This is a repository copy of *Models, measurement and inference in epithelial tissue dynamics*.

Article:

MODELS, MEASUREMENT AND INFERENCE IN EPITHELIAL TISSUE DYNAMICS

Oliver J. Maclaren

Wolfson Centre for Mathematical Biology, Mathematical Institute, University of Oxford,
Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK

Helen M. Byrne

Department of Computer Science, University of Oxford,
Wolfson Building, Parks Road, Oxford OX1 3QD, UK

Alexander G. Fletcher and Philip K. Maini

Wolfson Centre for Mathematical Biology, Mathematical Institute, University of Oxford,
Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK

Abstract. The majority of solid tumours arise in epithelia and therefore much research effort has gone into investigating the growth, renewal and regulation of these tissues. Here we review different mathematical and computational approaches that have been used to model epithelia. We compare different models and describe future challenges that need to be overcome in order to fully exploit new data which present, for the first time, the real possibility for detailed model validation and comparison.

1. **Introduction.** There are many exciting challenges facing mathematical modelling in biology. In this article we review some of these challenges in the context of the healthy and diseased dynamics of a specific biological system, the epithelial tissues that line internal and external body surfaces. Much of the interest in epithelial tissues arises because almost all solid cancers are epithelial in origin [84]; however, many of the modelling issues we discuss are relevant to a wider range of problems in mathematical biology.

The challenges we will highlight are: (I) *Scale and resolution of models*. The desire for a model to address a specific biological question often motivates model simplicity; however, it is now well-known that biological function arises from the integration of coupled processes acting across a range of spatial and temporal scales. Often these processes cannot be easily separated and this creates technical challenges for mathematical and computational modelling. (II) *Model-model comparison*. Currently, many different models have been developed to address the same problem, leading us to ask under what conditions do the different models generate the same predictions? Further, when models disagree, which approach is better? Questions of this nature motivate the development of new methods of model-model comparison. (III) *Model-data comparison*. Validation of multiscale models is typically performed in a coarse-grained manner. With the advent of improved imaging

---

techniques, there is now the potential to compare such models with high-resolution experimental data. Modern statistical inference methods, made available with increased computing power [3], have yet to be fully exploited, however, and offer exciting opportunities for the future.

   The remainder of the paper is structured as follows. Section 2 contains a brief biological introduction to epithelial tissues. We take the view that many questions regarding epithelial pathology such as tumorigenesis can be profitably understood in the context of failure of normal homeostasis and repair mechanisms. Thus we consider a range of phenomena in epithelial tissue dynamics. In Sections 3 and 4, respectively, we review compartmental and individual-based approaches that have been used to model epithelial cell populations, emphasising how the scale and resolution of the models (challenge I) have yielded insights into epithelial dynamics, and how some organising concepts have emerged within these model classes. At the same time, the absence of a common structure linking the different compartmental and cell-based models makes their comparison (challenge II) difficult. This motivates our discussion in Section 5 of continuum limits of discrete, cell-based models, considered in the context of mechanically-interacting epithelial cells. In the discussion in the final section, Section 6, we focus on the opportunities that new data raise for measurement and parameter estimation (challenge III), referencing examples from tumour dynamics and epithelial development and repair, while emphasising the intertwined nature of the three challenges I-III. We also speculate on possible approaches for developing unified frameworks for using modelling, measurement and inference to better understand epithelial tissue dynamics.

2. Epithelial biology.

2.1. **Motivation.** The epithelial tissues that line the internal and external surfaces of the body are a primary source of physical protection and act as barriers to unwarranted substances [47]. Maintenance of normal physiological function requires that many epithelia simultaneously carry out directed transport of substances such as water and electrolytes between bathing solutions on either side of the epithelial tissue; this is crucial for tasks such as regulating body fluid composition and absorbing nutrients [69]. As a result, many epithelial tissues, such as the skin and intestinal epithelia, have rapid rates of self-renewal and distinctive structural and functional organisation (see below). Importantly, dysregulation of the renewal process is thought to be associated with the large number (up to 90%) of cancers originating in epithelia [67, 84]. Epithelial tissues thus serve as important biological model systems for studying the dynamics of multicellular populations; further, understanding them under both healthy and diseased conditions is important for understanding the transition from one state to the other.

2.2. **General structure and function.** Epithelial tissues comprise sheets of polarised cells, as well as glandular ingrowths (or crypts) and finger-like protrusions (or villi) formed from these sheets; the arrangement of these layers is maintained by specialised cell-cell and cell-matrix junctions [69] (Figure 1A). These junctions, especially the tight junctions, are crucial not only for structural integrity of the epithelial layers, but also for maintaining cell polarity and, hence, the ability of epithelial tissues to perform directed transport of solutes and water. Individual epithelial cells are subject to strong mutual mechanical interactions mediated through these proteins.

The self-renewal of epithelial tissues is driven by stem cells which generate, in a tightly-controlled manner, a sequence of increasingly specialised descendants known collectively as a lineage structure (Figure 1B). This sequence must produce a sufficient number of descendants to maintain those parts of the epithelium exposed to the greatest levels of damage and so the spatial arrangement of the cell lineages frequently corresponds to functional requirements. An archetypal example is provided by intestinal crypts (Figure 1C), finger-like invaginations of the intestinal mucosa. The intestinal crypt is structured spatially into regions corresponding to a stem cell niche at the base, an intermediate region of rapidly proliferating transit-amplifying cells, and a region of differentiated, functional cells nearest to the crypt orifice, where most transport and damage occur [84]. This lineage structure maximises the number of functional descendants of stem cells while protecting the stem cells from damage. In addition to extensive experimental study, the intestinal epithelium has been the focus of numerous mathematical models, especially in the context of colorectal carcinogenesis [10, 14, 18, 38, 41]. In the sections that follow we will review some of these models.

3. Compartmental models.

3.1. Overview.
Compartmental models provide a flexible framework for studying cell population dynamics. As a result, they have been widely used to study epithelial tissue homeostasis, growth, regeneration and carcinogenesis, and the interplay between these dynamics. These models account for processes such as cell proliferation, differentiation and death in generic terms, without specifying details of the mechanisms (e.g. signalling pathways) by which they are regulated.

The resolution of these models is coarse; compartments typically correspond to either spatial domains or distinct cell populations, with movement or other exchange between compartments. Using too many compartments is not typically a good idea - it can lead to parameters that are redundant or poorly constrained. In this case one might seek to reduce the number of compartments or, alternatively, use partial differential equation (PDE) or delay-differential equation (DDE) models, which replace the discrete parameters with some constrained (e.g. smooth) function or 'lumping' of parameters. In what follows we restrict attention to deterministic compartment models comprising systems of ordinary differential equations (ODEs). Further details on more complex compartment models, including those that account for age-structure and stochastic effects, may be found in [10] and references therein. Later sections address alternative models for introducing finer-grained spatial resolution.

3.2. Framework.
The general form of an ODE-based compartment model for a sequence of n cell lineages (see Figure 1B), which incorporates the processes of cell

(A)

(B)

(C)
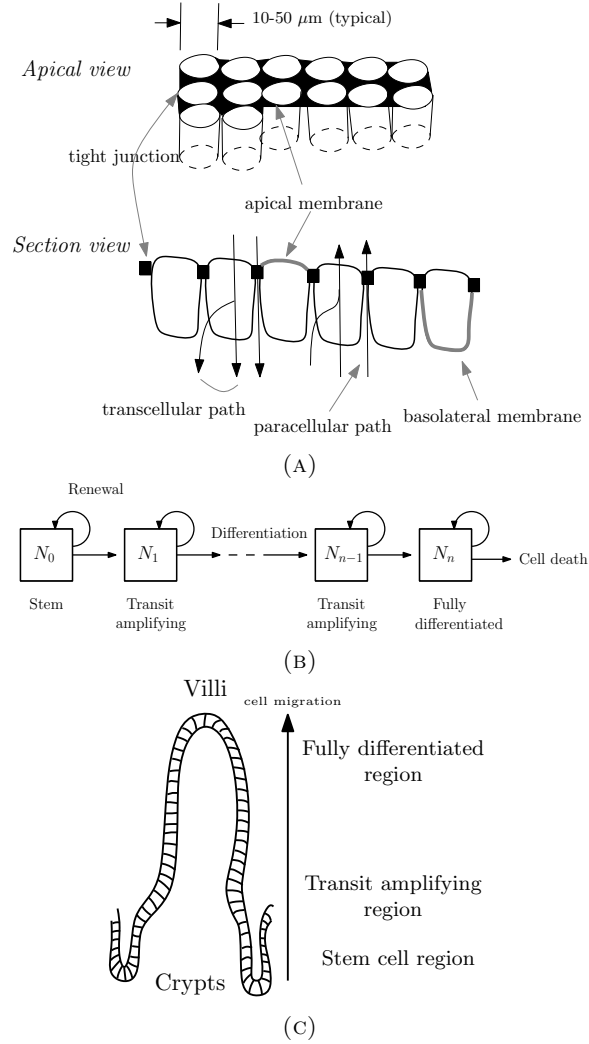
FIGURE 1. A: Schematic of the polarised structure of epithelial
sheets, adapted from [22]. B: Compartment description of the lin-
eage structure of epithelial cell populations. C: Schematic of the
structure of the small intestine, showing crypt invaginations where
stem and transit amplifying cells are located and villus protrusions,
responsible to the absorption of nutrients and water. Based on [8].

renewal, differentiation and death, is given by

$$\frac{dN_0}{dt} = (\alpha_0 - \beta_0 - \gamma_0)N_0, \tag{1}$$

$$\frac{dN_1}{dt} = (\alpha_1 - \beta_1 - \gamma_1)N_1 + \beta_0 N_0, \tag{2}$$

$$\vdots$$

$$\frac{dN_{n-1}}{dt} = (\alpha_{n-1} - \beta_{n-1} - \gamma_{n-1})N_{n-1} + \beta_{n-2}N_{n-2}, \tag{3}$$

$$\frac{dN_n}{dt} = -\gamma_n N_n + \beta_{n-1}N_{n-1}. \tag{4}$$

Here $N_i(t)$ denotes the number of cells in compartment i at time t, while the parameters α_i, β_i and γ denote the renewal, differentiation and death rates of cells in compartment i. These parameters may be constants (for a given biological scenario) or may be prescribed functions of the different cell populations. In the latter case, the functional forms encode key hypotheses about how signalling and other control mechanisms regulate cell kinetics; we consider some examples of this next.

3.3. **Applications to homeostasis, control and tumorigenesis.** Johnston et al. [39] construct a model of the form (1)–(4) to investigate homeostasis and the transition to tumorigenesis in the intestinal crypt. They consider three compartments, corresponding to stem, transit amplifying and differentiated cells respectively. They note that models in which the kinetic parameters α_i, β_i and γ_i are constants are *structurally unstable*: slightly different parameter values can lead to qualitatively different population dynamics. In their model, the total number of crypt cells may exhibit exponential growth, exponential decay or converge to a finite equilibrium, according to whether the net growth rate in the stem cell compartment, $\alpha = \alpha_0 - \beta_0 - \gamma_0$, is greater than, less than or equal to zero, respectively. Such behaviour is physically unrealistic, not least because the model parameters are idealisations, which are unlikely to remain constant during tissue homeostasis.

To obtain a model in which homeostasis is preserved as system parameters vary, Johnston et al. [39] incorporate local feedbacks within the stem and transit amplifying compartments, such that the proportion of cells differentiating in a given compartment varies with the compartment size. They find that a feedback which increases linearly with compartment population size N, taking the form $\alpha(N) = a + bN$ for constants a and b, leads to a model for which homeostasis is a stable steady state. While the total number of crypt cells at equilibrium changes as parameter values vary, representing the effect of pre-cancerous mutations on cell kinetics, the system does not exhibit the unbounded growth that is characteristic of tumorigenesis. Improving on this, the authors find that including a saturating local feedback of the form $\alpha^*(N) = a^* + b^* N/(1 + c^* N)$, for constants a^*, b^* and c^*, leads to a model that, as parameter values are varied to represent the effects of mutations, exhibits both changes to equilibrium cell numbers as well as eventual unbounded growth after sufficient mutations accumulate. The authors note that this behaviour recapitulates the natural pattern of tumorigenesis, where periods of rapid expansion alternate with long periods of equilibrium in a stepwise manner.

In other work, Lander et al. [42] (previously analysed in more generality in [44]) develop an ODE-based compartment model to study the renewal of mammalian olfactory neural epithelium. In contrast to the model of Johnston et al. [39], this model allows for feedback mechanisms whereby cells that are further along the lineage modulate the behaviour of cells earlier in the lineage, this feedback being mediated by chemicals sometimes given the generic name 'chalones'.

Lander et al. [42] consider the cases of two, three and an arbitrary number of compartments, and assume that cell death is restricted to the most differentiated compartment only: in all other compartments, dividing cells either renew or differentiate. Lander et al. denote by $0 \leq p \leq 1$ the renewal probability and by $0 \leq v$ the rate of cell division, and so we may identify the link between our notation and theirs through $\alpha - \beta = (2p - 1)v$.

Lander et al. [42] use an engineering approach to analyse their model. They consider a variety of control 'objectives' such as control of size, growth rate and cell-type proportions. In agreement with Johnston et al. [39], they observe that

without feedback the system is sensitive to variations in parameter values. They consider several feedback mechanisms by which chalones may achieve the control objectives. These include feedback on renewal probabilities, division rates, and the compartments between which the feedbacks operate. Feedback on renewal probabilities, rather than division rates, is found to give the most robust form of control, with reduced parameter sensitivity, faster regeneration in response to injury and less sensitive dependence on initial conditions.

The work of Lander et al. [42] illustrates how even simple theoretical analyses can generate new biological insight when carefully combined with experiments. The authors progress from a model prediction, which was unexpected without formal analysis, to experimental validation by considering the protein growth differentiation factor 11 (GDF11), a member of the transforming growth factor beta superfamily. While already known as an inhibitor of the mammalian olfactory neural epithelium, GDF11's primary target was believed to be the rate of cell division rather than the renewal probability. Lander et al. show experimentally that GDF11 appears negatively to regulate both rates, and conclude that many experimental observations are consistent with the theoretical expectations of 'best' control.

The generic nature of compartment models is both a strength and a weakness. Since they are formulated in phenomenological terms, they can be used to determine what types of feedback mechanisms will produce robust dynamic behaviour when model parameters fluctuate, and how these might fail. However, it can be difficult to relate these general features to specific measurable quantities and/or to rule out alternative mechanisms. Furthermore, while capable of representing coarse spatial structure, such as cell lineages in the intestinal crypt, compartmental models lack the spatial resolution needed to track cell movement or to account for mechanical effects that regulate such movement. This can have important implications for understanding both healthy and diseased state dynamics. We next consider a modelling approach that can address such issues.

4. Individual-based models.

4.1. Overview.
In order to account for phenomena such as cell sorting, rearrangements, mechanical interactions and pattern formation, mathematical models with spatial resolution are often required. Individual-based models (IBMs) offer a natural framework for describing the behaviour of a finite number of individual cells or cellular components. These models are typically termed either *'on lattice'* or *'off lattice'*, depending on how individual cells are represented (Figure 2). In the following we focus on general features of IBMs that have been used to study the dynamics of cell sorting in embryonic epithelia and the influence of mechanical forces on the dynamics of intestinal crypt cells: for more extensive reviews, see [14] and [41].

4.2. Applications to carcinogenesis, cell sorting and intestinal crypt dynamics.

On-lattice models. On-lattice models usually discretise space using a fixed, regular, lattice. For example, in cellular automata (CA), each cell occupies a single lattice site, which may change over time [16]. The system dynamics are encoded in a simple, rule-based manner, with changes in cell configuration occurring at discrete time steps, using synchronous or asynchronous updating or an event-driven approach based on the Gillespie algorithm [6, 43]. CA have no *a priori* requirement
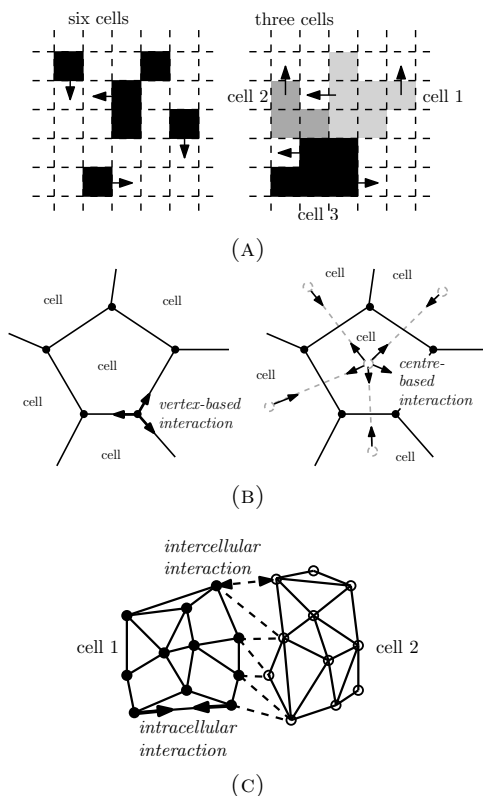
(A)

(B)

(C)

FIGURE 2. A: Schematic of a cellular automaton model (left), where each shaded lattice site corresponds to one cell, and a cellular Potts model (right), where each collection of similarly coloured lattice sites represents one distinct biological cell. B: Schematic of an off-lattice, vertex-based model (left), where the solid points are the cell vertices mediating the mechanical forces (indicated by the black arrows), and an off-lattice, cell centre-based model (right), where mechanical interactions (indicated by the black arrows) apply between cell centres. C. Schematic of a subcellular element model with two cells, each represented by a collection of similarly coloured vertices called 'subcellular elements', and where there are both intracellular (solid arrows) and intercellular (dashed arrows) interactions between elements. Images based on those in [41], [64] and [61].

to satisfy physical principles and hence may exhibit unphysical dynamics. For example, when a cell divides in a CA, its daughter cells may 'shove' their neighbours to make space, resulting in instantaneous displacements over many cell diameters. Model simulations may also suffer from artificial anisotropies induced by symmetries of the underlying lattice. Despite their relative simplicity, CA models have been successfully used to predict aspects of epithelial carcinogenesis. For example, Gatenby et al. [24] combined CA simulations with histological studies of tumour biopsies to demonstrate the key role that adaptation to hypoxia and acidosis may

play in the transition from a precancerous lesion (such as ductal carcinoma in situ in the breast) to invasive cancer.

The cellular Potts model (CPM) is another popular, on-lattice-model, based on simple models of magnetism from statistical mechanics, and adapted by Graner and Grazier [29, 27, 28] to study cell sorting in embryonic epithelia. In the CPM, cells may span several, connected lattice sites, rather than being in one-to-one correspondence as for the CA models. If lattice sites are indexed by a single integer i, then cells correspond to collections of lattice sites which share the same value of an integer label $\sigma(i)$. Using $\mathcal{N}$ to denote the set of all pairs of neighbouring lattice sites (i, j), the possible interactions between neighbouring cells are encoded via a Hamiltonian (interaction energy), such as [29, 27]

$$H = \sum_{(i,j) \in \mathcal{N}} \left(1 - \delta_{\sigma(i)\sigma(j)}\right) J(\tau(\sigma(i)), \tau(\sigma(j))) + \lambda_A \sum_{\sigma} (A(\sigma) - A_T(\tau(\sigma)))^2, \quad (5)$$

where the additional label τ defines cell type and δ_{mn} is the Kronecker delta. The first summation in equation (5) represents an adhesion interaction energy and the second term an energy associated with cell stretching. In (5), different adhesion energies $J(\tau(\sigma(i)), \tau(\sigma(j)))$ may be assigned to interactions involving different types of cells, but individual cells do not have adhesive interactions with themselves. The area of a cell (assuming we are modelling cells in two dimensions) is denoted by $A = A(\sigma)$, $A_T(\tau(\sigma))$ is its target area and λ_A is an tunable phenomenological parameter weighting the contribution of the target area energy to the adhesive energy. Additional interaction or cell property terms may be added as desired. The model dynamics are implemented via a Markov chain Monte Carlo (MCMC) updating method (the Metropolis algorithm) rather than an explicit equation of motion [29, 27]. We note that the original statistical mechanical models that the CPM is based on were developed to study equilibrium systems and there is some difficulty in interpreting the dynamic features of the CPM (see e.g. the discussion in [82]).

As mentioned, the CPM was introduced by Glazier and Graner [27, 29] to explain experiments performed on embryonic epithelial cell populations. After artificial mixing, the different types of cells rearrange until they recover their pre-mixed configurations. Steinberg [73, 74, 75, 76] introduced the differential adhesion hypothesis (DAH) to explain this phenomenon, the general principle being that intermixed cell populations rearrange so as to minimise their collective adhesive free energy. The CPM, as implemented by Graner and Glazier [27, 29], successfully, albeit qualitatively, reproduces observations of cell sorting [27, 29] based on the DAH. This concept has subsequently been applied to processes in development, wound repair, and tumorigenesis; for example, [80] used a CPM to investigate the influence of changes in cell-cell and cell-substrate adhesion on cancer invasion. Their model notably incorporates a description of cell proliferation in which the mitotic rate depends on the relative magnitudes of homotypic and heterotypic adhesion. Despite difficulties in precisely relating some aspects of this model framework to biological processes, the CPM remains an informative and popular framework [14, 28, 65].

Off-lattice models. Off-lattice IBMs describe the dynamics of discrete, mechanically-interacting cell entities. At their simplest these entities are typically either cell-centres or cell vertices, while more complex models allow for additional resolution of the cell interior in terms of further 'subcellular'-scale entities. The basic governing principles are the same in all cases, however. As for the CPM, these models can be

formulated in terms of potential functions encoding the interaction energy between cells or cellular elements. In contrast to the CPM, here an explicit equation of motion is specified for the cellular entities.

For vertex models we consider cells indexed by $\sigma = 1, ..., N$, as in the CPM, but now with vertices, rather than grid points, indexed by $i = 1, ..., V$ and with each vertex having an associated position vector $\mathbf{r}_i(t)$. Here cell vertices represent the tight junction proteins connecting cells. The free energy of each cell is then defined as a sum of contributions, such as adhesive energy U^J, target area (assuming for convenience that we are in two-dimensions) energy U^A and target perimeter energy U^P, giving [64, 20]

$$U = \sum_{\sigma} \left(U_\sigma^J + U_\sigma^A + U_\sigma^P \right) \tag{6}$$

$$= \sum_{\sigma} \left(\lambda_J \sum_{L \in \mathcal{L}(\sigma)} L^2 + \lambda_A (A(\sigma) - A_T(\sigma))^2 + \lambda_P (P(\sigma) - P_T(\sigma))^2 \right). \tag{7}$$

Here $\mathcal{L}(\sigma)$ is the set of lengths L of the lines, or 'bonds', connecting the vertices of cell σ; the first term, summing over these lengths, then represents the total vertex bond energy of cell σ. $A(\sigma)$ and $P(\sigma)$ are the area and perimeter of cell σ, respectively, and $A_T(\sigma)$ and $P_T(\sigma)$ are its target area and perimeter, respectively. Finally, λ_A, λ_P and λ_L, respectively, are phenomenological parameters weighting the contributions of the above energies.

As indicated, each cell σ has an associated set of vertices with position vectors $\mathbf{r} \in \mathcal{R}(\sigma)$ and each of the above cell properties are then defined by functions of these position vectors. This allows an equation of motion to be written for the coordinate vectors of the vertices, here an overdamped Newtonian equation. This is driven by (generalised) forces $\mathbf{F}_i$ obtained by differentiating the free energy function considered as a function of vertex coordinates. This gives

$$\eta_i \frac{d\mathbf{r}_i}{dt} = \mathbf{F}_i = -\nabla_{\mathbf{r}_i} U(\mathbf{r}_1, ..., \mathbf{r}_i, ...\mathbf{r}_V), \tag{8}$$

where η_i is the viscous drag coefficient of vertex i and $\nabla_{\mathbf{r}_i}$ the gradient operator with respect to $\mathbf{r}_i$.

While cell-centre models may be formulated in energetic terms, they are typically (but equivalently) formulated by applying force balances. Thus the equation governing the motion of $\mathbf{r}_i$, where i now indexes cell centres rather than the cell vertices, is similar in form to (8):

$$\eta_i \frac{d\mathbf{r}_i}{dt} = \sum_{j \neq i} \mathbf{F}_{ij}, \tag{9}$$

where $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$ is the assumed equal and opposite force that cell centre j exerts on cell centre i, an example being the linear spring law for which $\mathbf{F}_{ij} = k(|\mathbf{r}_i - \mathbf{r}_j - S_{Tij}|)(\mathbf{r}_i - \mathbf{r}_j)/(|\mathbf{r}_i - \mathbf{r}_j|)$, where S_{Tij} is the target length of the spring connecting cell centres i and j. In (9), the sum is over all cell centres j in a well-defined neighbourhood of the cell with centre i (e.g. by a Delaunay triangulation).

The 'subcellular element' approach was introduced by Newman [61] as a general framework for investigating the emergent behaviour of cells and tissues, including tumorigenesis; the mechanics of an epithelial sheet was chosen as an initial example but many other systems have been studied (see below). In this model, cells are defined by a finer set of vertices than just cell boundary or centre positions. In this

case i indexes a set of subcellular-scale vertices called 'subcellular elements', which have a phenomenological correspondence to internal cellular architecture. The total potential is split into a set of interaction potentials which here are further divided into intracellular and intercellular components depending only on the distance between element positions. A noise term ξ is often added (this may similarly be added to the above vertex or cell-centre models), and is typically chosen to be Gaussian with zero mean. Again, however, the principles are the same. Here the equation of motion for the position vector $\mathbf{r}_i$ of a subcellular element i, with associated cell σ, takes the form

$$\eta_i \frac{d\mathbf{r}_i}{dt} = -\nabla_{\mathbf{r}_i} U + \xi \tag{10}$$

$$= -\nabla_{\mathbf{r}_i} \left(\sum_{\mathbf{r_j} \in \mathcal{R}(\sigma)/\{\mathbf{r}_i\}} U_{\text{intra}}(|\mathbf{r}_i - \mathbf{r}_j|) + \sum_{\sigma' \neq \sigma} \sum_{\mathbf{r_j} \in \mathcal{R}(\sigma')} U_{\text{inter}}(|\mathbf{r}_i - \mathbf{r}_j|) \right) + \xi, \tag{11}$$

where the first sum of accounts for interactions between element i of cell σ and all other intracellular elements of the same cell, while the second sum governs interactions of element i with each element j of each other cell σ'.

Though the above models are all similar in form, each comes with its own conceptual framework for specifying the particular interaction potentials (i.e constitutive modelling); thus detailed comparisons of all three approaches under different circumstances would be helpful. This will be greatly aided by the development and sharing of open source software; an example is Chaste [55], which currently implements vertex and cell-centre models. For now we consider a range of different examples, where one particular approach was typically chosen.

Intestinal crypt mechanics have been widely studied with IBM models. Meineke et al. [51] developed the first lattice-free IBM model of cell motion in the intestinal crypt; the crypt is modelled as a cylindrical surface and cell centres interact via linear spring forces. The model was developed to overcome limitations of earlier, lattice-based CA models of the intestinal crypt such as [45]. In particular, the model [51] allows for continuous time and space steps (to computational resolution), avoids unrealistic shifting of columns of cells, and cell motion is not based on rules involving age or knowledge of neighbouring cell types, in contrast to [45]. Furthermore, the model can be adapted to determine how the cell dynamics change when the idealised cylindrical geometry of the crypt is replaced by a more realistic test-tube shaped geometry.

More recently, other authors have included more biological detail into IBM models of the intestinal crypt [81, 9, 19, 54, 2]. For example, van Leeuwen et al. [81] construct a multiscale model of the intestinal crypt which couples dynamic, subcellular models of the cell-cycle and Wnt signalling pathway with a cell-scale mechanical model based on Meineke et al.'s [51]. Their model simulations reveal that relating cell proliferation explicitly to the cells' biochemical environment or 'niche' more naturally explains the experimental data than the idea that cells are intrinsically 'stem' or 'non-stem' cells independent of their environment. Building on the van Leeuwen et al. model, Mirams et al. [54] studied how monoclonal conversion in the colonic crypt (i.e. when all cells in a crypt have emerged from the same stem cell) is affected by mutations causing changes in the proliferative and adhesive properties of cells, as well as the spatial location at which these mutations occur. A particularly

robust finding of theirs was that mutations must occur in the bottom few locations of the crypt in order to have any chance of their descendants taking over the crypt. Thus a better understanding of spatial dynamics can lead to a better understanding of the transitions from healthy to diseased states.

A notable early application of the subcellular element method was to individual cell rheology [72]; here the authors found evidence that they need to incorporate active cytoskeletal rearrangement to account for long-time/large-strain behaviour. The incorporation of active processes was subsequently investigated by the same authors in both a cellular and multicellular context [71]. They give favourable comparisons to experiments of a range of different types - cell stretching under large-strain conditions, tissue fluidity in embryos and streaming patterns in collective cell motion (see also [70]). Other applications of this method to epithelia include those by Nie and co-workers [12, 15], who applied the subcellular element method, giving particular attention to its implementation on graphics processing units (GPUs), to investigate the epidermal epithelium, finding for example that a polarized distribution of adhesion proteins, which inhibits cell detachment, in combination with asymmetric cell divisions provides a robust mechanism for the formation of the stratified structure of skin layers. The subcellular element method is one of the most detailed frameworks available for modelling epithelial dynamics and will likely continue to be applied to a range of systems.

5. Continuum models.

5.1. Overview and motivation.
Biological tissues such as epithelia have been modelled as continua for almost a century. Macroscale representations of growth have been of particular interest, stimulated initially by D'Arcy Thompson's [79] simple geometric theories. Recent reviews of continuum theories for growth and remodelling, and on continuum and discrete representations of growth, can be found in [1] and [40], respectively.

Simple continuum models, based on conservation principles and constitutive assumptions (e.g. Darcy's law for cell movement through a fluid), have been developed to describe epithelial tissues including the intestinal crypt. Such models are reviewed and compared to cell-based models in [64, 83], in the context of cancer in the colorectal crypt epithelium. Simple convection-diffusion-reaction models have recently been used to study the effects of chemical signalling on crypt formation in the intestine [85], tissue stratification in epithelia [11, 66] and for formulating an inverse problem for proliferation rates in the crypt [17].

It is often difficult, however, to relate mechanisms included in the continuum models to mesoscopic processes such as those involved in individual cellular rearrangements. In this section we primarily study epithelial relaxation and rearrangement by taking continuum limits of some of the IBMs discussed above, before returning briefly to recent developments in continuum modelling directly incorporating more mesoscopic information.

5.2. Relations between discrete and continuum modelling.
Murray et al. [58] argue that nonlinear diffusion equations may represent a conceptual continuum framework for comparing different discrete models of epithelial cell movement, and also for 'inverting' discrete to continuum derivations in order to formulate discrete

models that are compatible with a particular nonlinear diffusion equation. In particular, they note [58] that continuum limits of IBM models frequently result in equations of the form, here in one-dimension,

$$\frac{\partial n}{\partial t} = \frac{\partial}{\partial x}\left(D(n)\frac{\partial n}{\partial x}\right), \tag{12}$$

where $n(x,t)$ is the cellular density at position x and time t and $D(n)$ is a nonlinear diffusion coefficient. Here we have neglected growth, and just consider the relaxation of tissues towards equilibrium; we return to this issue in the discussion. A benefit of deriving such a coarse-grained description is the availability of mathematical analysis tools for the study of such models. Before we discuss this further, we consider some examples.

Murray et al. [57] consider a discrete model of one-dimensional cell motion, in the overdamped limit for which viscous forces are dominant and similar in form to the cell-centre model in [51]. By converting the discrete cell index into a continuous index and assuming that cell numbers are large, they obtain an equation of the form (12), with

$$D(n) = \frac{k}{\eta n^2} \tag{13}$$

where, as above, n represents the cell density, k is the spring constant associated with cell-cell interactions and η is a viscous drag coefficient characterising resistance to cell motion. Murray et al. [57] demonstrate good agreement between simulations of the discrete and continuum models and note that models of this sort have appeared in the context of 'fast' diffusion phenomena.

Using an alternative approach, Lushnikov et al. [46] consider continuum limits for a (lattice-based) cellular Potts model of collective cell movement. As discussed above, interactions are incorporated via an energetic (Hamiltonian) approach; for example, cell-cell adherence, interaction with the extracellular matrix and an excluded volume constraint (ensuring cells do not overlap) are all included in the 'effective energy'. They take the surrounding medium to be viscous and neglect inertial effects. They first derive a chemical master equation before taking a continuum limit. Despite their different approach and model type, this leads to a nonlinear diffusion equation of the same form as (12), this time with

$$D(n) = C\frac{1 + (\frac{n}{n_0})^2}{(1 - \frac{n}{n_0})^2} \tag{14}$$

where n is the cell density, C is a constant and n_0 is the equilibrium cell density, equal to one over the local cell rest length (see also [58, 48]).

As noted above, Murray et al. [58] explain how different discrete models can be compared by analysing the functional form of the nonlinear diffusion coefficient that is generated when a continuum limit of the cell-based equations is taken. As further support for this approach, they found that they could derive an equivalent Newtonian force law at the cell level which recovers the same continuum description of the CPM as Lushknikov et al. [46]. This illustrates how information at the continuum level can be used to infer information about possible behaviour at the cell scale and to compare alternative cell-based descriptions. Furthermore continuum models are generally much easier to analyse than cell-based models; they can be used

either independently or in combination with cell-level simulations. For example, Murray et al. [59] found, by adding a cell proliferation term to equation (12) with diffusion coefficient (13), that a bifurcation in model behaviour occurs - if the diffusion is too slow relative to cell division rate, the system does not reach a steady state but instead 'blows up'. Since this is biologically unrealistic under normal conditions, this aids in understanding the important balances of different biological processes, as well as providing guidance for parameterisation of the IBM model.

There are several directions that require further work, however. It is not clear, in particular, how growth, topological changes, subcellular signalling and energy transduction should be rigorously incorporated. Extensions of these derivations to higher dimensions also remain to be established. While Fozard et al. [21] also investigate continuum limits of a discrete mechanical model of cell motion, systematically considering both slowly varying and heterogeneous cell parameter limits, they also limit their analysis to one dimension. In addition, the approaches of Murray et al. [58, 59] and Fozard et al. [21] are based on limits of deterministic models; it is likely valuable to retain more explicit considerations of stochastic effects, as in Lushknikov et al.'s approach [46]. Middleton et al. [53], building on the work of Newman and Grima [60], present a similar approach for a stochastically-forced point-mass cell-centre model (i.e. a Langevin equation model) of cell migration in the presence of adhesive effects, taking careful account of the correlations induced by strong mechanical interactions between cells. In principle this approach is applicable in three-dimensions but only one-dimensional comparisons are carried out.

Contrasting with the above 'bottom-up' approach, we briefly mention the more 'top-down' generalised continuum theories, which may provide one alternative or, more appropriately, a complement to 'bottom-up' modelling [40, 1]. The sophistication of continuum models has also increased significantly in parallel with discrete models, with the introduction of additional mesoscopic state variables, multiple phases and extensions of continuum kinematics. These generalised continuum theories are an important approach to representing phenomena such as growth, rearrangements, plastic deformations and other dissipative phenomena, as well as coupling to processes such as chemical signalling and mechanotransduction. Here we can only briefly review this large topic and refer the reader to the books [86, 50, 34, 31, 32] for excellent introductions to such generalised continuum theories (here ultilising subtle concepts such as 'internal state variables' and 'configurational forces'), and the articles [23, 40, 52, 1] reviewing, among other topics, multiphase theories, continuum thermodynamics and the so-called multiplicative decomposition of the continuum deformation gradient in the context of growth and remodelling of biological materials.

This approach can help bypass the technical difficulties of homogenisation of underlying discrete models; however it is probably best combined with partial homogenisation, statistical mechanics and simulation targeted at modelling particular state variables such as the 'growth tensor' often used in continuum biomechanics [40, 1]. This filling in of mesoscopic details, providing appropriate interpretations and determining the validity of a continuous description is the central difficulty involved in applying these generalised continuum models . An interesting practical example of tackling this challenge is [5] (see also [4]) where the kinematic tools of continuum field theory were used, along with statistical fitting methods, to quantify deformation rates at both the tissue and mesoscopic cellular level (cell shape

changes, sliding and rearrangements) in epithelial tissues to assess their impact on development. The rearrangement/intercalation deformation tensor they define provides an example of a mesoscopic 'microstructural' variable that fits well with the approach of generalised continuum theories. Similarly, Bonnet et al. [7] investigated the applicability of continuum theory to epithelia through the use of laser ablation experiments, live imaging and statistical fitting methods, finding that the continuum description was well justified (see also their related earlier work on connecting statistical processing of foam structures to continuum theories [30, 49, 68]).

6. **Discussion.** As detailed in this review, epithelial tissue provides an important experimental model system for investigating a broad range of intertwined biological phenomena, including cell lineage dynamics, tissue repair, tissue homeostasis, tissue development and tumorigenesis. Many features of these processes are still not fully understood, however, and in particular we face three key challenges, raised in the introduction: I) choosing which of the diverse biological processes, spanning multiple spatial and temporal scales, to include in our mathematical models (*scale and resolution of models*); II) how to best compare models of different types (*model-model comparison*) and III) how to best connect our models to data (*model-data comparison*).

We first summarise the features of the various modelling approaches considered in this review, and then discuss some possible approaches for tackling the above challenges more systematically.

6.1. **Summary of modelling approaches considered.** Compartmental models provide a robust and high-level approach suitable for investigating feedbacks and signalling, while having limitations on the resolution of spatial and mechanical factors. Individual-based models allow for higher-resolution descriptions of cellular motion and mechanics. As such they offer many new insights but can be difficult to analyse and are still subject at present to computational limitations for large simulations. Continuum models provide an approach with a spatial and mechanistic resolution lying between broad-brush compartment models and detailed discrete models, and standard formulations typically lack the detail to permit comparisons with important cellular-level data. While each approach has its strengths, and in many cases each can be used independently of the other, there is much to be gained from methods allowing the comparison of different models. Continuum limits of discrete models are one approach to tackling this; alternatively, modern continuum models tend to include much more 'microstructural' information than they previously did. This apparent 'meeting in the middle' offers many opportunities to incorporate and combine the strengths of different approaches, but requires practitioners to first be aware of the range of alternatives - one aim of this review is to help further this goal.

6.2. **Future directions - inference.** Addressing questions about how much detail to include, which model class(es) to use and how to relate models to each other and to real biological phenomena remains an art rather than science. These questions have typically been tackled in an *ad-hoc* manner in mathematical biology; on the other hand, advances in computation, formalisation of inference methods and new imaging methods should enable some aspects of these questions to be addressed more systematically.

In particular, in statistics, the rise in computational power has seen a resurgence of interest in Bayesian inference methods, typically through Markov chain Monte Carlo sampling methods (see for example the account in [3]). One of the biggest strengths of the Bayesian approach is that it provides a formalised inference framework for combining models with disparate data sources and other information and for characterising the uncertainty and/or 'underdetermination' in parameters [25, 78]. This provides one possible inference-based perspective for assessing models and relating them to data[1]. The key here is to interpret probabilistically *both* observable data and model parameters (this latter part being particular to the Bayesian inference approach), which allows/forces one to define a 'full probability model' or joint distribution for observable data and model parameters [25, 78]. By conditioning on a sample of observed data, one can update *prior* estimates of model parameters (expressed in terms of probability distributions) and as-yet unobserved data to obtain a new (*posterior*) predictive model. One appealing aspect of this approach is that this posterior can then be used either as a prior to be updated on the basis of further data or as a predictive model, which can also be subject to model checking/tests[2]. There are of course a number of mathematical and modelling complexities, especially when dealing with the high or even infinite-dimensional nature of parameter estimation and uncertainty quantification for truly spatio-temporal models. This represents an exciting interface between applied mathematics and statistics; we refer the reader to [77, 13] for more mathematical details in the context of PDE models. We illustrate the general Bayesian approach in Figure 3.

The recent work of Oden et al. [62, 63, 33] illustrates this approach in the context of tumour modelling. They combine a modelling framework of appropriate (spatio-temporal) resolution - the generalised phenomenological framework of continuum thermodynamics of mixtures - with the formalised methods of Bayesian inference and uncertainty quantification as outlined above. Their approach is still rather general and we refer the reader to this work for more information. Of particular note, however, is that even a deterministic continuum model such as that developed in [62, 63, 33] is now interpreted stochastically, here as a set of stochastic partial differential equations. With this in mind, we speculate that this 'inference-centric' approach may further bridge the gap between model types, with both initially discrete or continuous and deterministic or stochastic models ultimately having common interpretations in terms of probabilistic models linking both observable data and model parameters. By translating models into this common language[3] one could, in principle, facilitate model comparisons and derivations of one type from another (both 'bottom-up' and 'top-down').

---

[1]There are of course many sophisticated 'frequentist' approaches, but we do not consider these here.

[2]We note that parameter estimation, uncertainty quantification (probabilistic predictions from uncertain input) and model checking can all be carried out within the Bayesian framework, as nicely described for example in [25, 26]. We also note that, as argued by these authors, while it may be difficult to obtain good prior information about some parameters, this approach forces one to explicitly acknowledge and (at least formally) account for this uncertainty. See also the high-level overview [35] and more mathematically-detailed [77, 13] linking Bayesian methods to the field of uncertainty quantification.

[3]One well-known advocate of an approach such as this was the physicist E.T. Jaynes, who (re-)interpreted a number of solution methods such as the methods of statistical mechanics and regularization methods in inverse problem theory as problems of (Bayesian) inference (see e.g. [37, 36]). Similar influential views may be found in the work of geophysicist A. Tarantola, e.g. [78, 56], along with some proposed improvements to aspects of the traditional Bayesian approach.
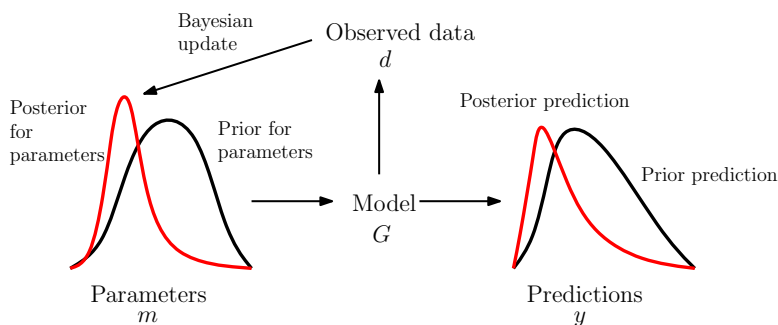
FIGURE 3. In the Bayesian approach one starts from a fully-specified probabilistic model, which can generate predictions based on assumptions about model parameters which are expressed in terms of a *prior* probability distribution. If new data are obtained, one can use the Bayesian conditioning procedure to update the information on model parameters into the form of a *posterior* probability distribution. This can then be used to generate new predictions or carry out model checks, and this process can be repeated if further data are obtained. Based on [35].

Mathematical biology is an area ripe for application of methods such as these - not just to the newest or most complicated models, though these are certainly of major interest, but also to return to previous models or theories that often far outstripped available data. This will should enable us to systematically test, validate and improve on these models, starting from good empirical information.

6.3. **Final thoughts.** In general, advances in imaging, computation and statistical methods have brought models and data closer together and this means that we can develop better ways of testing, comparing and choosing between different modelling approaches. Methods are needed for organising and summarising, in an understandable way, different models, interesting data features and complex couplings of biological processes. In this discussion we have speculated that treating specific problems of mathematical biology as problems of inference provides a promising guide to choosing and comparing appropriate models, characterising uncertainty and even deriving one model type from another.

This perspective is but one promising approach to be further developed and employed. Furthermore, as we have noted with earlier models, modeller creativity will often outstrip available data; in the right balance this is not necessarily a bad thing. While we feel that some of the most exciting challenges presently involve directly relating models and data, novel models will likely continue to provide conceptual foundations extending even further into the future.

REFERENCES

[1] D. Ambrosi, G.A. Ateshian, E.M. Arruda, S.C. Cowin, J. Dumais, A. Goriely, G.A. Holzapfel, J.D. Humphrey, R. Kemkemer, E. Kuhl, J.E. Olberding, L.A. Taber, and K. Garikipati. Perspectives on biological growth and remodeling. *J. Mech. Phys. Solids*, 59(4):863–883, 2011.

[2] A.-M. Baker, B. Cereser, S. Melton, A.G. Fletcher, M. Rodriguez-Justo, P.J. Tadrous, A. Humphries, G. Elia, S.A.C. McDonald, N.A. Wright, B.D. Simons, M. Jansen, and T.A. Graham. Quantification of crypt and stem cell evolution in the normal and neoplastic human colon. *Cell Rep.*, 8(4):940–947, 2014.

[3] J.O. Berger. Bayesian analysis: A look at today and thoughts of tomorrow. *J. Am. Statist. Assoc.*, 95(452):1269–1276, 2000.

[4] G.B. Blanchard and R.J. Adams. Measuring the multi-scale integration of mechanical forces during morphogenesis. *Curr. Opin. Genet. Dev.*, 21(5):653–663, 2011.

[5] G.B. Blanchard, A.J. Kabla, N.L. Schultz, L.C. Butler, B. Sanson, N. Gorfinkiel, L. Mahadevan, and R.J. Adams. Tissue tectonics: morphogenetic strain rates, cell shape change and intercalation. *Nat. Methods*, 6(6):458–464, 2009.

[6] M. Block, E. Schöll, and D. Drasdo. Classifying the expansion kinetics and critical surface dynamics of growing cell populations. *Phys. Rev. Lett.*, 99, 248101, 2007.

[7] I. Bonnet, P. Marcq, F. Bosveld, L. Fetler, Y. Bellaïche, and F. Graner. Mechanical state, material properties and continuous description of an epithelial tissue. *J. R. Soc. Interface*, page rsif20120263, 2012.

[8] N.F. Britton, N.A. Wright, and J.D. Murray. A mathematical model for cell population kinetics in the intestine. *J. Theor. Biol.*, 98(3):531–541, 1982.

[9] P. Buske, J. Galle, N. Barker, G. Aust, H. Clevers, and M. Loeffler. A comprehensive model of the spatio-temporal stem cell and tissue organisation in the intestinal crypt. *PLoS Comput. Biol.*, 7(1):e1001045, 2011.

[10] A.J. Carulli, L.C. Samuelson, and S. Schnell. Unraveling intestinal stem cell behavior with models of crypt dynamics. *Integr. Biol.*, 6(3):243–257, 2014.

[11] C.-S. Chou, W.-C. Lo, K.K. Gokoffski, Y.-T. Zhang, F.Y. Wan, A.D. Lander, A.L. Calof and Q. Nie. Spatial dynamics of multistage cell lineages in tissue stratification. *Biophy. J.*, 99(10):3145–3154, 2010.

[12] S. Christley, B. Lee, X. Dai and Q. Nie. Integrative multicellular biological modeling: a case study of 3d epidermal development using gpu algorithms. *BMC Sys. Biol.*, 4(1):107, 2010.

[13] M. Dashti, and A.M. Stuart. The bayesian approach to inverse problems. *arXiv preprint arXiv:1302.6989v3*

[14] G. De Matteis, A. Graudenzi, and M. Antoniotti. A review of spatial computational models for multi-cellular systems, with regard to intestinal crypts and colorectal cancer development. *J. Math. Biol.*, 66(7):1409–1462, 2013.

[15] A. Gord, W.R. Holmes, X. Dai, and Q. Nie. Computational modelling of epidermal stratification highlights the importance of asymmetric cell division for predictable and robust layer formation. *J. R. Soc. Interface*, 11(99):20140631, 2014.

[16] A. Deutsch, and S. Dormann *Cellular Automaton Modeling of Biological Pattern Formation: Characterization, Applications, and Analysis.* Springer, 2005.

[17] I.N. Figueiredo and C. Leal. Physiologic parameter estimation using inverse problems. *SIAM J. Appl. Math.*, 73(3):1164–1182, 2013.

[18] A.G. Fletcher, G.R. Mirams, P.J. Murray, A. Walter, J.-W. Kang, K.-H. Cho, P.K. Maini, and H.M. Byrne. Multiscale modeling of colonic crypts and early colorectal cancer. In *Multiscale Cancer Modeling*, Editor: TS Deisboeck, ISBN: 9781439814406, 2010.

[19] A.G. Fletcher, C.J.W. Breward, and S.J. Chapman. Mathematical modeling of monoclonal conversion in the colonic crypt. *J. Theor. Biol.*, 300:118–133, 2012.

[20] A.G. Fletcher, J.M. Osborne, P.K. Maini, and D.J. Gavaghan. Implementing vertex dynamics models of cell populations in biology within a consistent computational framework. *Prog. Biophys. Mol. Bio.*, 113(2):299–326, 2013.

[21] J.A. Fozard, H.M. Byrne, O.E. Jensen, and J.R. King. Continuum approximations of individual-based models for epithelial monolayers. *Math. Med. Biol.*, 27(1):39–74, 2010.

[22] M.H. Friedman. *Principles and Models of Biological Transport.* Springer, New York, 2008.

[23] K. Garikipati. The kinematics of biological growth. *Appl. Mech. Rev.*, 62(3):030801, 2009.

[24] R.A Gatenby, K. Smallbone, P.K. Maini, F. Rose, J. Averill, R.B. Nagle, L. Worrall, and R.J. Gillies. Cellular adaptations to hypoxia and acidosis during somatic evolution of breast cancer. *Brit. J. Cancer, BJC*, 97(5):646–653, 2007.

[25] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari and D.B. Rubin. *Bayesian data analysis*. CRC press, 2013.

[26] A. Gelman and C.R. Shalizi. Philosophy and the practice of bayesian statistics. *Br. J. Math. Stat. Psychol.*, 66(1):8–38, 2013.

[27] J.A. Glazier and F. Graner. Simulation of the differential adhesion driven rearrangement of biological cells. *Phys. Rev. E*, 47(3):2128, 1993.

[28] J.A. Glazier, A. Balter, and N.J. Popławski. Magnetization to morphogenesis: a brief history of the Glazier-Graner-Hogeweg model. In *Single-Cell-Based Models in Biology and Medicine*, pages 79–106. Springer, 2007.

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

[30] F. Graner, B. Dollet, C. Raufaste, and P. Marmottant. Discrete rearranging disordered patterns. Part I: Robust statistical tools in two or three dimensions. *Eur. Phys. J. E Soft Matter*, 25(4):349–369, 2008.

[31] M.E. Gurtin *Configurational forces as basic concepts of continuum physics*, volume 137. Springer, 2000.

[32] M.E. Gurtin, E. Fried, L. and Anand. *The mechanics and thermodynamics of continua*. Cambridge University Press, 2010.

[33] A. Hawkins-Daarud, S. Prudhomme, K.G. van der Zee, and J.T. Oden. Bayesian calibration, validation, and uncertainty quantification of diffuse interface models of tumor growth. *J. Math. Biol.*, 67(6-7):1457–1485, 2013.

[34] G.T. Houlsby and A.M. Puzrin. *Principles of Hyperplasticity*. Springer, 2007.

[35] M. Iglesias, M. A. and Stuart. Inverse problems and uncerrainty quantification. In *SIAM News*, volume July/August, 2014.

[36] E.T. Jaynes. *ET Jaynes: Papers on probability, statistics, and statistical physics*, volume 50. Springer, 1989

[37] E.T. Jaynes. *Probability theory: the logic of science*. Cambridge University Press, 2003

[38] M.D. Johnston, C.M. Edwards, W.F. Bodmer, P.K. Maini, and S.J. Chapman. Examples of mathematical modeling: tales from the crypt. *Cell Cycle*, 6(17):2106, 2007.

[39] M.D. Johnston, C.M. Edwards, W.F. Bodmer, P.K. Maini, and S.J. Chapman. Mathematical modeling of cell population dynamics in the colonic crypt and in colorectal cancer. *Proc. Natl. Acad. Sci. USA*, 104(10):4008–4013, 2007.

[40] G.W. Jones and S.J. Chapman. Modeling growth in biological materials. *SIAM Rev.*, 54(1):52–118, 2012.

[41] S.K. Kershaw, H.M. Byrne, D.J. Gavaghan, and J.M. Osborne. Colorectal cancer through simulation and experiment. *IET Syst. Biol.*, 7(3), 2013.

[42] A.D. Lander, K.K. Gokoffski, F.Y.M. Wan, Q. Nie, and A.L. Calof. Cell lineages and the logic of proliferative control. *PLoS Biol.*, 7(1):e1000015, 2009.

[43] Y. Lee, S. Kouvroukoglou, L. McIntire, and K. Zygourakis. , A cellular automaton model for the proliferation of migrating contact-inhibited cells. *Biophys. J.*, 69:1284e1298, 2005.

[44] W.-C Lo, C.-S Chou, K.K Gokoffski, F.Y.-M. Wan, A.D. Lander, A.L. Calof, and Q. Nie. Feedback regulation in multistage cell lineages. *Math. Biosci. Eng.: MBE*, 6(1):59, 2009.

[45] M. Loeffler, R. Stein, H.-E. Wichmann, C.S. Potten, P. Kaur, and S. Chwalinski. Intestinal cell proliferation. I. a comprehensive model of steady-state proliferation in the crypt. *Cell Prolif.*, 19(6):627–645, 1986.

[46] P.M. Lushnikov, N. Chen, and M. Alber. Macroscopic dynamics of biological cells interacting via chemotaxis and direct contact. *Phys Rev. E*, 78(6):061904, 2008.

[47] A.M. Marchiando, W.V. Graham, and J.R. Turner. Epithelial barriers in homeostasis and disease. *Annu. Rev. Pathol. - Mech.*, 5:119–144, 2010.

[48] D.C. Markham, R.E. Baker, and P.K. Maini. Modelling collective cell behaviour. *Disc. Cont. Dyn. Syst.*, 34(12):5123–5133, 2014.

[49] P. Marmottant, C. Raufaste, and F. Graner. Discrete rearranging disordered patterns, part II: 2D plasticity, elasticity and flow of a foam. *Eur. Phys. J.E* 25(4):371–384, 2008.

[50] G.A. Maugin. *The Thermomechanics of Nonlinear Irreversible Behaviors*. World Scientific, 1999.

[51] F.A. Meineke, C.S. Potten, and M. Loeffler. Cell migration and organization in the intestinal crypt using a lattice-free model. *Cell Prolif.*, 34(4):253–266, 2001.

[52] A. Menzel and E. Kuhl. Frontiers in growth and remodeling. *Mech. Res. Commun.*, 42:1–14, 2012.

[53] A.M. Middleton, C. Fleck, and R. Grima. A continuum approximation to an off-lattice individual-cell based model of cell migration and adhesion. *J. Theor. Biol.*, 359:220–232, 2014.

[54] G.R. Mirams, A.G. Fletcher, P.K. Maini, and H.M. Byrne. A theoretical investigation of the effect of proliferation and adhesion on monoclonal conversion in the colonic crypt. *J. Theor. Biol.*, 312:143–156, 2012.

[55] G.R. Mirams, C.J. Arthurs, M.O. Bernabeu, R. Bordas, J. Cooper, A. Corrias, Y. Davit, S.-J. Dunn, A.G. Fletcher, D.G. Harvey, M.E. Marsh, J.M. Osborne, P. Pathmanathan, J.M. Pitt-Francis, J. Southern, N. Zemzemi, and D.J. Gavaghan. Chaste: an open source c++ library for computational physiology and biology. *PLoS Comput. Biol.*, 9(3):e1002970, 2013.

[56] K. Mosegaard, A. Tarantola. Probabilistic approach to inverse problems. *Int. Geophys. Series*, 81(A):237–268, 2002.

[57] 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. *Phys. Rev. E*, 80(3):031912, 2009.

[58] P.J. Murray, C.M. Edwards, M.J. Tindall, and P.K. Maini. Classifying general nonlinear force laws in cell-based models via the continuum limit. *Phys. Rev. E*, 85(2):021921, 2012.

[59] 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. *Phys. Biol.*, 8(2):026011, 2011.

[60] T. Newman and R. Grima. Many-body theory of chemotactic cell-cell interactions. *Phys. Rev. E*, 70(5):051916, 2004.

[61] T.J. Newman Modeling multicellular systems using subcellular elements. *Math. Biosci. Eng.: MBE*, 2(3):613-624, 2005.

[62] J.T Oden, A. Hawkins, and S. Prudhomme. General diffuse-interface theories and an approach to predictive tumor growth modeling. *Math. Mod. Meth. Appl. S.*, 20(03):477–517, 2010.

[63] J.T. Oden, E.E. Prudencio and A. Hawkins-Daarud. Selection and assessment of phenomenological models of tumor growth. *Math. Mod. Meth. Appl. S.*, 23(07):1309–1338, 2013.

[64] 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. *Phil. Trans. R. Soc. A*, 368(1930):5013–5028, 2010.

[65] N.B. Ouchi, J.A. Glazier, J.-P. Rieu, A. Upadhyaya, and Y. Sawada. Improving the realism of the cellular Potts model in simulations of biological cells. *Physica A*, 329(3):451–458, 2003.

[66] J. Ovadia and Q. Nie. Stem cell niche structure as an inherent cause of undulating epithelial morphologies. *Biophys. J.*, 104(1):237–246, 2013.

[67] F. Radtke and H. Clevers. Self-renewal and cancer of the gut: two sides of a coin. *Science*, 307(5717):1904–1909, 2005.

[68] C. Raufaste, S.J. Cox, P. Marmottant, and F. Graner. Discrete rearranging disordered patterns: Prediction of elastic and plastic behavior, and application to two-dimensional foams. *Phys. Rev. E*, 81(3):031404, 2010.

[69] L. Reuss. Epithelial Transport. *Compr. Physiol.*, 2011.

[70] S.A. Sandersius, M. Chuai, C. Weijer and T.J. Newman. A'chemotactic dipole'mechanism for large-scale vortex motion during primitive streak formation in the chick embryo. *Phys. Biol.*, 8(4):045008, 2011.

[71] S.A. Sandersius, C. Weijer, and T.J. Newman. Emergent cell and tissue dynamics from subcellular modeling of active biomechanical processes. *Phys. Biol.*, 8(4):045007, 2011.

[72] S.A. Sandersius and T.J. Newman. Modeling cell rheology with the subcellular element model. *Phys. Biol.*, 5(1):015002, 2008.

[73] M.S. Steinberg. Mechanism of tissue reconstruction by dissociated cells. II. Time-course of events. *Science*, 137(3532):762–763, 1962.

[74] M.S. Steinberg. On the mechanism of tissue reconstruction by dissociated cells. I. population kinetics, differential adhesiveness, and the absence of directed migration. *Proc. Natl. Acad. Sci. USA*, 48(9):1577, 1962.

[75] M.S. Steinberg. On the mechanism of tissue reconstruction by dissociated cells. III. Free energy relations and the reorganization of fused, heteronomic tissue fragments. *Proc. Natl. Acad. Sci. USA*, 48(10):1769, 1962.

[76] M.S. Steinberg. Reconstruction of tissues by dissociated cells. *Science*, 141(3579):401–408, 1963.

[77] A.M. Stuart. Inverse problems: a bayesian perspective. *Acta Numerica*, 19:451–559, 2010.

[78] A. Tarantola. *Inverse problem theory and methods for model parameter estimation*. SIAM, 2005

[79] W.T. DArcy Thompson. On growth and form. *Cambridge University Press*, 1917.

[80] S. Turner and J.A. Sherratt. Intercellular adhesion and cancer invasion: a discrete simulation using the extended potts model. *J. Theor. Biol.*, 216(1):85–100, 2002.

[81] I.M.M. van Leeuwen, G.R. Mirams, A. Walter, A.G. Fletcher, P.J. Murray, J.M. Osborne, S. Varma, S.J. Young, J. Cooper, B. Doyle, J.M. Pitt-Francis, P. Pathmanathan, L. Momtahan, J.P. Whiteley, S.J. Chapman, D.J. Gavaghan, O.E. Jense, J.R. King, P.K. Maini, S.L. Waters, and H.M. Byrne. An integrative computational model for intestinal tissue renewal. *Cell Prolif.*, 42(5):617–636, 2009.

[82] A. Voss-Böhme. Multi-scale modeling in morphogenesis: a critical analysis of the cellular potts model. *PloS one*, 7(9):e42852, 2012.

[83] A. Walter. *A comparison of continuum and cell-based models of colorectal cancer*. PhD thesis, University of Nottingham, 2009.

[84] N.A. Wright and M. Alison. *The Biology of Epithelial Cell Populations*, Volume 1. Clarendon Press Oxford, 1984.

[85] L. Zhang, A.D. Lander, and Q. Nie. A reaction–diffusion mechanism influences cell lineage progression as a basis for formation, regeneration, and stability of intestinal crypts. *BMC Syst. Biol.*, 6(1):93, 2012.

[86] H. Ziegler. *An Introduction to Thermomechanics*. Elsevier, 1983.

E-mail address: maclaren@maths.ox.ac.uk

E-mail address: fletcher@maths.ox.ac.uk

E-mail address: byrneh@maths.ox.ac.uk

E-mail address: maini@maths.ox.ac.uk