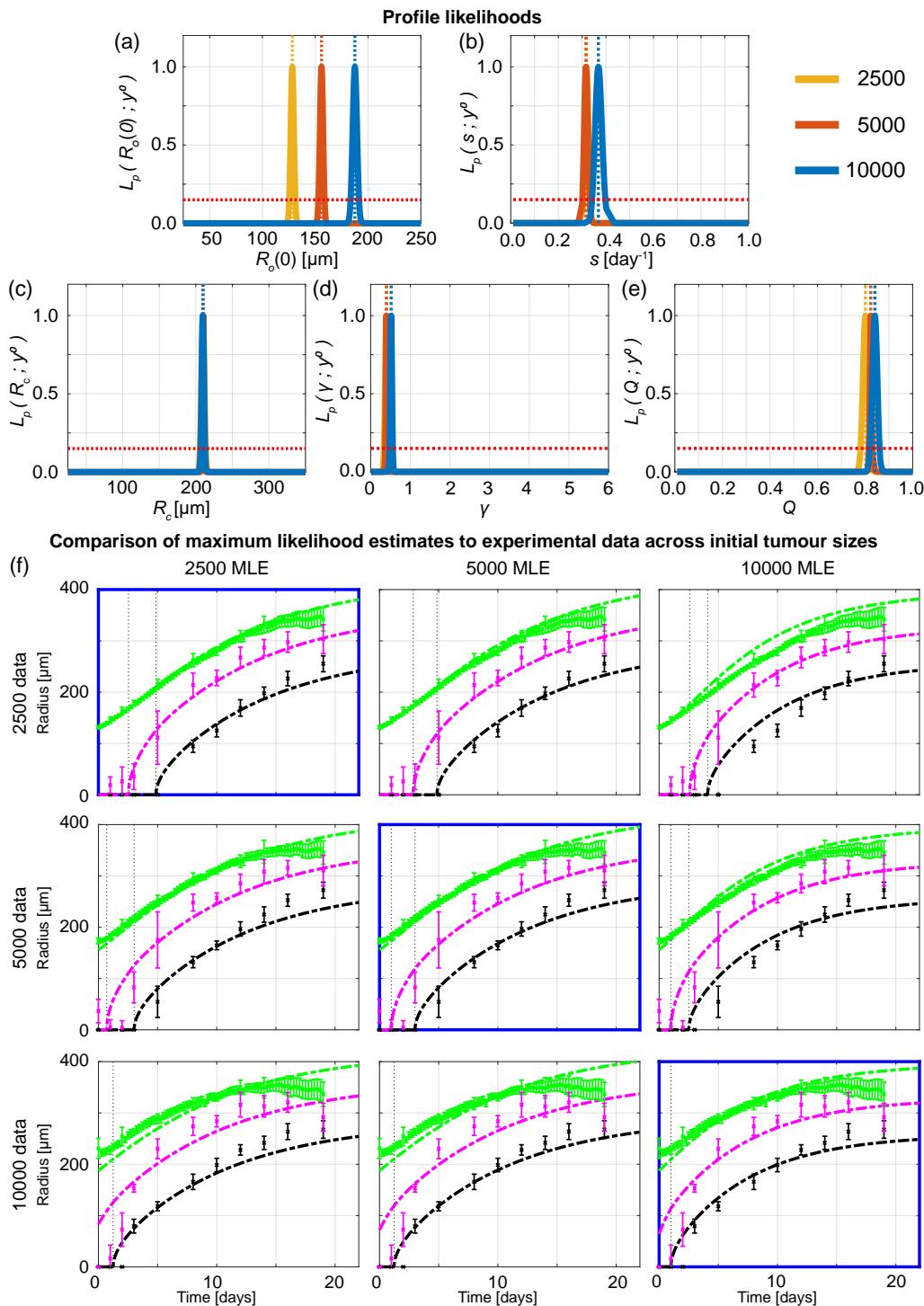


**Figure 6A.37:** Measuring the necrotic and inhibited radius provides valuable information. (a)-(c) Experimental data used in Designs 1, 2 and 3. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set for Designs 1, 2 and 3, where error bars show standard deviation. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, orange, blue lines in (g)-(j) represent profile likelihoods from Designs 1, 2, and 3, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. Results shown for WM983b spheroids formed with 5000 cells per spheroid.



**Figure 6A.38:** Measuring the necrotic and inhibited radius provides valuable information. (a)-(c) Experimental data used in Designs 1, 2 and 3. (d)-(f) Comparison of Greenspan model simulated with maximum likelihood estimate compared to full experimental data set for Designs 1, 2 and 3, where error bars show standard deviation. Profile likelihoods for (g)  $s$ , (h)  $R_c$ , (i)  $\gamma$ , (j)  $Q$ . Yellow, orange, blue lines in (g)-(j) represent profile likelihoods from Designs 1, 2, and 3, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. Results shown for WM983b spheroids formed with 10000 cells per spheroid.

### 6A.6.3 Information gained across spheroid sizes is consistent



**Figure 6A.39:** Information gained from experiments across different initial tumour spheroid sizes is mostly consistent. Profile likelihoods for (a)  $R_o$ , (b)  $s$ , (c)  $R_c$ , (d)  $\gamma$ , (e)  $Q$ . Yellow, orange, and blue lines in (a)-(e) represent profile likelihoods from tumour spheroids initially with approximately 2500, 5000, and 10000 cells per spheroid, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. (f) Comparison of Greenspan model simulated with maximum likelihood estimates compared to full experimental data sets across initial tumour spheroid size, where error bars show standard deviation. Results shown for WM983b cell line.

## 6A.7 Parameter identifiability analysis for WM164

The main manuscript focuses on results for the human melanoma WM793b cell line. Here, we present analogous results for the human melanoma WM164 spheroids formed with 1250, 2500, 5000, and 10000 cells per spheroid. These spheroids are more challenging to interpret as we will now explain.

In experiments WM164 spheroids formed after 3 days. These spheroids were larger in size than other spheroids considered in this work, with the initial radius of WM164 spheroids formed with 1250 cells per spheroid larger than and similar size to WM983b and WM793b spheroids formed with 10000 cells per spheroid, respectively. The WM164 spheroids had relatively poor spherical symmetry [219], grew rapidly and many spheroids lost their structural integrity nine days after seeding formed with 1250, 2500, and 5000 cells per spheroid, and seven days after seeding for spheroids formed with 10000 cells spheroid. In addition, confocal microscopy could not be performed on day 7 after seeding for spheroids formed with 5000 and 10000 cells per spheroid due loss of structural integrity during harvesting. Identification of the necrotic region using image processing was more challenging, than for other cell lines, as a well-defined necrotic region did not form prior to the termination of the experiment. Therefore, necrotic region measurements for these spheroids are more subjective and uncertain. Spheroid boundaries were less smooth, so settings on the IncuCyte S3 live cell imaging system were updated to measure the largest brightfield object area with max eccentricity to 0.75 and sensitivity 20. These outer radius measurements were then manually reviewed to confirm accuracy.

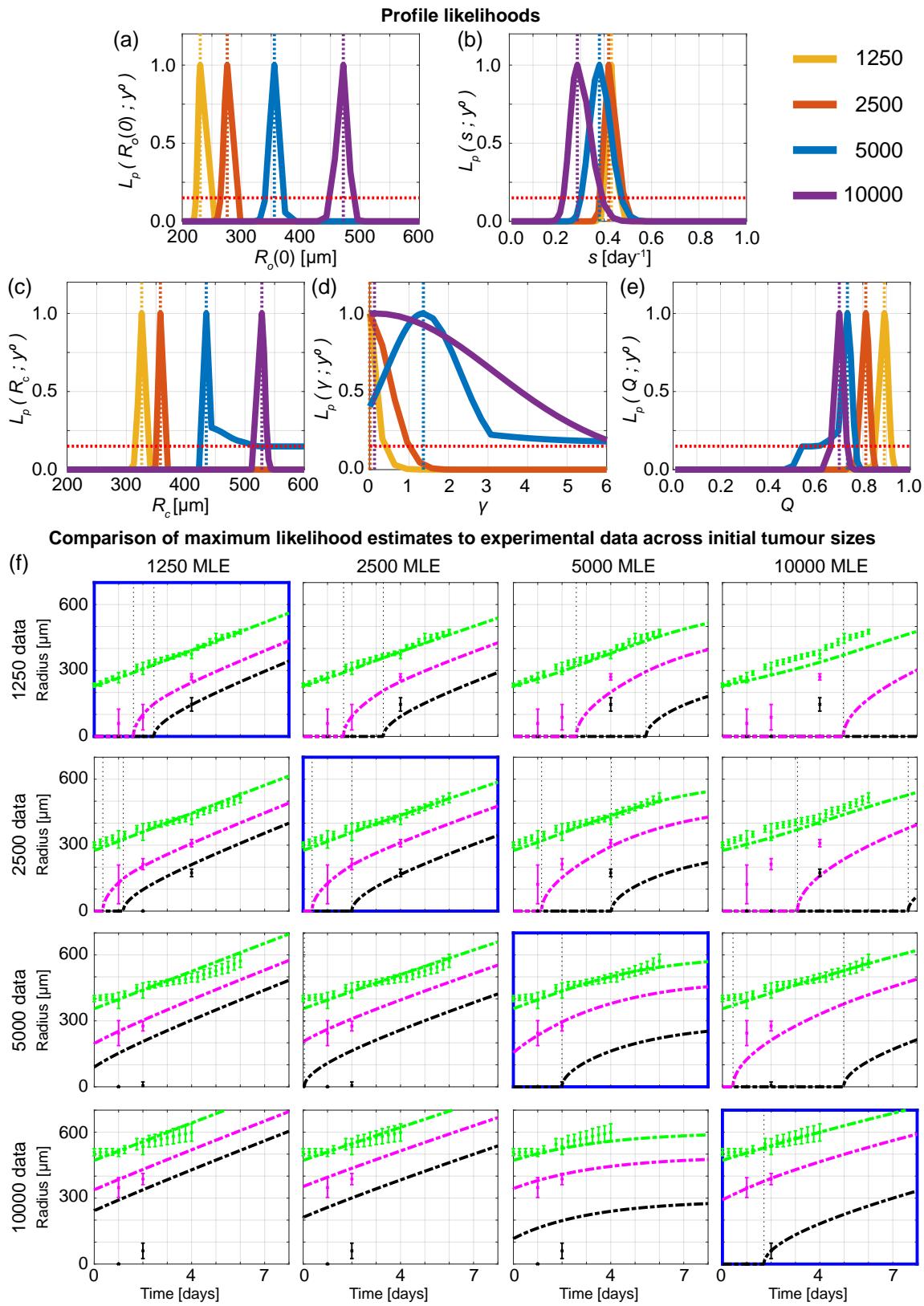
We perform analysis for WM164 spheroids using Days 1, 2, 3, 4 and 5 after formation, where measurements could be obtained. This means that we do not include the last day of outer radius measurements for spheroids formed with 1250, 2500, and 5000 cells per spheroid. This allows us to compare the final outer radius measurement to Greenspan's model simulated with the MLE as a predictive test. However, for spheroids formed with 10000 cells per spheroid we include all data points so cannot form a predictive test, but this is because we seek to obtain as much information as possible in the shorter experimental duration. While the experimental duration for WM164 spheroids is relatively short in comparison to the WM793b and WM983b experiments, these experiments are still performed for multiple days longer than previous WM164 spheroid experiments [87].

To perform the analysis we update initial parameter bounds, used for practical parameter identifiability analysis, for as  $200 < R_o(0) < 600 \text{ } [\mu\text{m}]$  and  $200 < R_c < 700 \text{ } [\mu\text{m}]$ . We update FirstGuess to  $(Q, \gamma, s, R_c, R_o(0)) = (0.8, 0.1, 0.5, 400, 210)$  for spheroids formed with 1250

and 2500 cells per spheroid, and to  $(Q, \gamma, s, R_c, R_o(0)) = (0.8, 0.1, 0.4, 400, 350)$  for spheroids formed with 5000 and 10000 cells per spheroid. Due to the reduced experimental duration for WM164 spheroids, and as we have already demonstrated with two other cell lines and synthetic data that Design 3 provides most insight, here we compare results obtained from spheroids with different initial sizes using Design 3.

In Figure 6A.40a, we observe four distinct narrow peaks for  $R_o(0)$  corresponding to the four initial spheroid sizes, which is expected. For  $s$ , we observe that profile likelihoods overlap showing information obtained for  $s$  is relatively consistent for different initial spheroid sizes (Figure 6A.40b). Interestingly and importantly, we observe four distinct peaks for  $R_c$  (Figure 6A.40c). This lack of consistency is different to the other two cell lines considered and strongly suggests information gained across initial spheroid sizes is not consistent. This is supported by direct observation of the experimental data where spheroids formed with 2500 cells have necrotic cores on day 4, whereas similar sized spheroids on day 2 formed with 5000 and 10000 cells per spheroid do not. Profile likelihoods for  $\gamma$  are wide for spheroids formed with 5000 and 10000 cells per spheroid, and narrow and overlapping for spheroids formed with 1250 and 2500 cell per spheroid, showing that  $\gamma$  requires more necrotic core measurements to be identified (Figure 6A.40d). Profile likelihoods for  $Q$  suggest that  $Q$  decreases as the initial spheroid size increases (Figure 6A.40e). This result for  $Q$  is less consistent and in contrast to results from other cell lines, where the profiles for  $Q$  overlapped for all spheroid sizes. Overall, we conclude that, possibly due to their lack of spherical symmetry, WM164 spheroids are more challenging to interpret and information gained using spheroids of different sizes is not consistent.

To support these results, we show along the diagonal of Figure 6A.40f the solution of the mathematical model evaluated at the MLE associated with each initial spheroid size compared to the experimental measurements. In doing so we demonstrate that we accurately predict the last outer radius measurement using previous days measurements for spheroids formed with 1250, 2500, and 5000 cells per spheroid. However, on the off-diagonals of Figure 6.4f, we compare how the Greenspan model simulated with the MLE from one initial spheroid size predicts data from different initial spheroid sizes by only changing the initial radius. These off-diagonal results show that using information from one spheroid size to predict the behaviour of a different spheroid size is not always accurate. For example, using information gained from spheroids formed with 10000 cell per spheroid poorly predicts the behaviour of spheroids formed with 1250 cell per spheroid, as the time evolution of the outer radius is not accurately predicted at late time and the inhibited and necrotic regions form much earlier than predicted (top right of Figure 6A.40f).



**Figure 6A.40:** Information gained from WM164 experiments across different initial tumour spheroid sizes is inconsistent. Profile likelihoods for (a)  $R_o(0)$ , (b)  $s$ , (c)  $R_c$ , (d)  $\gamma$ , (e)  $Q$ . Yellow, orange, blue, and purple lines in (a)-(e) represent profile likelihoods from WM164 spheroids formed with 1250, 2500, 5000, 10000 cells per spheroid, respectively, and the red-dashed line shows the approximate 95% confidence interval threshold. (f) Comparison of Greenspan model simulated with maximum likelihood estimates compared to full experimental data sets across initial tumour spheroid size, where error bars show standard deviation.



## **Chapter 7**

# **Conclusion**

## 7.1 Summary of the research

In this thesis we address two key aims. The first key aim is to develop a mathematical modelling framework to study epithelial tissue dynamics incorporating key biological processes, such as cell movement due to mechanical interactions, mechanical relaxation, cell proliferation, cell death, mechanical cell competition, mechanochemical coupling, and cell detachment at tissue boundaries due to epithelial-mesenchymal transitions. This framework is applicable to heterogeneous populations on free and fixed domains, and where the discrete model is prescribed based on biological observations and the corresponding continuum-limit model derived. The second key aim is to use mathematical models to quantitatively compare experimental designs for tumour spheroid experiments to reveal those design choices that are important and lead to reliable biological insight. This enables us to provide recommendations for future studies. To achieve these aims we develop new mathematical models, perform tumour spheroid experiments in the laboratory, and use statistical analysis. These two key aims correspond to Part 1 and Part 2 of this thesis and objectives 1-4, and 5, respectively, of this thesis:

### **Part 1: Mathematical model development**

1. Develop a discrete mathematical model to describe cell movement due to mechanical interactions in heterogeneous epithelial tissues and derive and compare to the corresponding continuum-limit model,
2. Extend objective 1 to develop a discrete mathematical model that includes cell proliferation and cell death in order to describe mechanical cell competition in heterogeneous epithelial tissues and derive and compare to the corresponding continuum model,
3. Extend objectives 2 and 3 to examine the travelling wave behaviour of the free boundary continuum model, incorporating cell movement due to mechanical interactions, cell proliferation, and cell death,
4. Extend objectives 2, 3, and 4 to develop a discrete mathematical model to describe the role of mechanical interactions in epithelial-mesenchymal transitions and derive and compare to the corresponding continuum model,

### **Part 2: Experimental design and mathematical modelling**

5. Perform tumour spheroid experiments in the laboratory and use statistical analysis with the Greenspan model to identify experimental design choices that are important and lead to reliable biological insight. Provide recommendations for future studies and demonstrate how to quantitatively compare data collected across different experimental designs.

These five objectives are successfully addressed in the five articles presented in Chapters 2-6 of this thesis, which we now summarise.

In Chapter 2, we develop a discrete to continuum modelling framework to study cell movement due to mechanical interactions in heterogeneous epithelial tissues. We provide a novel derivation that connects the discrete model to the continuum model by increasing the number of springs per cell, while maintaining the number of cells in the tissue and its fixed total length, and by considering spatial averages over length scales involving a large enough number of cells to define continuous densities but small enough to retain spatial heterogeneities. Excellent agreement is observed between solutions from the discrete model and the corresponding continuum model for slowly-varying-in-space and rapidly-varying-in-space heterogeneity, as the number of springs per cell increases. By dealing explicitly with heterogeneous cell populations this model has many potential applications. We consider two applications. First, we provide guidance how to infer cellular-level properties by tracking the interface between two distinct adjacent populations. We suggest that it is easier to determine the relative cell stiffnesses than it is to determine the relative resting cell lengths. Second, we use our model to interpret recent breast cancer detection experiments. Our results show that while a cell's resting cell length is not an easily measured experimental quantity it could be an important variable to consider.

In Chapter 3, we extend the model developed in Chapter 2 to incorporate cell proliferation and cell death, allowing us to explore mechanical cell competition. After prescribing the discrete model we derive the corresponding continuum model. In deriving the continuum model, the discrete mechanisms and assumptions that underpin the continuum model are made explicit by presenting the details of the coarse-graining derivations. This enables us to provide insights into when the continuum model is and is not valid. Therefore, under certain conditions we can recommend the discrete model is more appropriate. Further, we stress the limitations of developing continuum models by simply adding source and sink terms to an existing model without considering the underlying discrete model in complex biological systems. By coupling mechanics with cell proliferation and cell death we explore biological scenarios that could not be described with previous modelling frameworks. We focus on mechanical cell competition in heterogeneous epithelial tissues. By choosing mechanical relaxation rates sufficiently fast relative to proliferation rates we observe good agreement between the average of many identically prepared stochastic realisations of the discrete model and the corresponding solutions of the continuum model, even when there are a low number of cells in the tissue. We explore mechanical cell competition applied to cancer invasion by considering cancer cells adjacent

to healthy cells which compete for space. Interestingly, when we only allow cancer cells and healthy cells to differ in their cell stiffnesses, as a result of mechanical coupling, we observe that the cancer cells have more opportunities to proliferate and are less likely to die than healthy cells. We can then identify the cancer cells, as a result of the property of lower cell stiffness, as being the *winner* cells which invade the full domain. Cell stiffness and cell size may therefore be important factors to include when interpreting proliferation and death rates in experimental data.

In Chapter 4, we extend the work presented in Chapters 2-3 to include a free boundary. By focusing on the continuum model and using numerical simulations, phase plane and perturbation analysis, we find travelling wave solutions with negative, zero, and positive wavespeeds. Whether the cell population invades or retreats corresponds to whether cells at the carrying capacity density are in compression or in extension, respectively. Furthermore, unlike classical reaction-diffusion models, travelling wave solutions for this model have well-defined fronts and do not correspond to a heteroclinic orbit in the phase plane. We also obtain exact expressions for the speed of travelling wave solutions together with useful approximations of the shape of the travelling wave solutions.

In Chapter 5, we extend Chapters 2, 3, and 4 to incorporate chemical diffusion, and to incorporate a new free boundary condition describing cells detaching due to chemically-dependent epithelial-mesenchymal transitions (EMT). The coupling of mechanical interactions between cells and chemically-dependent EMT gives rise to a mechanochemical model. We focus on exploring the role of mechanical interactions in epithelial-mesenchymal transitions (EMT). We add chemically-dependent cell detachment, to represent EMT, into the modelling framework. This leads to a novel nonlinear free boundary problem where the boundary condition is derived from cell-level biological processes. This is unlike many other models in mathematical biology where the physical interpretation of the boundary conditions can be unclear. Both the discrete and continuum models provide useful information: discrete models show the important role of stochastic effects while continuum models help classify possible behaviours. Our results show good agreement between the continuum model and appropriately averaged quantities from many discrete realisations. Our models suggest that the coupling of mechanical interactions with EMT is important, can change the probability of long-term extinction significantly, and give rise to different rates of cell detachment. Using our model we postulate how one could encourage a wound to heal faster and how one may prevent cancer cells detaching from primary cancer tumours and spreading to other locations to form secondary tumours.

In Chapter 6, we transition to Part 2 of this thesis and focus on tumour spheroid experiments

where many mathematical models, with mechanochemical mechanisms, have been proposed but very few have been experimentally tested. Tumour spheroid experiments are routinely performed to study cancer progression and cancer treatments. However, experimental designs are inconsistent, leading to challenges in interpretability and reproducibility. First, we collect experimental data from tumour spheroid experiments across a range of experimental designs. Next, we develop an objective theoretical framework, using the seminal Greenspan mathematical model and statistical identifiability analysis, to quantitatively compare these data and provide recommendations for future experimental studies. We find that most insight is gained from measuring spheroids internal structure, and that frequency of measurement is less important. Furthermore, as information from tumour spheroids across varying initial spheroid sizes appears consistent, provided sufficient measurements at later times (in phase (iii)) are obtained, we recommend that performing experiments with larger tumour spheroids can be beneficial to obtain useful information in a shorter experimental duration. This framework can be applied to other tumour spheroid experiments grown in different conditions or with different cell types. The insights of this study provide a platform for future studies that develop, test, and examine cancer treatments with spheroids. In doing this work we also quantitatively test the seminal Greenspan mathematical model, which has been highly influential in mathematical biology, for the first time since its publication 50 years ago.

In preparation for Chapter 6, I was trained in a wet-laboratory to perform tumour spheroid experiments from start to finish including: cell culturing; spheroid formation; spheroid harvesting; spheroid fixing and mounting in preparation for imaging; confocal microscopy; and, image acquisition, processing, and analysis. Having started this PhD with no laboratory experience, the ability to perform biological experiments in combination with mathematical modelling and statistical analysis has been enlightening. For example, simple and fundamental biological questions were unknown, such as how does the internal structure of spheroid, and spatiotemporal heterogeneity with respect to the cell cycle change with time? how does the choice of the initial number of cells used to construct spheroid experiments influence the outcome of the experiment? and how do we compare and interpret tumour spheroid results when different measurements are made at different time points? We address these questions by developing an objective quantitative framework connecting experimental data, mathematical modelling, and statistical analysis, in Chapter 6.

Overall, the work in this thesis has explored mechanochemical and experimental models in mathematical biology. In the first part of this thesis we develop a discrete to continuum mechanochemical modelling framework for epithelial tissue dynamics incorporating key bio-

logical features. Using these models we interpret experimental data, provide guidance on how to infer cellular-level properties, identify that cell stiffness and cell size may be important factors when interpreting proliferation and death rates in experimental data, develop a quantitative framework to test hypotheses on mechanical cell competition, and highlight the role of mechanical interactions in epithelial-mesenchymal transitions. Without the modelling framework developed here such interpretations, observations, and recommendations would be difficult and not clear. Further this modelling framework is a platform for future studies on epithelial tissue dynamics, as outlined in Section 7.2. In the second part of the thesis we develop an objective theoretical framework, using the seminal Greenspan mathematical model and statistical identifiability analysis, to quantitatively compare experimental data that I collect in the laboratory across a range of experimental designs. In doing so we provide recommendations for future studies, such as measurements of internal spheroid structure provide the most insight, whereas varying initial spheroid size and temporal measurement frequency is less important. As tumour spheroid experiments are routinely performed to study cancer progression and treatment, the insights of this work are widely applicable and lay a foundation for future work incorporating drug treatments, spheroids grown with different conditions and with different cell types, and heterogeneous populations, as outlined in Section 7.2.

## 7.2 Future work

Both Part 1 and Part 2 of this thesis, studying epithelial tissue dynamics and tumour growth, respectively, provide a platform for many future studies. We now discuss some potential extensions.

### 7.2.1 Epithelial tissue dynamics

The theoretical foundations presented in this thesis for building a discrete model and constructing the continuum limit of that discrete model can be used to describe many additional mechanisms in future analyses.

#### Cell cycle

While some features of the cell cycle and cell ageing are implicit in this work, it would be of interest to explicitly incorporate the cell cycle and associated cell ageing processes such as growth in the resting cell length [147]. Furthermore, experimental results with real-time cell cycle imaging, using the same FUCCI technology described in Chapter 6, suggest that cell cycle progression in epithelial tissues is regulated cooperatively by forces between the dividing cell and its neighbours [234]. However, these observations appear not to have been incorporated into models of heterogeneous epithelial tissue dynamics. In this thesis we provide the foundations for such mechanisms to be included and explored.

#### Intestinal crypts, curved substrates, and tissue engineering

The work presented in this thesis focuses on one-dimensional epithelial tissues on flat substrates. However, this work can be extended to non-flat geometries, for example curved substrates and intestinal crypts [166, 167]. One expects that many results obtained in this thesis for one-dimensional flat substrates apply to curved substrates. For example, the continuum model is expected to hold as taking the limit as the number of springs per cell tends to infinity would correspond to approximating the curved substrate as flat. In contrast, the role of curvature of the substrate may be important in determining when the continuum model agrees with the discrete model. Extending this work to substrates that evolve in time, for example in bone [9, 94] or tissue engineering [32, 35], may also be interesting where the role of mechanics can be overlooked.

### Higher dimensions

The one-dimensional approach taken in this work has many advantages in its predictive power, interpretability, and relative computational simplicity in comparison to two- or three-dimensional models. Furthermore, cell-length-dependent proliferation may be thought of as an approximation for cell-volume-dependent proliferation which occurs for cells that move in three-dimensional environments. However, real cells can also spread without changing volume, so it may be beneficial to explore the role of the cell cycle in this one-dimensional framework [234]. A significant extension of this work would be to consider higher dimensions. The discrete model could be extended by considering a cell-centre or vertex model which introduces questions regarding cell shape and how neighbours can be identified, along with increased computational expense [67, 175, 179]. A corresponding continuum model in higher dimensions is less clear. The one-dimensional model enforces an ordering of neighbouring cells, which is important when deriving a continuum model [68, 163]. However, in higher dimensions cells can change their neighbours which poses significant challenges [68, 163]. The work completed in this thesis will provide great insights for such studies.

### Quantitatively connect to experimental data

Recent experimental studies have explored the role of mechanochemical processes in epithelial tissues [28, 98], mechanical waves in epithelial tissues [198], colliding epithelial tissues [95], and cell competition [130]. However, the corresponding mathematical models tend to neglect cell proliferation and cell death. In this thesis we have developed a mathematical framework that provides the platform to study these in greater depth. Therefore it would be of great interest to extend this work to such experimental studies. Further, it would be interesting to experimentally test predictions made in this work. One approach to quantitatively connect this model to experimental data could be to use statistical analysis, as described in Chapter 6 for tumour spheroid experiments.

### 7.2.2 Tumour growth

In the second part of this thesis we present an objective theoretical framework to quantitatively compare experimental data from tumour spheroid experiments performed across a range of experimental designs. This framework is suitable for many extensions.

#### Spheroids grown with different conditions

Our modelling framework is suitable to be extended to consider spheroids grown with different cell types and in different conditions. Such experiments could provide valuable insights. For example, Greenspan's model [79], which we use to interpret the experimental data, assumes that spatial gradients of diffusing nutrient and/or waste drive the governing behaviour. Therefore, performing perturbation experiments, such as varying the external nutrient concentration, would test key assumptions in the Greenspan model.

In addition, many studies in experimental biology focus on biological pathways from the molecular scale upwards. Using our framework one could interpret tumour spheroid experiments with and without gene knockdowns, use statistical identifiability analysis to determine the parameters of the model, and identify if they vary significantly. If parameters are significantly different one could suggest the gene that was knocked down plays an important role on tumour growth and internal structure. This provides a top down interpretation of experimental results as opposed to the traditional bottom up approach. This could be powerful in future studies and would be difficult without using the quantitative framework we outline in this thesis.

#### Heterogeneous populations

Much of the first part of this thesis focuses on developing a model for heterogeneous epithelial tissues. In the second part of this thesis we use a mathematical model for homogeneous populations, since the experiments that we perform use tumour spheroids grown from cell lines so treating the spheroid as a homogeneous population is realistic. However, *in vivo* tumours are heterogeneous. Therefore, considering heterogeneous populations both experimentally, using tumour spheroids generated from multiple cell lines or patient derived organoids, and mathematically, possibly by extending Greenspans model, may more accurately mimic real life scenarios and provide even greater insight. This work would directly build upon the work presented in this thesis and would be difficult without using the quantitative framework we outline in this thesis.

### 7.3 Final remarks

In this thesis we explore mechanochemical and experimental models in mathematical biology. We develop new mathematical models for epithelial tissue dynamics, and perform tumour spheroid experiments in a wet-laboratory to then examine a range of experimental designs using mathematical modelling and statistical analysis. As a result of the work in this thesis we reveal that many measurements and processes, often neglected, may be important. We demonstrate that power of the discrete-to-continuum modelling approach, and provide a platform for many future studies exploring mechanochemical processes. Furthermore, by developing an objective theoretical framework, using the seminal Greenspan mathematical model and statistical identifiability analysis, we demonstrate how to quantitatively compare experimental data across experimental designs and address basic and fundamental biological questions. We provide recommendations that will benefit and guide routinely performed tumour spheroid experiments to accelerate the study of cancer progression and treatment. For future studies in these research areas, we recommend continued cross-disciplinary collaboration using tools from experimental biology, mathematical modelling, and statistical analysis.

# Bibliography

- [1] MathWorks `fsolve`. <https://au.mathworks.com/help/optim/ug/fsolve.html>. Accessed August 2021.
- [2] MathWorks `interp1`. <https://au.mathworks.com/help/matlab/ref/interp1.html>. Accessed August 2021.
- [3] MathWorks `normpdf`. <https://au.mathworks.com/help/stats/normpdf.html>. Accessed August 2021.
- [4] MathWorks `roots`. <https://au.mathworks.com/help/matlab/ref/roots.html>. Accessed August 2021.
- [5] Optimization toolbox: User's guide (R2019a). [https://au.mathworks.com/help/pdf\\_doc/optim/optim\\_tb.pdf](https://au.mathworks.com/help/pdf_doc/optim/optim_tb.pdf). Accessed: March 27, 2019.
- [6] M. L. Ackland, D. F. Newgreen, M. Fridman, M. C. Waltham, A. Arvanitis, J. Minichiello, J. T. Price, and E. W. Thompson. Epidermal growth factor-induced epithelial-mesenchymal transition in human breast carcinoma cells. *Laboratory Investigation*, 83:435—448, 2003.
- [7] J. A. Adam and S. A. Maggelakis. Diffusion regulated growth characteristics of a spherical prevascular carcinoma. *Bulletin of Mathematical Biology*, 52:549–582, 1990.
- [8] A. Albanese, A. K. Lam, E. A. Sykes, J. V. Rocheleau, and W. C. W. Chan. Tumour-on-a-chip provides an optical window into nanoparticle tissue transport. *Nature Communications*, 4:2718, 2013.
- [9] M. A. Alias and P. R. Buerzli. A level-set method for the evolution of cells and tissue during curvature-controlled growth. *International Journal for Numerical Methods in Biomedical Engineering*, 36:e3279, 2019.
- [10] C. Alibert, B. Goud, and J-B. Manneville. Are cancer cells really softer than normal cells? *Biology of the Cell*, 109:167–189, 2017.
- [11] S. J. Altschuler and L. F. Wu. Cellular heterogeneity: when do differences make a difference? *Cell*, 141:559–563, 2010.
- [12] V. Andasari, R. T. Roper, M. H. Swat, and M. A. J. Chaplain. Integrating intracellular dynamics using CompuCell3D and Bionetsolver: Applications to multiscale modelling of cancer cell growth and invasion. *PLoS One*, 7:e33726, 2012.
- [13] A. R. A. Anderson. A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion. *Mathematical Medicine and Biology*, 22:163–186, 2005.
- [14] S. S. Antman. *Nonlinear Problems of Elasticity*. Springer, New York, NY, vol 107 of Applied Mathematical Sciences edition, 2005.
- [15] J. Antony, J. P. Theiry, and R. Y. Huang. Epithelial-to-mesenchymal transition: lessons from development, insights into cancer and the potential of EMT-subtype based therapeutic intervention. *Physical Biology*, 16:041004, 2019.

- [16] R. P. Araujo and D. L. S. McElwain. A history of the study of solid tumour growth: the contribution of mathematical modelling. *Bulletin of Mathematical Biology*, 66:1039, 2004.
- [17] N. J. Armstrong, K. J. Painter, and J. A. Sherratt. A continuum approach to modelling cell-cell adhesion. *Journal of Theoretical Biology*, 243:98–113, 2006.
- [18] K. Bacevic, R. Noble, A. Soffar, O. W. Ammar, B. Boszonyik, S. Prieto, C. Vincent, M. E. Hochberg, L. Krasinska, and D. Fisher. Spatial competition constrains resistance to targeted cancer therapy. *Nature Communications*, 8:1995, 2017.
- [19] R. E. Baker, A. Parker, and M. J. Simpson. A free boundary model of epithelial dynamics. *Journal of Theoretical Biology*, 481:61–74, 2019.
- [20] R. E. Baker, C. A. Yates, and R. Erban. From microscopic to macroscopic descriptions of cell migration on growing domains. *Bulletin of Mathematical Biology*, 72:719—762, 2010.
- [21] M Basan, T. Risler, J. Joanny, X. Sastre-Garau, and J. Prost. Homeostatic competition drives tumor growth and metastasis nucleation. *HFSP Journal*, 3:265–272, 2009.
- [22] S. Benzekry, C. Lamont, A. Beheshti, A. Tracz, J. M. L. Ebos, L. Hlatky, and P. Hahnfeldt. Classical mathematical models for description and prediction of experimental tumor growth. *PLOS Computational Biology*, 10:e1003800, 2014.
- [23] S. Bhatia, J. Monkman, A. K. L. Toh, S. H. Nagaraj, and E. W. Thompson. Targeting epithelial–mesenchymal plasticity in cancer: clinical and preclinical advances in therapy and monitoring. *Biochemical Journal*, 474:3269—3306, 2008.
- [24] S. Bhatia, P. Wang, A. Toh, and E. W. Thompson. New insights into the role of phenotypic plasticity and emt in driving cancer progression. *Frontier in Molecular Biosciences*, 7:71, 2020.
- [25] F. Bocci, L. Gearhart-Serna, M. Boareto, M. Ribeiro, E. Ben-Jacob, G. R. Devi, H. Levine, J. N. Onuchic, and M. K. Jolly. Toward understanding cancer stem cell heterogeneity in the tumor microenvironment. *Proceedings of the National Academy of Sciences of the United States of America*, 116:148–157, 2019.
- [26] F. Bocci, M. K. Jolly, and J. N. Onuchic. A biophysical model uncovers the size distribution of migrating cell clusters across cancer types. *Cancer Research*, 79:5527–5535, 2019.
- [27] M. Bodnar and J. Velazquez. Derivation of macroscopic equations for individual cell-based models: a formal approach. *Mathematical Methods in the Applied Sciences*, 28:1757–1779, 2005.
- [28] D. Boocock, N. Hino, N. Ruzickova, T. Hirashima, and E. Hannezo. Theory of mechano-chemical patterning and optimal migration in cell monolayers. *Nature Physics*, 95, 2020.
- [29] C. Bras-Pereira and E. Moreno. Mechanical cell competition. *Current Opinion in Cell Biology*, 51:15–21, 2018.
- [30] N. F. Britton. *Essential Mathematical Biology*. Springer, London, UK, 2003.
- [31] A. P. Browning, P. Haridas, and M. J. Simpson. A Bayesian sequential learning framework to parameterise continuum models of melanoma invasion into human skin. *Bulletin of Mathematical Biology*, 81:676–698, 2019.
- [32] A. P. Browning, O. J. Maclaren, P. R. Buerzli, M. Lanaro, M. C. Allenby, M. A. Woodruff, and M. J. Simpson. Model-based data analysis of tissue growth in thin 3D printed scaffolds. *Journal of Theoretical Biology*, 528:110852, 2021.

- [33] A. P. Browning and R. J. Murphy. Image processing algorithm to identify structure of tumour spheroids with cell cycle labelling. *Zenodo*, 2021.
- [34] P. R. Buerzli. Governing equations of tissue modelling and remodelling: A unified generalised description of surface and bulk balance. *PLOS ONE*, 11:e0152582, 2016.
- [35] P. R. Buerzli, M. Lanaro, C. S. Wong, M. P. McLaughlin, M. C. Allenby, M. A. Woodruff, and M. J. Simpson. Cell proliferation and migration explain pore bridging dynamics in 3D printed scaffolds of different pore size. *Acta Biomaterialia*, 114:285–295, 2020.
- [36] J. A. Bull, F. Mech, T. Quasier, S. L. Waters, and H. M. Byrne. Mathematical modelling reveals cellular dynamics within tumour spheroids. *PLoS Computational Biology*, 16:e1007961, 2020.
- [37] H. M. Byrne. Dissecting cancer through mathematics: from the cell to the animal model. *Nature Reviews Cancer*, 10:221–230, 2010.
- [38] C. Cadard, L. Venkova, P. Recho, M. C. Lagomarsino, and M. Piel. The physics of cell-size regulation across timescales. *Nature Physics*, 15:993–1004, 2019.
- [39] M. J. Carr, M. J. Simpson, and C. Drovandi. Estimating parameters of a stochastic cell invasion model with fluorescent cell cycle labelling using approximate Bayesian computation. *To appear, Journal of the Royal Society Interface*, 2021.
- [40] J. J. Casciari, S. V. Sotirchos, and R. M. Sutherland. Mathematical modelling of microenvironment and growth in EMT6/Ro multicellular tumour spheroids. *Cell Proliferation*, 25:1–22, 1992.
- [41] T. Celià-Terrassa, C. Bastain, D. D. Liu, B. Ell, N. M. Aiello, Y. Wei, J. Zamalloa, A. M. Blanco, X. Hang, D. Kunisky, et al. Hysteresis control of epithelial-mesenchymal transition dynamics conveys a distinct program with enhanced metastatic ability. *Nature Communications*, 9:5005, 2018.
- [42] C. L. Chaffer and R. A. Weinberg. A perspective on cancer cell metastasis. *Science*, 331:1559–1564, 2011.
- [43] L. A. Chapman. *Mathematical modelling of cell growth in tissue engineering bioreactors*. PhD thesis, Oxford University, UK, 2012.
- [44] S. C. Chapra and R. P. Canale. *Numerical Methods for Engineers*. McGraw-Hill, New York, NY, 6th edition, 2010.
- [45] J. Collis, D. L. Brown, M. E. Hubbard, and R. D. O’Dea. Effective equations governing an active poroelastic medium. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 473:20160755, 2017.
- [46] E. C. Costa, A. F. Moreira, D. de Melo-Diogo, V. M. Gaspar, M. P. Carvalho, and I. J. Correia. 3D tumor spheroids: an overview on the tools and techniques used for their analysis. *Biotechnology Advances*, 34:1427–1441, 2016.
- [47] E. C. Costa, D. B. Silva, A. F. Moreira, and I. J. Correia. Optical clearing methods: an overview of the techniques used for the imaging of 3D spheroids. *Biotechnology & Bioengineering*, 116:2742–2763, 2019.
- [48] E. J. Crampin, E. A. Gaffney, and P. K. Maini. Reaction and diffusion on growing domains: Scenarios for robust pattern formation. *Bulletin of Mathematical Biology*, 61:1093—1120, 1999.
- [49] E. J. Crampin, W. W. Hackborn, and P. K. Maini. Pattern formation in reaction-diffusion models with nonuniform domain growth. *Bulletin of Mathematical Biology*, 64:747—769, 2002.
- [50] J. Crank. *The Mathematics of Diffusion*. Oxford University Press, Oxford, UK, 2nd ed. edition, 1975.

- [51] J. Crank. *Free and Moving Boundary Problems*. Clarendon Press, Oxford, 1984.
- [52] S. E. Cross, Y. Jin, J. Rao, and J. K. Gimzewski. Nanomechanical analysis of cells from cancer patients. *Nature Nanotechnology*, 2:780–783, 2007.
- [53] M. Delarue, F. Montel, D. Vignjevic, J. Prost, J. Joanny, and G. Cappello. Compressive stress inhibits proliferation in tumor spheroids through a volume limitation. *Biophysical Journal*, 107:1821–1828, 2014.
- [54] M. E. Dolega, M. Delarue, F. Ingremeau, J. Prost, A. Delon, and G. Cappello. Cell-like pressure sensors reveal increase of mechanical stress towards the core of multicellular spheroids under compression. *Nature Communications*, 8:14056, 2016.
- [55] A. Dongre and R. A. Weinberg. New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. *Nature Reviews Molecular Cell Biology*, 20:69–84, 2019.
- [56] L. Edelstein-Keshet. *Mathematical Models in Biology*. SIAM Classics in Applied Mathematics, Philadelphia, PA, 2005.
- [57] M. El-Hachem, S. W. McCue, W. Jin, Y. Du, and M. J. Simpson. Revisiting the Fisher–Kolmogorov–Petrovsky–Piskunov equation to interpret the spreading–extinction dichotomy. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 475:20190378, 2019.
- [58] M. El-Hachem, S. W. McCue, and M. J. Simpson. A sharp-front moving boundary model for malignant invasion. *Physica D: Nonlinear Phenomena*, 412:132639, 2020.
- [59] S. Engblom, L. Ferm, A. Hellander, and P. Lötstedt. Simulation of stochastic reaction-diffusion processes on unstructured meshes. *SIAM Journal on Scientific Computing*, 31:1774–1797, 2009.
- [60] D. J. Evans and G. Morriss. *Statistical Mechanics of Nonequilibrium Liquids*. Cambridge University Press, Cambridge, UK, 2008.
- [61] N. D. Evans, R. O. C. Oreffo, E. Healy, P. J. Thurner, and Y. H. Man. Epithelial mechanobiology, skin wound healing, and the stem cell niche. *Journal of the Mechanical Behavior of Biomedical Materials*, 28:397–409, 2013.
- [62] N. T. Fadai. Semi-infinite travelling waves arising in a general reaction–diffusion Stefan model. *Nonlinearity*, 34:725–743, 2021.
- [63] N. T. Fadai and M. J. Simpson. New travelling wave solutions of the Porous–Fisher model with a moving boundary. *Journal of Physics A: Mathematical and Theoretical*, 53:095601, 2020.
- [64] R. A. Fisher. The wave of advance of advantageous genes. *Annals of Eugenics*, 7:355–369, 1937.
- [65] J. A. Flegg and N. Nataraj. Mathematical modelling and avascular tumour growth. *Resonance*, 24:313–325, 2019.
- [66] A. G. Fletcher, F. Cooper, and R. E. Baker. Mechanocellular models of epithelial morphogenesis. *Philosophical Transactions of the Royal Society B*, 327:20150519, 2017.
- [67] A. G. Fletcher, M. Osterfield, R. E. Baker, and S. Y. Shvartsman. Vertex models of epithelial morphogenesis. *Biophysical Journal*, 106:2291–2304, 2014.
- [68] J. A. Fozard, H. M. Byrne, O. E. Jensen, and J. R. King. Continuum approximations of individual-based models for epithelial monolayers. *Mathematical Medicine and Biology*, 27:39–74, 2010.

- [69] L. C. Franssen and M. A. J. Chaplain. A mathematical multi-organ model for bidirectional epithelial-mesenchymal transitions in the metastatic spread of cancer. *IMA Journal of Applied Mathematics*, 85:724–761, 2020.
- [70] L. C. Franssen, T. Lorenzi, A. E. F. Burgess, and M. A. J. Chaplain. A mathematical framework for modelling the metastatic spread of cancer. *Bulletin of Mathematical Biology*, 81:1965–2010, 2019.
- [71] J. Friedrich, C. Seidel, R. Ebner, and L. A. Kunz-Schughart. Spheroid-based drug screen considerations and practical approach. *Nature Protocols*, 4:309–324, 2009.
- [72] J. A. Gallaher, P. M. Enriquez-Nava, K. A. Luddy, R. A. Gatenby, and A. R. A. Anderson. Spatial heterogeneity and evolutionary dynamics modulate time to recurrence in continuous and adaptive cancer therapies. *Cancer Research*, 78:2127–2139, 2017.
- [73] 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. *88*, 88:62–75, 2005.
- [74] A. Gelman, J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin. *Bayesian Data Analysis*. Chapman and Hall/CRC, New York, NY, 3rd edition, 2013.
- [75] S. A. H. Geritz and A. Kisdi. Mathematical ecology: why mechanistic models? *Journal of Mathematical Biology*, 65:1411–1415, 2012.
- [76] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81:2340–2361, 1977.
- [77] N. Gjorevski, E. Boghaert, and CM. Nelson. Regulation of the epithelial-mesenchymal transition by transmission of mechanical stress through epithelial tissues. *Cancer Microenvironment*, 5:29–38, 2012.
- [78] A. Goriely. *The Mathematics and Mechanics of Growth*. Springer, New York, USA, 2017.
- [79] H. P. Greenspan. Models for the growth of a solid tumor by diffusion. *Studies in Applied Mathematics*, 51:317–340, 1972.
- [80] D. R. Grimes, C. Kelly, K. Bloch, and M. Partridge. A method for estimating the oxygen consumption rate in multicellular tumour spheroids. *Journal of the Royal Society Interface*, 11:20131124, 2014.
- [81] J. Guck, S. Schinkinger, B. Lincoln, F. Wottawah, S. Ebert, M. Romeyke, D. Lenz, H. M. Erickson, R. Ananthakrishnan, D. Mitchell, J. Käs, S. Ulwick, and C. Bilby. Optical deformability as an inherent cell marker for testing malignant transformation and metastatic competence. *Biophysical Journal*, 88:3689–6398, 2005.
- [82] S. A. Gudipaty, J. Lindblom, P. D. Loftus, M. J. Redd, K. Edes, C. F. Davey, F. Krishnegowda, and J. Rosenblatt. Mechanical stretch triggers rapid epithelial cell division through piezo1. *Nature*, 543:118–121, 2017.
- [83] L. Guillaume, L. Rigal, J. Fehrenbach, C. Severac, B. Ducommun, and V. Lobjois. Characterization of the physical properties of tumor-derived spheroids reveals critical insights for pre-clinical studies. *Scientific Reports*, 9:6597, 2019.
- [84] N. P. A. D. Gunasinghe, A. Wells, E. W. Thompson, and H. J. Hugo. Mesenchymal-epithelial transition (MET) as a mechanism for metastatic colonisation in breast cancer. *Cancer and Metastasis Reviews*, 31:469–478, 2012.
- [85] M. Gyllenberg and G. F. Webb. Quiescence as an explanation of Gompertzian tumour growth. *Growth, Development, and Aging: GDA*, 53:25–33, 1989.

- [86] N. K. Haass. Dynamic tumor heterogeneity in melanoma therapy: how do we address this in a novel model system? *Melanoma Management*, 2:93–95, 2015.
- [87] N. K. Haass, K. A. Beaumont, D. S. Hill, A. Anfosso, P. Mrass, M. A. Munoz, I. Kinjyo, and W. Weninger. Real-time cell cycle imaging during melanoma growth, invasion, and drug response. *Pigment Cell & Melanoma Research*, 27:764–776, 2014.
- [88] K. Han, S. E. Pierce, A. Li, K. Spees, G. R. Anderson, J. A. Seoane, Y. Lo, M. Dubreuil, M. Olivas, R. A. Kamber, M. Wainberg, K. Kostyrko, M. R. Kelly, M. Yousefi, S. W. Simpkins, D. Yao, K. Lee, C. J. Kuo, P. K. Jackson, A. Sweet-Cordero, A. Kundaje, A. J. Gentles, M. M. Curtis, C. amd Winslow, and M. C. Bassik. CRISPR screens in cancer spheroids identify 3D growth-specific vulnerabilities. *Nature*, 580:136–141, 2020.
- [89] Y. L. Han, A. F. Pegoraro, H. Li, K. Li, Y. Yuan, G. Xu, Z. Gu, J. Sun, Y. Hao, S. K. Gupta, Y. Li, W. Tang, H. Kang, L. Teng, J. J. Fredberg, and M. Guo. Cell swelling, softening and invasion in a three-dimensional breast cancer model. *Nature Physics*, 16:101–108, 2019.
- [90] D. Hanahan and R. A. Weinberg. Hallmarks of cancer: the next generation. *Cell*, 144:646–674, 2011.
- [91] E. Hannezo and C-P. Heisenberg. Mechanochemical feedback loops in development and disease. *Cell*, 178:12–25, 2019.
- [92] P. Haridas, J. A. McGovern, D. L. S. McElwain, and M. J. Simpson. Quantitative comparison of the spreading and invasion of radial growth phase and metastatic melanoma cells in an three-dimensional human skin equivalent model. *Peer J*, 5:e3754, 2017.
- [93] P. Haridas, C. J. Penington, J. A. McGovern, D. L. S. McElwain, and M. J. Simpson. Quantifying rates of cell migration and cell proliferation in co-culture barrier assays reveals how skin and melanoma cells interact during melanoma spreading and invasion. *Journal of Theoretical Biology*, 423:13–25, 2017.
- [94] S. G. D. Hegarty-Cremer, M. J. Simpson, T. L. Andersen, and P. R. Buenzli. Modelling cell guidance and curvature control in evolving biological tissues. *Journal of Theoretical Biology*, 520:110658, 2021.
- [95] M. A. Heinrich, R. Alert, A. E. Wolf, A. Košmrlj, and D. J. Cohen. Self-assembly of tessellated tissue sheets by growth and collision.
- [96] D. Hilhorst and Y. Kim. Diffusive and inviscid traveling waves of the Fisher equation and the nonuniqueness of wave speed. *Applied Mathematics Letters*, 60:28–35, 2016.
- [97] K. E. Hines, T. R. Middendorf, and R. W. Aldrich. Determination of parameter identifiability in nonlinear biophysical models: a Bayesian approach. *The Journal of General Physiology*, 143:401–416, 2014.
- [98] N. Hino, L. Rossetti, A. Marín-Llauradó, K. Aoki, X. Trepat, M. Matsuda, and T. Hirashima. ERK-mediated mechanochemical waves direct collective cell polarization. *Developmental Cell*, 53:646–660, 2020.
- [99] F. Hirschhaeuser, H. Menne, C. Dittfeld, J. West, W. Mueller-Klieser, and L. A. Kunz-Schughart. Multicellular tumor spheroids: An underestimated tool is catching up again. *Journal of Biotechnology*, 148:3–15, 2010.
- [100] K. S. Hoek, N. C. Schlegel, P. Brafford, A. Sucker, S. Ugurel, R. Kumar, B. L. Weber, K. L. Nathanson, D. J. Philips, M. Herlyn, D. Schadendorf, and R. Dummer. Metastatic potential of melanomas defined by specific gene expression profiles with no BRAF signature. *Pigment Cell Research*, 19:290–302, 2006.
- [101] W. R. Holmes and L. Edelstein-Keshet. Analysis of a minimal Rho-GTPase circuit regulating cell shape. *Physical Biology*, 13:046001, 2016.

- [102] H. Hugo, M. L. Ackland, T. Blick, M. G. Lawrence, J. A. Clements, E. D. Williams, and E. W. Thompson. Epithelial-mesenchymal and mesenchymal-epithelial transitions in carcinoma progression. *Journal of Cellular Physiology*, 213:374–383, 2007.
- [103] C. W. Huo, G. L. Chew, K. L. Britt, W. V. Ingman, M. A. Henderson, J. L. Hopper, and E. W. Thompson. Mammographic density-a review on the current understanding of its association with breast cancer. *Breast Cancer Research and Treatment*, 144:479–502, 2014.
- [104] D. E. Ingber. Tensegrity I. cell structure and hierarchical systems biolog. *Journal of Cell Science*, 116:1157–1173, 2003.
- [105] N. Jagiella, B. Müller, M. Müller, I. E. Vignon-Clementel, and D. Drasdo. Inferring growth control mechanisms in growing multi-cellular spheroids of NSCLC cells from spatial-temporal image data. *PLOS Computational Biology*, 12:e1004412, 2016.
- [106] Y. Jiang, J. Pjesivac-Grbovic, C. Cantrell, and J. P. Freyer. A multiscale model for avascular tumor growth. *Biophysical Journal*, 89:3884–3894, 2005.
- [107] S. Jin. Bipotent stem cells support the cyclical regeneration of endometrial epithelium of the murine uterus. *Proceedings of the National Academy of Sciences of the United States of America*, 116:6848–6857, 2019.
- [108] W. Jin, L. Spoerri, N. K. Haass, and M. J. Simpson. Mathematical model of tumour spheroid experiments with real-time cell cycle imaging. *Bulletin of Mathematical Biology*, 83:44, 2021.
- [109] S. T. Johnston, M. J. Simpson, and M. J. Plank. Lattice-free descriptions of collective motion with crowding and adhesion. *Physical Review E*, 88:062720, 2013.
- [110] M. K. Jolly, R. J. Murphy, S. Bhatia, H. J. Whitfield, A. Redfern, M. J. Davis, and E. W. Thompson. Measuring and modelling the epithelial mesenchymal hybrid state in cancer: clinical implications. *Cells Tissue Organs*, 2021.
- [111] M. K. Jolly, S. C. Tripathi, J. A. Somarelli, S. M. Hanash, and H. Levine. Epithelial/mesenchymal plasticity how have quantitative mathematical models helped improve our understanding? *Molecular Oncology*, 11:739–754, 2017.
- [112] Y. Katsuno, D. S. Meyer, Z. Zhang, K. M. Shokat, R. J. Akhurst, K. Miyazono, and R. Deryck. Chronic TGF- $\beta$  exposure drives stabilized EMT, tumor stemness, and cancer drug resistance with vulnerability to bitopic mTOR inhibition. *Science Signaling*, 12:eaau8544, 2019.
- [113] D. Khuperkar, M. Helen, I. Magre, and J. Joseph. Inter-cellular transport of Ran GTPase. *PLOS One*, 10:e0125506, 2015.
- [114] A. Kienzle, S. Kurch, J. Schlöder, C. Berges, R. Ose, J Schupp, A. Tuettenberg, H. Weiss, J. Schultze, S. Winzen, M. Schinnerer, K. Koynov, M. Mezger, N. K. Haass, W. Tremel, and H. Jonuleit. Dendritic mesoporous silica nanoparticles for pH-stimuli-responsive drug delivery of TNF-alpha. *Advanced Healthcare Materials*, 6:1700012, 2017.
- [115] H. Y. Kim, T. R. Jackson, and L. A. Davidson. On the role of mechanics in driving mesenchymal-to-epithelial transitions. *Seminars in Cell and Developmental Biology*, 67:113—122, 2017.
- [116] G. J. Kimmel, P. Gerlee, and P. M. Altrock. Time scales and wave formation in non-linear spatial public good games. *PLoS Computational Biology*, 15:e1007361–22, 2019.

- [117] K. L. Kiran, D. Jayachandran, and S. Lakshminarayanan. Mathematical modelling of avascular tumour growth based on diffusion of nutrients and its validation. *The Canadian Journal of Chemical Engineering*, 87:732–740, 2009.
- [118] M. Kot. *Elements of Mathematical Ecology*. Cambridge University Press, Cambridge, UK, 2001.
- [119] L. A. Kunz-Schughart, M. Kreutz, and R. Knuechel. Multicellular spheroids: a three-dimensional *in vitro* culture system to study tumour biology. *International Journal of Experimental Pathology*, 79:1–23, 1998.
- [120] S. Kutluay, A. R. Bahadir, and A. Özdeş. The numerical solution of one-phase classical Stefan problem. *Journal of Computational and Applied Mathematics*, 81:135–144, 1997.
- [121] S. Lagies, M. Schlimpert, S. Neumann, A. Wäldin, B. Kammerer, C. Borner, and L. Peintner. Cells grown in three-dimensional spheroids mirror *in vivo* metabolic response of epithelial cells. *Communications Biology*, 3:246, 2020.
- [122] K. A. Landman and C. P. Please. Tumour dynamics and necrosis: surface tension and stability. *IMA Journal of Mathematics Applied in Medicine and Biology*, 18:131–158, 2001.
- [123] K. A. Landman, M. J. Simpson, J. L. Slater, and D. F. Newgreen. Diffusive and chemotactic cellular migration: smooth and discontinuous traveling wave solutions. *SIAM Journal on Applied Mathematics*, 65:1420–1442, 2005.
- [124] J. M. Lee, S. Dedhar, R. Kalluri, and E. W. Thompson. The epithelial–mesenchymal transition: new insights in signaling, development, and disease. *The Journal of Cell Biology*, 172:973–981, 2006.
- [125] S. Lee and Y. Morishita. Possible roles of mechanical cell elimination intrinsic to growing tissues from the perspective of tissue growth efficiency and homeostasis. *PLOS Computational Biology*, 13:e1005651, 2017.
- [126] M. Lekka. Atomic force microscopy: a tip for diagnosing cancer. *Nature Nanotechnology*, 7:691–692, 2012.
- [127] M. Lekka. Discrimination between normal and cancerous cells using AFM. *BioNanoScience*, 6:65–80, 2016.
- [128] M. Lekka, D. Gil, K. Pogoda, J. Dulińska-Litewka, R. Jach, J. Gostek, O. Klymenko, S. Prauzner-Bechcicki, Z. Stachura, J. Wiltowska-Zuber, K. Okoń, and P. Laidler. Cancer cell detection in tissue sections using afm. *Archives of Biochemistry and Biophysics*, 518:151–156, 2012.
- [129] R. Levayer. Solid stress, competition for space and cancer: The opposing roles of mechanical cell competition in tumour initiation and growth. *Seminars in Cancer Biology*, 63:69–80, 2019.
- [130] R. Levayer. Cell competition: Bridging the scales through cell-based modeling. *Current Biology*, 31:R856–R858, 2021.
- [131] I. Levental, P. C. Georges, and P. A. Janmey. Soft biological materials and their impact on cell function. *Soft Matter*, 3:299–306, 2006.
- [132] Y. Li, P. van Heijster, R. Marangell, and M. J. Simpson. Travelling wave solutions in a negative nonlinear diffusion-reaction model. *Journal of Mathematical Biology*, 81:1495–1522, 2020.
- [133] M. J. Lighthill. *An Introduction to Fourier Analysis and Generalised Functions*. Cambridge University Press, Cambridge, UK, 1958.
- [134] S. Liu, Y. Cong, D. Wang, Y. Sun, L. Deng, Y. Liu, R. Martin-Trevino, L. Shang, S. P. McDermott, M. D. Landis, et al. Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. *Stem Cell Reports*, 2:78–91, 2014.

- [135] D. Loessner, J. A. Flegg, H. M. Byrne, J. A. Clements, and D. W. Hutmacher. Growth of confined cancer spheroids: a combined experimental and mathematical modelling approach. *Integrative Biology*, 5:597–605, 2013.
- [136] T. Lorenzi, P. Murray, and M. Ptashnyk. From individual-based mechanical models of multicellular systems to free-boundary problems. *Interfaces and Free Boundaries*, 22:205–244, 2020.
- [137] M. Lu, M. K. Jolly, H. Levine, J. N. Onuchic, and E. Ben-Jacob. MicroRNA-based regulation of epithelial-hybrid-mesenchymal fate determination. *Proceedings of the National Academy of Sciences*, 110:18144–18149, 2013.
- [138] K. M. Lucas, N. Mohana-Kumara, D. Lau, X. D. Zhang, P. Hersey, D. C. Huang, W. Weninger, N. K. Haass, and J. D. Allen. Modulation of NOXA and MCL-1 as a strategy for sensitizing melanoma cells to the BH3-Mimnetic ABT-737. *Clinical Cancer Research*, 18:783–795, 2012.
- [139] R. Lui and Z. A. Wang. Traveling wave solutions from microscopic to macroscopic chemotaxis models. *Journal of Mathematical Biology*, 61:739–761, 2010.
- [140] P. K. Maini, D. L. S. McElwain, and D. Leavesley. Travelling wave model to interpret a wound-healing cell migration assay for human peritoneal mesothelial cells. *Tissue Engineering*, 10:475–482, 2004.
- [141] P. K. Maini, D. L. S. McElwain, and D. Leavesley. Travelling waves in a wound healing assay. *Applied Mathematics Letters*, 17:575–580, 2004.
- [142] F. Marcucci, G. Stassi, and R. De Maria. Epithelial-mesenchymal transition: a new target in anticancer drug discovery. *Nature Reviews Drug Discovery*, 15:311–325, 2016.
- [143] D. C. Markham, M. J. Simpson, P. K. Maini, E. A. Gaffney, and R. E. Baker. Comparing methods for modelling spreading cell fronts. *Journal of Theoretical Biology*, 353:95–103, 2014.
- [144] A. Matamoro-Vidal and R. Levayer. Multiple influences of mechanical forces on cell competition. *Current Biology*, 29:R762–R774, 2019.
- [145] MATLAB. 9.6.0.1150989 (R2019a) Update 4. The MathWorks Inc., Natick, Massachusetts, 2019.
- [146] MATLAB Global Optimization Toolbox, 2019. The MathWorks, Natick, MA, USA.
- [147] O. M. Matsiaka, C. J. Penington, R. E. Baker, and M. J. Simpson. Discrete and continuum approximations for collective cell migration in a scratch assay with cell size dynamics. *Bulletin of Mathematical Biology*, 80:738–757, 2018.
- [148] D. L. S. McElwain, R. Callcott, and L. E. Morris. A model of vascular compression in solid tumours. *Journal of Theoretical Biology*, 78:405–415, 1979.
- [149] D. L. S. McElwain and L. E. Morris. Apoptosis as a volume loss mechanism in mathematical models of solid tumor growth. *Mathematical Biosciences*, 39:147–157, 1977.
- [150] G. Mehta, A. Y. Hsiao, M. Ingram, G. D. Luker, and S. Takayama. Opportunities and challenges for use of tumor spheroids as models to test drug delivery and efficacy. *Journal of Controlled Release*, 10:192–204, 2012.
- [151] F.A. Meineke, C. S. Potten, and M. Loeffler. Cell migration and organisation in the intestinal crypt using a lattice-free model. *Cell Proliferation*, 34:253–266, 2001.
- [152] T. Michel, J. Fehrenbach, V. Lobjois, J. Laurent, A. Gomes, T. Colin, and C. Poignard. Mathematical modeling of the proliferation gradient in multicellular tumor spheroids. *Journal of Theoretical Biology*, 458:133–147, 2018.

- [153] A. M. Middleton, C. Fleck, and R. Grima. A continuum approximation to an off-lattice individual-cell based model of cell migration and adhesion. *Journal of Theoretical Biology*, 359:220–232, 2014.
- [154] D. E. Moulton, T. Lessinnes, and A. Goriely. Morphoelastic rods: Part I a single growing elastic rod. *Journal of the Mechanics and Physics of Solids*, 61:398–427, 2013.
- [155] K. C. Murphy, B. P. Hung, S. Browne-Bourne, D. Zhou, J. Yeung, D. C. Genitos, and J. K. Leach. Measurement of oxygen tension within mesenchymal stem cell spheroids. *Journal of the Royal Society Interface*, 14:20160851, 2017.
- [156] R. J. Murphy, A. P. Browning, G. Gunasingh, N. K. Haass, and M. J. Simpson. Designing and interpreting 4D tumour spheroid experiments. *Preprint on bioRxiv*, 2021.
- [157] R. J. Murphy, P. R. Buerzli, R. E. Baker, and M. J. Simpson. A one-dimensional individual-based mechanical model of cell movement in heterogeneous tissues and its coarse-grained approximation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 475:20180838, 2019.
- [158] R. J. Murphy, P. R. Buerzli, R. E. Baker, and M. J. Simpson. Mechanical cell competition in heterogeneous epithelial tissues. *Bulletin of Mathematical Biology*, 82:130, 2020.
- [159] R. J. Murphy, P. R. Buerzli, R. E. Baker, and M. J. Simpson. Travelling waves in a free boundary mechanobiological model of an epithelial tissue. *Applied Mathematics Letters*, 111:106636, 2021.
- [160] R. J. Murphy, P. R. Buerzli, T. A. Tambyah, E. W. Thompson, H. J. Hugo, R. E. Baker, and M. J. Simpson. The role of mechanical interactions in emt. *Physical Biology*, 18:046001, 2021.
- [161] J. D. Murray. *Mathematical Biology I: An Introduction*. Springer-Verlag, New York, USA, 2003.
- [162] J. D. Murray. *Mathematical Biology II: Spatial Models and Biomedical Applications*. Springer-Verlag, New York, USA, 2003.
- [163] P. J. Murray. *From discrete to continuum models of tumour growth*. PhD thesis, Oxford University, UK, 2008.
- [164] P. J. Murray, C. M. Edwards, M. J. Tindall, and P. K. Maini. From a discrete to a continuum model of cell dynamics in one dimension. *Physical Review E*, 80:031912, 2009.
- [165] P. J. Murray, C. M. Edwards, M. J. Tindall, and P. K. Maini. Classifying general non-linear force laws in cell-based models via the continuum limit. *Physical Review E*, 85:021921, 2012.
- [166] P. J. Murray, J. Kang, G. R. Mirams, S. Shin, H. M. Byrne, P. K. Maini, and K. H. Cho. Modelling spatially regulated  $\beta$ -catenin dynamics and invasion in intestinal crypts. *Biophysical Journal*, 99:716–725, 2010.
- [167] P. J. Murray, A. Walter, A. G. Fletcher, C. M. Edwards, M. J. Tindall, and P. K. Maini. Comparing a discrete and continuum model of the intestinal crypt. *Physical Biology*, 8:026011, 2011.
- [168] T. Muthusamy, T. Cordes, M. K. Handzlik, L. You, E. W. Lim, J. Gengatharan, A. F. M. Pinto, M. G. Badur, M. J. Kolar, M. Wallace, A. Saghatelyan, and C. M. Metallo. Serine restriction alters sphingolipid diversity to constrain tumour growth. *Nature*, 586:790–795, 2020.
- [169] S. Nath and G. R. Devi. Three-dimensional culture systems in cancer research: Focus on tumor spheroid model. *Pharmacology & Therapeutics*, 163:94–108, 2016.
- [170] A. V. Nguyen, K. D. Nyberg, M. B. Scott, A. M. Welsh, A. H. Nguyen, N. Wu, S. V. Hohlbauch, N. A. Geisse, E. A. Gibb, A. G. Robertson, T. R. Donahue, and A. C. Rowat. Stiffness of pancreatic cancer cells is associated with increased invasive potential. *Integrative Biology*, 8:1232–1245, 2016.

- [171] D. L. Nikolić, A. N. Boettiger, D. Bar-Sagi, J. D. Carbeck, and S. Y. Shvartsman. Role of boundary conditions in an experimental model of epithelial wound healing. *American Journal of Physiology-Cell Physiology*, 291:C68–C75, 2006.
- [172] S. Nisikawa, A. Takamatsu, S. Ohsawa, and T. Igaki. Mathematical model for cell competition: predator-prey interactions at the interface between two groups of cells in monolayer tissue. *Journal of Theoretical Biology*, 404:40–50, 2016.
- [173] L. Norton, R. Simon, H. D. Brereton, and A. E. Bogden. Predicting the course of Gompertzian growth. *Nature*, 264:542–545, 1976.
- [174] R. D. O’Dea and J. R. King. Continuum limits of pattern formation in hexagonal-cell monolayers. *Journal of Mathematical Biology*, 64:579–610, 2012.
- [175] J. M. Osborne, A. G. Fletcher, J. M. Pitt-Francis, P. K. Maini, and D. J. Gavaghan. Comparing individual-based approaches to modelling the self-organization of multicellular tissues. *PLOS Computational Biology*, 13:e1005387, 2017.
- [176] J. M. Osborne, A. Walter, S. K. Kershaw, G. R. Mirams, A. G. Fletcher, P. Pathmanathan, D. Gavaghan, O. E. Jensen, P. K. Maini, and H. M. Byrne. A hybrid approach to multi-scale modelling of cancer. *Philosophical Transactions of the Royal Society A*, 368:5013–5028, 2010.
- [177] M. Paczkowski, W. W. Kretzschmar, B. Markelc, S. K. Liu, L. A. Kunz-Schughart, A. L. Harris, M. Partridge, H. M. Byrne, and P. Kannan. Reciprocal interactions between tumour cell populations enhance growth and reduce radiation sensitivity in prostate cancer. *Communications Biology*, 4:6, 2021.
- [178] I. Pastushenko and C. Blanpain. EMT transition states during tumor progression and metastasis. *Trends in Cell Biology*, 29:212–226, 2019.
- [179] P. Pathmanathan, J. Cooper, A. G. Fletcher, G. Mirams, P. J. Murray, J. M. Osborne, J. Pitt-Francis, A. Walter, and S. J. Chapman. A computational study of discrete mechanical tissue models. *Physical Biology*, 6:036001, 2009.
- [180] Y. Pawitan. *In all likelihood: statistical modelling and inference using likelihood*. Oxford University Press, Oxford, UK, 2001.
- [181] R. Penta, D. Ambrosi, and R. J. Shipley. Effective governing equations for poroelastic growing media. *The Quarterly Journal of Mechanics and Applied Mathematics*, 67:69–91, 2014.
- [182] M. Plodinec, M. Loparic, C. Monnier, E. Obermann, R. Zanetti-Dallenbach, P. Oertle, J. Hyotyla, U. Aebi, M. Bentires-Alj, R. Lim, and C. Schoenberger. The nanomechanical signature of breast cancer. *Nature Nanotechnology*, 7:757–764, 2012.
- [183] M. Poujade, E. Grasland-Mongrain, A. Hertzog, J. Jouanneau, P. Chavrier, B. Ladoux, A. Buguin, and P. Silberzan. Collective migration of an epithelial monolayer in response to a model wound. *Proceedings of the National Academy of Sciences of the United States of America*, 104:15988–15993, 2007.
- [184] K. Powell. These secret battles between your body’s cells might just save your life. *Nature*, 574:310–312, 2019.
- [185] A. Puliafito, L. Hufnagel, P. Neveu, S. Streichan, D. Sigal, A. Kuchnir Fygenson, and B. I. Shraiman. Collective and single cell behavior in epithelial contact inhibition. *Proceedings of the National Academy of Sciences of the United States of America*, 109:739–744, 2012.

- [186] V. Rajagopal, W. R. Holmes, and P. V. S. Lee. Computational modeling of single-cell mechanics an cytoskeletal mechanobiology. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 10:5–7, 2018.
- [187] 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. *Biophysical Journal*, 95:155–165, 2008.
- [188] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25:1923–1929, 2009.
- [189] P. Recho, J. Ranft, and P. Marcq. One-dimensional collective migration of a proliferating cell monolayer. *Soft Matter*, 8:2381–2391, 2016.
- [190] S. Redner. *A guide to first-passage processes*. Cambridge University Press, Cambridge, UK, 2001.
- [191] T. Roose, S. J. Chapman, and P. K. Maini. Mathematical models of avascular tumor growth. *SIAM Review*, 49:179–208, 2007.
- [192] A. Sakaue-Sawano, H. Kurokawa, T. Morimura, A. Hanyu, H. Hama, H. Osawa, S. Kashiwagi, K. Fukami, T. Miyata, H. Miyoshi, T. Imamura, M. Ogawa, H. Masai, and A. Miyawaki. Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell*, 132:487–498, 2008.
- [193] M. S. Samuel, J. Lopez, E. McGhee, D. Croft, D. Strachan, P. Timpson, J. Munro, E. Schroder, J. Zhou, V. Brunton, N. Baker, H. Clevers, O. Sansom, K. Anderson, V. Weaver, and M. Olson. Actomyosin-mediated cellular tension drives increased tissue stiffness and  $\beta$ -catenin activation to induce epidermal hyperplasia and tumor growth. *Cancer Cell*, 19:776–791, 2011.
- [194] F. Sanchez-Garduno and P. K. Maini. An approximation to a sharp type solution of a density-dependent reaction-diffusion equation. *Applied Mathematics Letters*, 7:47–51, 1994.
- [195] F. Sanchez-Garduno and P. K. Maini. Travelling wave phenomena in some degenerate reaction-diffusion equations. *Journal of Differential Equations*, 117:281–389, 1995.
- [196] S. A. Sandersius, C. J. Weijer, and T. J. Newman. Emergent cell and tissue dynamics from subcellular modeling of active biomechanical processes. *Physical Biology*, 8:045007, 2011.
- [197] M . T. Santini and G. Rainaldi. Three-dimensional spheroid model in tumor biology. *Pathobiology*, 67:148–157, 1999.
- [198] X. Serra-Picamal, V. Conte, R. Vincent, E. Anon, D. T. Tambe, E. Bazellieres, J. P. Butler, J. J. Fredberg, and X. Trepat. Mechanical waves during tissue expansion. *Nature Physics*, 8:628–634, 2012.
- [199] L. F. Shampine and M. W. Reichelt. The matlab ode suite. *SIAM Journal on Scientific Computing*, 18:1–22, 1997.
- [200] J. A. Sherratt and B. P. Marchant. Nonsharp travelling wave fronts in the Fisher equation with degenerate nonlinear diffusion. *Applied Mathematics Letters*, 9:33–38, 1996.
- [201] J. A. Sherratt and J. D. Murray. Models of epidermal wound healing. *Proceedings of the Royal Society B: Biological Sciences*, 241:29–36, 1990.
- [202] B. I. Shraiman. Mechanical feedback as a possible regulator of tissue growth. *Proceedings of the National Academy of Sciences of the United States of America*, 102:3318–3323, 2005.

- [203] M. J. Simpson and R. E. Baker. Exact calculations of survival probability for diffusion on growing lines, disks, and spheres: the role of dimension. *The Journal of Chemical Physics*, 143:094109, 2015.
- [204] M. J. Simpson, R. E. Baker, S. T. Vittadello, and O. J. Maclaren. Practical parameter identifiability for spatio-temporal models of cell invasion. *Journal of the Royal Society Interface*, 17:2020055, 2020.
- [205] M. J. Simpson, A. P. Browning, C. Drovandi, E. J. Carr, O. J. Maclaren, and R. E. Baker. Profile likelihood analysis for a stochastic model of diffusion in heterogeneous media. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 477:20210214, 2021.
- [206] M. J. Simpson, A. P. Browning, D. J. Warne, O. J. Maclaren, and R. E. Baker. Parameter identifiability and model selection for sigmoid population growth models. *Preprint on bioRxiv*, 2021.
- [207] M. J. Simpson, K. A. Landman, and D. F. Newgreen. Chemotactic and diffusive migration on a nonuniformly growing domain: numerical algorithm development and applications. *Journal of Computational and Applied Mathematics*, 192:282–300, 2006.
- [208] M. J. Simpson, J. A. Sharp, and R. E. Baker. Survival probability for a diffusive process on a growing domain. *Physical Review E*, 91:042701, 2014.
- [209] M. J. Simpson, J. A. Sharp, L. C. Morrow, and R. E. Baker. Exact solutions of coupled multispecies linear reaction–diffusion equations on a uniformly growing domain. *PLOS One*, 10, 2015.
- [210] P. Simpson. Parameters of cell competition in the compartments of the wing disc of *Drosophila*. *Developmental Biology*, 69:182–193, 1979.
- [211] P. Simpson and G. Morata. Differential mitotic rates and patterns of growth in compartments in the *Drosophila* wing. *Developmental Biology*, 85:299–308, 1981.
- [212] J. G. Skellam. Random dispersal in theoretical populations. *Biometrika*, 38:196–218, 1951.
- [213] K. S. M. Smalley, R. Contractor, N. K. Haass, A. N. Kulp, G. E. Atilla-Gokcumen, D. S. Williams, H. Bregman, K. T. Flaherty, M. S. Soengas, E. Meggers, and M. Herlyn. An organometallic protein kinase inhibitor pharmacologically activates p53 and induces apoptosis in human melanoma cells. *Cancer Research*, 67:209–217, 2007.
- [214] K. S. M. Smalley, R. Contractor, N. K. Haass, J. T. Lee, K. L. Nathanson, C. A. Medina, K. T. Flaherty, and M. Herlyn. Ki67 expression levels are a better marker of reduced melanoma growth following MEK inhibitor treatment than phospho-ERK levels. *British Journal of Cancer*, 96:445–449, 2007.
- [215] K. S. M. Smalley, M. Lioni, K. Noma, N. K. Haass, and M. Herlyn. *In vitro* three-dimensional tumour microenvironment models for anticancer drug discovery. *Expert Opinion on Drug Discovery*, 3:1–10, 2008.
- [216] C. A. Smith and C. A. Yates. Spatially extended hybrid methods: a review. *Journal of the Royal Society Interface*, 15:20170931, 2018.
- [217] L. Spoerri, K. A. Beaumont, A. Anfosso, and N. K. Haass. Real-time cell cycle imaging in a 3D cell culture model of melanoma. *Methods in Molecular Biology*, 1612:401–416, 2017.
- [218] L. Spoerri, G. Gunasingh, and N. K. Haass. Fluorescence-based quantitative and spatial analysis of tumour spheroids: a proposed tool to predict patient-specific therapy response. *Frontiers in Digital Health*, 3:668390, 2021.
- [219] L. Spoerri, C. A. Tonnessen-Murray, G. Gunasingh, D. S. Hill, K. A. Beaumont, R. J. Jurek, J. Chauhan, G. C. Vanwalleghem, M. E. Fane, S. M. Daignault-Mill, N. Matigian, G. M. Boyle, E. K. Scott, A. G. Smith, S. J.

- Stehbens, H. Schaider, B. Gabrielli, W. Weninger, C. R. Goding, and N. K. Haass. Phenotypic melanoma heterogeneity is regulated through cell-matrix interaction-dependent changes in tumor microarchitecture. *Preprint on bioRxiv*, 2021.
- [220] A. Surendran, M. J. Plank, and M. J. Simpson. Population dynamics with spatial structure and an allee effect. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 476:20200501, 2020.
- [221] S. Suresh. Elastic clues in cancer detection. *Nature Nanotechnology*, 2:748–749, 2007.
- [222] E. A. Susaki, C. Shimizu, A. Kuno, K. Tainaka, X. Li, K. Nishi, K. Morishima, H. Ono, K. L. Ode, Y. Saeki, K. Miyamichi, K. Isa, C. Yokoyama, H. Kitaura, M. Ikemura, T. Ushiku, Y. Shimizu, T. Saito, T. C. Saido, M. Fukayama, H. Onoe, K. Touhara, T. Isa, A. Kakita, M. Shibayama, and H. R. Ueda. Versatile whole-organ/body staining and imaging based on electrolyte-gel properties of biological tissues. *Nature Communications*, 11:1982, 2020.
- [223] K. Takahashi, R. Tanabe, S. Ehata, S. I. Kubota, Y. Morishita, H. R. Ueda, and K. Miyazono. Visualization of the cancer cell cycle by tissue-clearing technology using the Fucci reporter system. *Cancer Science*, 0:1–14, 2021.
- [224] D. T. Tambe, C. C. Hardin, T. E. Angelini, K. Rajendran, C. Y. Park, X. Serra-Picamal, E. H. Zhou, M. H. Zaman, J. P. Butler, D. A. Weitz, J. J. Fredberg, and X. Trepat. Collective cell guidance by cooperative intercellular forces. *Nature Materials*, 10:469–475, 2011.
- [225] T. A. Tambyah, R. J. Murphy, P. R. Buenzli, and M. J. Simpson. A free boundary mechanobiological model of epithelial tissues. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 476:20200528, 2020.
- [226] X. Tian, H. Zhang, and J. Xing. Coupled reversible and irreversible bistable switches underlying  $TGF\beta$ -induced epithelial to mesenchymal transition. *Biophysical Journal*, 105:1079–1089, 2013.
- [227] X. Trepat and E. Sahai. Mesoscale physical principles of collective cell organisation. *Nature Physics*, 14:671–682, 2018.
- [228] X. Trepat, M. R. Wasserman, T. E. Angelini, E. Millet, D. A. Weitz, J. P. Butler, and J. J. Fredberg. Physical forces during collective cell migration. *Nature Physics*, 5:426–430, 2009.
- [229] S. Tripathi, H. Levine, and M. K. Jolly. The physics of cellular decision making during epithelial–mesenchymal transition. *Annual Review of Biophysics*, 49:1–18, 2020.
- [230] T. N. TruongVo, R. M. Kennedy, H. Chen, A. Chen, A. Berndt, M. Agarwal, L. Zhu, H. Nakshatri, J. Wallace, S. Na, H. Yokota, and J. E. Ryu. Microfluidic channel for characterizing normal and breast cancer cells. *Journal of Micromechanics and Microengineering*, 27:035017, 2017.
- [231] J. M. Tse, G. Cheng, J. A. Tyrrell, S. A. Wilcox-Adelman, Y. Boucher, R. K. Jain, and L. L. Munn. Mechanical compression drives cancer cells toward invasive phenotype. *Proceedings of the National Academy of Sciences of the United States of America*, 109:911–916, 2012.
- [232] A. Tsuboi, S. Ohsawa, D. Umetsu, Y. Sando, E. Kuranaga, T. Igaki, and K. Fujimoto. Competition for space is controlled by apoptosis-induced change of local epithelial topology. *Current Biology*, 28:2115–2128, 2018.
- [233] D. Tuveson and H. Clevers. Cancer modeling meets human organoid technology. *Science*, 364:952–955, 2019.

- [234] M. Uroz, S. Wistorf, X. Serra-Picamal, V. Conte, M. Sales-Pardo, P. Roca-Cusachs, R. Guimerá, and X. Trepat. Regulation of cell cycle progression by cell–cell and cell–matrix forces. *Nature Cell Biology*, 20:646–654, 2018.
- [235] P. Van Meurs and M. Marondotti. Discrete-to-continuum limits of particles with an annihilation rule. *SIAM Journal on Applied Mathematics*, 79:1940–1966, 2019.
- [236] V. Vavourakis, P. A. Wijeratne, R. Shipley, M. Loizidou, T. Stylianopoulos, and D. J. Hawkes. A validated multiscale in-silico model for mechano-sensitive tumour angiogenesis and growth. *PLoS Computational Biology*, 13:e1005259, 2017.
- [237] S. T. Vittadello, S. W. McCue, G. Gunasingh, N. K. Haass, and M. J. Simpson. Mathematical models for cell migration with real-time cell cycle dynamics. *Biophysical Journal*, 114:1241–1253, 2018.
- [238] S. T. Vittadello, S. W. McCue, G. Gunasingh, N. K. Haass, and M. J. Simpson. Mathematical models incorporating a multi-stage cell cycle replicate normally-hidden inherent synchronisation in cell proliferation. *Journal of the Royal Society Interface*, 16:20190382, 2019.
- [239] S. T. Vittadello, S. W. McCue, G. Gunasingh, N. K. Haass, and M. J. Simpson. A novel mathematical model of heterogeneous cell proliferation. *Journal of Mathematical Biology*, 82:34, 2021.
- [240] M. Voliotis, P. Thomas, R. Grima, and C. G. Bowsher. Stochastic simulation of biomolecular networks in dynamic environments. *PLOS Computational Biology*, 6:e1004923, 2016.
- [241] L. Wagstaff, M. Goschorksa, K. Kozyrksa, G. Duclos, I. Kucinski, A. Chessel, L. Hampton-O’Neil, C. R. Bradshaw, G. E. Allen, E. L. Rawlins, P. Silberzan, R. E. Carazo-Salas, and E. Piddini. Mechanical cell competition kills cells via induction of lethal p53 levels. *Nature Communications*, 7:11373, 2016.
- [242] J. P. Ward and J. R. King. Mathematical modelling of avascular-tumour growth. *IMA Journal of Mathematics Applied in Medicine and Biology*, 14:39–69, 1997.
- [243] J. P. Ward and J. R. King. Mathematical modelling of avascular-tumour growth II: modelling growth saturation. *IMA Journal of Mathematics Applied in Medicine and Biology*, 16:171–211, 1999.
- [244] D. J. Warne, R. E. Baker, and M. J. Simpson. Using experimental data and information criteria to guide model selection for reaction–diffusion problems in mathematical biology. *Bulletin of Mathematical Biology*, 81:1760–1804, 2019.
- [245] L. Weiswald, D. Bellet, and V. Dangles-Marie. Spherical cancer models in tumor biology. *Neoplasia*, 17:1–15, 2015.
- [246] F. Wieland, A. L. Hauber, M. Rosenblatt, C. Tonsing, and J. Timmer. On structural and practical identifiability. *Current Opinion in Systems Biology*, 25:60–69, 2021.
- [247] T. P. Witelski. Merging traveling waves for the Porous-Fisher’s equation. *Applied Mathematics Letters*, 8:57–62, 1995.
- [248] T. Wyatt, B. Baum, and G. Charras. A question of time: tissue adaptation to mechanical forces. *Current Opinion in Cell Biology*, 38:68–73, 2016.
- [249] J. Yang, P. Antin, G. Berx, C. Blanpain, T. Brabletz, M. Bronner, K. Campbell, A. Cano, J. Casanova, G. Christofori, et al. Guidelines and definitions for research on epithelial-mesenchymal transition. *Nature Reviews Molecular Cell Biology*, 21:341–352, 2020.

- [250] S. Yano, S. Miwa, Y. Mii, S. anmd Hiroshima, F. Uehara, M. Yamammota, H. Kishimoto, H. Tazawa, M. Bouvet, T. Fujiwara, and R. M. Hoffman. Invading cancer cells are predominantly in G0/G1 resulting in chemoresistance demonstrated by real-time FUCCI imaging. *Cell Cycle*, 13:953–960, 2014.
- [251] S. Yano, H. Tazawa, S. Kagawa, T. Fujiwara, and R. M. Hoffman. FUCCI real-time cell-cycle imaging as a guide for designing improved cancer therapy: A review of innovative strategies to target quiescent chemoresistant cancer cells. *Cancers*, 12:2655, 2020.
- [252] C. A. Yates, R. E. Baker, R. Erban, and P. K. Maini. Going from microscopic to macroscopic on nonuniform growing domains. *Physical Review E*, 86:021921, 2012.
- [253] M. A. Yereniuk and S. D. Olson. Global density analysis for an off-lattice agent-based model. *SIAM Journal on Applied Mathematics*, 79:1700–1721, 2019.
- [254] C. Zheng and G. Bennett. *Applied Contaminant Transport Modeling*. Wiley-Blackwell, New York, NY, 2002.
- [255] N. Zielke and B. A. Edgar. FUCCI sensors: powerful new tools for analysis of cell proliferation. *WIREs Developmental Biology*, 4:469–487, 2015.
- [256] C. Zmurchok, D. Bhaskar, and L. Edelstein-Keshet. Coupling mechanical tension and GTPase signaling to generate cell and tissue dynamics. *Physical Biology*, 15:046004, 2018.
- [257] C. Zmurchok and W. R. Holmes. Simple Rho GTPase dynamics generate a complex regulatory landscape associated with cell shape. *Biophysical Journal*, 118:1438–1454, 2020.