

Understanding tissue morphology: model repurposing using the CoSMoS process

Ye Li · Adam T. Sampson · James Bown ·
Hilal S. Khalil · Yusuf Deeni

Received: date / Accepted: date

Abstract We present CoSMoS as a way of structuring thinking on how to reuse parts of an existing model and simulation in a new model and its implementation. CoSMoS provides a lens through which to consider, post-implementation, the assumptions made during the design and implementation of a software simulation of physical interactions in the formation of vascular structures from endothelial cells. We show how the abstract physical model and its software implementation can be adapted for a different problem: the growth of cancer cells under varying environmental perturbations. We identify the changes that must be made to adapt the model to its new context, along with the gaps in our knowledge of the domain that must be filled by wet-lab experimentation when recalibrating the model. Through parameter exploration, we identify the parameters that are critical to the dynamic physical structure of the modelled tissue, and we calibrate these parameters using a series of *in vitro* experiments. Drawing inspiration from the CoSMoS project structure, we maintain confidence in the repurposed model, and achieve a satisfactory degree of model reuse within our *in silico* experimental system.

Keywords calibration · CoSMoS · model reuse · simulation

Y. Li · A.T. Sampson (corresponding) · J. Bown · H.S. Khalil · Y. Deeni
Abertay University, Bell Street, Dundee, DD1 1HG, UK
Tel.: +44-1382-308170
Fax: +44-1382-308177
E-mail: a.sampson@abertay.ac.uk

1 Introduction

The CoSMoS process (Andrews et al, 2010; Stepney et al, 2015) describes a principled approach to scientific modelling and simulation: it provides a structure for managing and documenting the iterative development of a simulation, and gives scientists and simulation developers tools to reason – with an appropriate balance of confidence and scepticism – about how their simulation’s results relate to the domain of study. CoSMoS is an agile approach to support development of models and simulations of complex systems: a user may organise their project entirely following the CoSMoS principles, or they may integrate some of the CoSMoS process as appropriate into an existing project. In brief, CoSMoS identifies, among others, the following concepts: (1) Research Context, which notes the high-level motivations and hypotheses; (2) Domain, the subject of scientific research; (3) Domain Model, defining the purpose, scale and scope of the simulation activity; (4) Platform Model, an engineering derivation of the Domain Model; (5) Simulation Platform, an encoding of the Platform Model; and (6) Results Model, capturing understanding of the simulation based on output.

Reusability of software components is a key concern of software engineering. Reusable components can – ideally – avoid the difficulty and expense of developing and validating substantial amounts of new software. But software developed for one purpose may not be reusable for a different purpose without substantial modification. In particular, a simulation component developed for one *in silico* experiment may rely on assumptions (parameter values, model simplifications, etc.) that are only valid within the context of that experiment. Adapting such a component for reuse in a different context requires careful consideration of the assumptions made during its design. In this paper, we use the general structure of the process of modelling and simulating complex systems as considered by CoSMoS to organise our thinking around how a model of physical interactions among cells can be adapted from one context – the formation of vascular structures from endothelial cells – to a different context, that is the effect of anti-cancer drugs on the growth of cells in *in vitro* systems. Neither of these projects was initially developed using a CoSMoS approach. To apply the CoSMoS process, we must first effectively reverse-engineer our work to date, and attempt to organise

the information we have about the systems under study and our models and simulations of them broadly in terms of the CoSMoS concepts (1) to (6) as listed above. We expect that this step in itself will prove valuable.

Our objective is specifically to reuse the software components that implement the physical aspects of the model of vascular formation, referred to as Model 1, within the implementation of the model for cell growth, referred to as Model 2, since this implementation required considerable development effort and are critical to the overall performance of the simulation. As the modes of physical interaction among cells are broadly similar between Models 1 and 2, this seems *intuitively* to be an appropriate approach – but identifying and revalidating our assumptions will help to build our confidence in our simulation’s results, and enable the future reuse of the implementation of physical aspects of cell interactions in other contexts. In respect of this, we use CoSMoS to guide our reuse.

In addition, we have domain expertise in the field of cell biology and have conducted wet-lab experimentation studying the growth of cancer cells in response to perturbation by the therapeutic drug 5-fluorouracil (5-FU), hypoxia (oxygen deprivation) and the combination of both effects. Importantly, under hypoxic stress cells can respond differently to therapeutic intervention (Papageorgis et al, 2011). We present a new model, Model 2, of the cell and its response to anti-cancer drugs and hypoxia, and identify those components of Model 1 that can be reused in the implementation of Model 2. We then fit Model 2 to the experimentally measured cell growth rate in our control condition: this serves as the baseline calibration. We then determine the *in silico* changes we need to make to reproduce other *in vitro* experimental conditions, specifically the effects of 5-FU and hypoxia as single interventions, and these changes are linked to plausible biological mechanisms. Finally we test the amended model against other experimental data describing the effect of 5-FU and hypoxia in combination.

The paper is structured as follows. Sect. 2 describes the original model (Model 1) of vascular structures. We describe the model through CoSMoS concepts as a structured means to report on the model context, assumptions, results and so on. In Sect. 3 we describe Model 2 using the same stepwise, CoSMoS structured approach to lay out the new research context,

domain, etc. in order that we can expose those parts of the original vascular structure model that may be re-used and those that may not, together with new additions required. In this way, CoSMoS is used as a framework to support model reuse. In Sect. 4 we describe our first efforts in using the new model to understand cellular response to therapeutic insult and environmental stress.

2 Model 1: Vascular Formation

2.1 Research Context

Mammalian cells require blood vessels to deliver oxygen and other nutrients in order to survive. Capillaries are made up of endothelial cells and one of the key elements that drives their growth is growth factor VEGF (Karamysheva, 2008). We seek to build a model to understand not only the range of structures that endothelial cells can form but also the process by which those structures are formed. We will then compare simulation results with the results of an *in vitro* experiment from the literature, demonstrating the formation of capillary structures from endothelial cells (Serini et al, 2003).

The experiments by Serini et al (2003) explore how the physical interactions among the cells, and their low-level physical properties, affect the larger-scale structural patterns in the resulting capillary network. The effects of varying concentrations of growth factors – which have a direct effect on the low-level physical properties of the cells – are of particular interest.

2.2 Domain

The vascular system is complex and made up of various kinds of cells, and it grows in three dimensions (Cavalli et al, 2007). We seek to simplify the study of vascular formation and study the behaviour of endothelial cells and their interactions. It is challenging to study vascular formation in three dimensions, and experimental studies have explored this in *in vitro* systems in two dimensions (Gamba et al, 2003). We start from the position of *in vitro* capillary formation experiments: endothelial cells are planted on a substrate and self-organise

to form capillaries; note we recognise the processes of capillary formation may differ *in vivo* (Karamysheva, 2008). In the *in vitro* experiment, microvessels are formed by the aggregation of endothelial cells, which themselves are formed by the differentiation of stem cells. This formation process has three stages (Folkman and Haudenschild, 1980; Cavalli et al, 2007):

- cell migration and early network formation;
- network remodelling, where cells connect to each other;
- further differentiation into tubular structures.

For the purposes of this experiment, we are only concerned with the first stage, which takes place between six and nine hours *in vitro* (Serini et al, 2003). At the end of this stage, the basic network structure has formed, but cells have not yet begun to bind to each other or to differentiate further. All cells are similar in general terms during this stage, although their individual properties may vary.

We believe that in this stage the most significant forces are those resulting from physical interactions: between pairs of cells, between cells and their surrounding medium (Batchelor, 2000; Szabó et al, 2011), and between cells and the substrate, Matrigel film (Serini et al, 2003, p. 1778). As the surrounding medium is relatively thin and the interactions with the substrate are strong, there is only limited potential for cell movement away from the substrate, and experiment imaging can be used effectively to capture cell positions in real-world systems. Time-series imagery can be used to characterise cell interactions – for example, Niles (2012) shows physical interactions among stem cells *in vitro*, including attraction between cells, and cell shape changes after differentiation and binding.

As cells follow growth factor gradients, the density of cells tends to be higher where there is a higher concentration of growth factors in the environment. The *in vitro* experiment examined the effects of an artificial reduction of growth factor levels across the environment, imaging control and reduced-factor experiments at 3 h intervals.

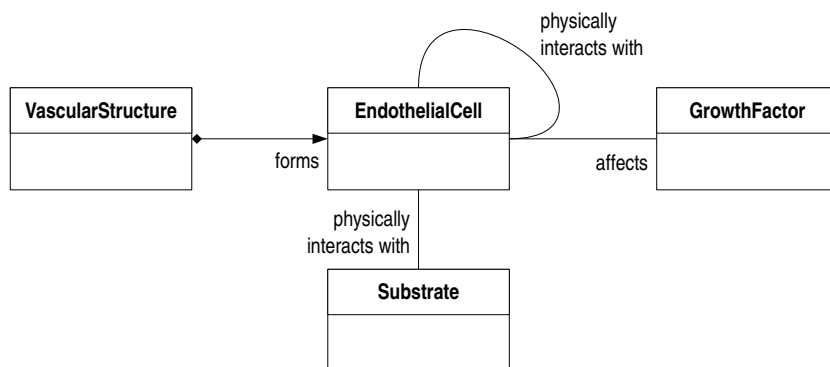


Fig. 1 Domain model: the entities of concern in the domain, shown as a UML-style class diagram

2.3 Domain Model

Within the Domain described in Sect. 2.2 we wish to study the dynamics of interactions of endothelial cells *in vitro*. This requires us to consider the cells as individual, interacting agents and we limit those interactions to inter-cell physics. We must build in a substrate on which cells grow and this substrate must include a growth factor to promote cell movement – cells follow growth factor gradients and this makes possible cell aggregation including but not limited to vascular formation (Karamysheva, 2008). We are only interested in simulating the first phase of vascular formation and during this time the total number of endothelial cells does not change (Serini et al, 2003). Thus it is reasonable to assume endothelial cells do not die or divide in the *in vitro* experiment. We do not therefore need to simulate cell differentiation or the cell cycle. Further, cells do not measurably change their size or shape over this time period.

Fig. 1 shows the entities within the *in vitro* experiment that we are attempting to reproduce, and their interactions. This includes both the biological entities under study and their experimental environment. Vascular structures are also included here as an emergent behaviour of the cells. Note that we have used UML-style diagrams in a rather informal way throughout our description, using familiar syntax but adapted semantics – for example, while growth factors are indeed individual molecules, it would clearly be inappropriate to think about them that way when modelling the system. Such diagrams are still useful as

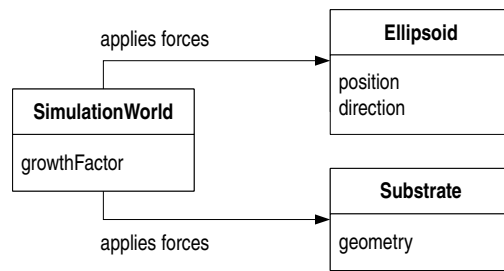


Fig. 2 Platform model: the entities of concern in the simulation, shown as a class diagram

“cartoon models” (Étienne, 2006), capturing our (necessarily limited) understanding of the system in a convenient but loosely-specified notation.

2.4 Platform Model

Existing discrete models of capillary formation can be divided to two types (Welter et al, 2009): those that consider cells as continuous density distributions, e.g. Holmes and Sleeman (2000); and those that consider cells as individual agents moving on a lattice, e.g. Turner and Sherratt (2002). We choose to consider cells as agents with particular shapes and seek to assign the same rule to all agents to reproduce the vessel structure as an emergent behaviour. However we will let the agents move in continuous space rather than a discretised environment. In this way we hope to capture more detailed cell movement, so that as a result we will have a better understanding of the vessel formation process, not only the vessel structure.

Fig. 2 gives an overview of how the domain model has been simplified for the purposes of the simulation. Cells show a direct correspondence between the domain and platform models. This allows us to define interacting rules for single agents, and examine both the lower-level properties of individual agents and the higher-level behaviour of the system as a whole. The simulation proceeds in discrete timesteps, with all agents updating their positions and orientations atomically at the end of a timestep.

Vascular structures are not represented in the Platform Model since these are the emergent property that we are attempting to reproduce. The Domain Model has been simplified to allow implementation of the physical interactions within the system as described below.

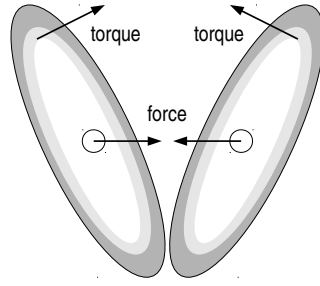


Fig. 3 Idealised ellipsoid agents within the platform model, showing torque and force

While cells can take a wide variety of shapes in the real world, we must model these as simpler shapes in order to practically simulate physical interactions at realistic scales. Modelling cells as simple spheres simplifies reasoning, but it does so by discarding information about the orientation of the cell, which limits the types of physical interactions that are possible. Initial prototyping showed that it was not possible to reproduce vascular formation behaviour using spheroid agents. Instead of forming the net-shaped structure, the spheroid agents form separate clusters.

We therefore represent cells as ellipsoids (Fig. 3). The shapes of cells observed in the *in vitro* experiment by Serini et al (2003) are roughly ellipsoidal (in the first phase), and previous models have assumed that ellipsoids are an appropriate representation of cell shape, e.g. Palsson (2001). An ellipsoid has three orthometric semi-axes, which can be used as a local coordinate system. The rotation of an ellipsoid can be represented by the change of this local coordinate system, and the direction of an ellipsoid can be represented by the transformation from the local coordinate system to the global coordinate system. The position of the centre of an ellipsoid represents the position of the whole ellipsoid. For simplicity we assume all the agents have the same size and shape.

We assume that the density of agents is evenly distributed, so forces can be modelled as acting on the centre of the agent, and changes in agent orientation can be modelled as torques acting on the agent. This is a modelling convenience and difficult to validate against experimental data, as cell rotations are hard to distinguish in time-series images.

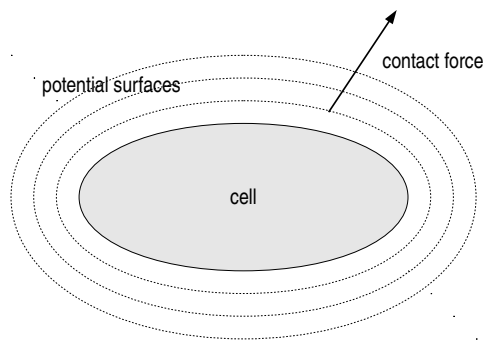


Fig. 4 An ellipsoidal agent showing potential surfaces, and the vector along which contact force is computed

We model physical interactions between cells in terms of forces between agents. The *adhesion force* attracts agents to each other; the *contact force* repels them and prevents them from overlapping; there is also a *resistance force* resulting from agents' interactions with the surrounding medium. For each force, there is an corresponding torque that is computed in an analogous way.

The contact force only takes effect when agents are in physical contact; the greater the overlap, the greater the contact force. It is introduced to prevent agents from totally overlapping. As the ellipsoid is not an isotropic shape, we cannot simply use the distance between two ellipsoids to calculate the contact force and torque. Instead, we compute a potential for each interaction: a path-independent potential energy. In Fig. 4, dotted lines represent potential surfaces around the agent – the potential is constant for any point on the same surface, although the distance to the agent centre will vary as a result of the agent shape. The potential is calculated following the approach given in Perram et al (1996), using the direction, position and length of the semi-axes of the interacting ellipsoids.

The potential is then transformed into energy using the Hertz formula. The magnitude of the resulting force or torque is the same for all points on a potential surface; the direction is computed based on the partial derivative of the energy field towards the centre of the interacting ellipsoid (Fig. 4).

The growth factor is considered to be a key factor in endothelial cell activation *in vitro* (Karamysheva, 2008). We model this mechanism as an adhesion force between agents,

which is modelled as a constant force attracting the centres of every pair of agents towards each other. This is the simplest approach that reproduces the behaviour observed in time-series images of the *in vitro* experiments. However early prototyping showed that the vessel structure was not stable due to interactions between far away agents, so that we must limit the distance that the adhesion force can take effect: if agents are beyond this distance they have no physical interactions; if they are within range, they move towards each other, until they become close enough to overlap, stopping at the point at which the adhesion force and contact force balance each other.

The relative strengths of the two forces may be calibrated so that this balance happens at a potential corresponding to that observed in cells *in vitro*. Unlike real cells, our agents cannot change their shape on contact. Yet we still need to model the elasticity of the real cell. The potential of this balancing point can be considered as the elasticity of the agents, with higher balancing potential levels indicating more rigid agents. Some elasticity is necessary to obtain realistic cell interactions: an early prototype of the model used a simpler approximation to the Hertz function which effectively gave inelastic collisions between agents, and resulted in agents visibly “bouncing off” each other – which did not match what we see in time-series images.

As cells move at relatively low speeds within the medium, according to classical mechanics (Goldstein et al, 2001) their acceleration can be approximated as zero – which means the sum of the forces upon them is also zero:

$$\sum \mathbf{F} = 0 = \mathbf{F}_{\text{contact}} + \mathbf{F}_{\text{adhesion}} + \mathbf{F}_{\text{resistance}} \quad (1)$$

We can therefore compute the resistance force in terms of the contact and adhesion forces – and, from this, compute the velocity of the agent using Stokes’ law, based on the known size and shape of the agent and the properties of the medium. The angular velocity can be found using a similar technique; from these, the position and orientation of the agent on the next timestep can be computed.

The substrate itself is modelled as a plane. The physical interaction between an agent and the substrate is modelled as the interaction between an agent and its mirror image in the

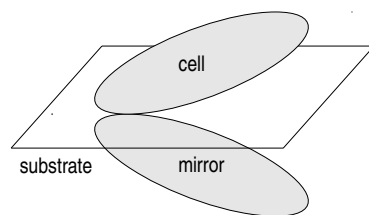


Fig. 5 Agent-substrate interaction, modelled as interaction with a copy of the same agent, mirrored in the substrate plane

plane (Fig. 5). Having observed that cells generally move in the plane of the substrate and do not grow on top of other cells under *in vitro* experimental conditions, we assume that the interactions between cells and the substrate are stronger than between cells and other cells. Therefore the adhesion force between an agent and its mirror image is scaled up to account for that.

The model is dimensionless, being defined in terms of a unit time (the simulation timestep) and a unit length (the radius of a typical agent). These two quantities are related, in that computing the velocity of an agent within the fluid medium depends on both the timestep and the shape of the agent. However, making an assumption about the maximum velocity of an agent allows us to find reasonable bounds for one unit knowing the other, and in our case choosing a unit timestep of 1 s gives a physically-plausible maximum velocity for endothelial cells.

2.5 Simulation Platform

The simulation is implemented by a program written in C++ in line with the assumptions and detail described in the platform model (Fig. 2). On program start up, the Simulation World initialises a set of agents and a plane that represents the substrate. According to the geometry definition of an ellipsoid, each agent is assigned with a random position within the range of the plane, as well as three orthometric directions for the three semi-axes. The lengths of the semi-axes, however, are fixed in the proportion 1 : 1 : 4. In this way, all the agents are identical with random position and direction. Typically there are 2000 to 4000 agents in one simulation.

After initialisation, the Simulation World enters a loop in which it calculates the forces and torques for agents, update their position and direction, and outputs the geometry data to text files. This loop repeats for the desired period of simulated time. Visualisation and analysis are carried out by external tools.

In each loop, the simulation goes through all the agents. For each agent, the Simulation World calculates the overall force and torque that other agents work on this agent, then calculates the force and torque between this agent and the plane. The rules to calculate the force and torque are defined within the agent object, so that the program may simulate emergent behaviours of a heterogeneous population of agents without changing the behaviour of the simulation rule set.

The Simulation World then estimates the total force and torque on the agent, thereby estimating the velocity and angular velocity of the agent, and uses these values to determine the new position and direction of the agent.

The position and direction information of all the agents are stored regularly into text files, resulting in a time-lapse record of the geometric information of all the agents. A visualisation program (Zhu et al, 2006) reads these text files in the same order, and produces 3D imagery using the OpenGL framework to render the agents as ellipsoids. When the visualisation program is running, we have the ability to observe the agents at an arbitrary distance and at an arbitrary angle. However for further study, we must save the scenario as images, in which the angle of observation is fixed. To best visualise the overall vessel structure, we choose the top-down angle. In each image, to separate agents and the plane, they are drawn with different colours: the agents are in green and the substrate is in blue.

As there are thousands of agents and a large amount of calculation, a full simulation would take up to 6 days to run on a single core. We made use of the Threading Building Blocks library (Reinders, 2007) to parallelise the simulation, achieving good parallel speedup.

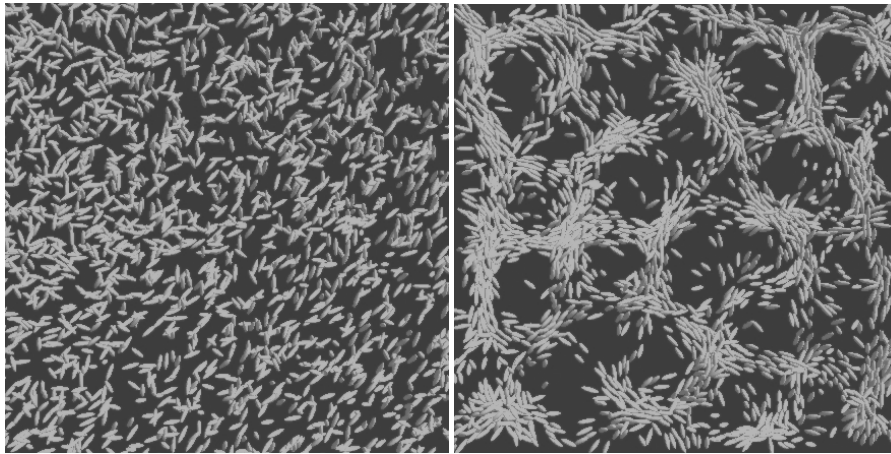


Fig. 6 Visualisation showing agent positions and orientations at the start (left) and end (right) of the simulation; “unstable” pattern

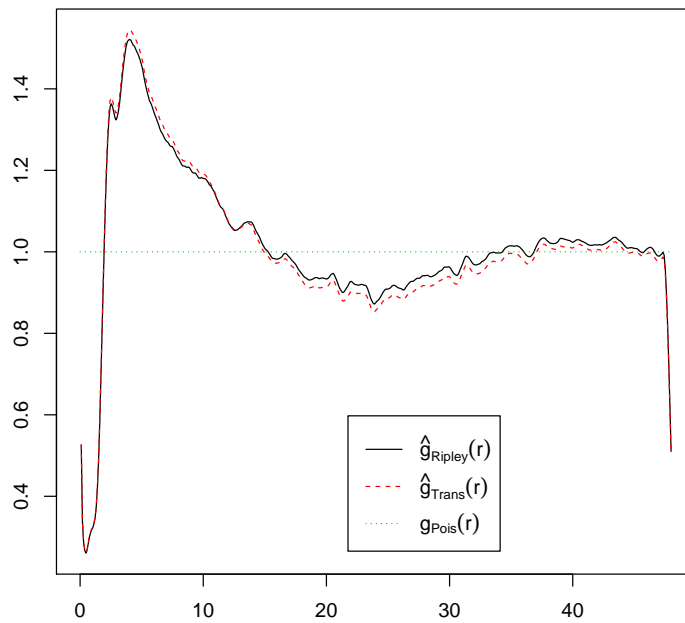


Fig. 7 Radial distribution of agents in Fig. 6 (right); X axis is distance between agents in simulation units, and Y axis is normalised probability of finding another agent at that distance

2.6 Results Model

Fig. 6 shows the starting and ending conditions of the simulation. This certainly resembles the vascular network we are trying to reproduce – but we need a quantitative measure of this, in order to relate the results back to the changes in the level of growth factor.

There is a quantised method to describe the pattern of this structure, which is called the radial distribution function. The radial distribution function is a tool to describe space distribution of a system that consists of particles (here agents), by describing the chance of finding another particle within an arbitrary distance from the reference particle. In the form of the distribution curve, normally the X axis is distance, and the Y axis is the function value. If the function value is larger than 1.0 at a certain distance, it means the agent density is higher than average at that distance; if the function value is smaller than 1.0, it means the agents are more sparse at this distance. Fig. 7 shows the radial distribution of agents at the endpoint of the simulation.

The minimum near distance 0 shows that agents tend not to have very close neighbours; the second minimum near distance 25 shows the typical size of hole in the net-shaped structure. As the distance from the reference agent increases, the value of the distribution function varies around 1.0, which means over longer distances the agents tend to be distributed evenly. We observe that our simulation based on physical interactions yields broadly similar results to those that Serini et al (2003) demonstrated using a model based on growth factor.

If we allow the simulation to continue past the state shown in Fig. 6 – i.e. past the period of time covered in the original model design – the pattern will collapse into a few large clusters of agents. Fig. 8 shows the results of an simulation where the physical parameters have been adjusted to produce a stable pattern that does not collapse; while some network structure is evident, these structures are formed at the scale of the individual cells. This scale reduction is evident in Fig. 9: the sharply-defined minimum around distance 12 (1.5 cell lengths), which corresponds to the typical vascular net size. We also see a more pronounced double peak resulting from cells positioned end-to-side (with cell centres 0.5 cell lengths apart) and end-to-end (1 cell length apart).

The validation of the model requires more work, but now we can show that the physical interactions among individual agents are linked to the emergent group pattern: by changing the physical parameters of individuals we can control the emergent structure. The simulation components relating to physical interactions are the parts that are most likely to be reused

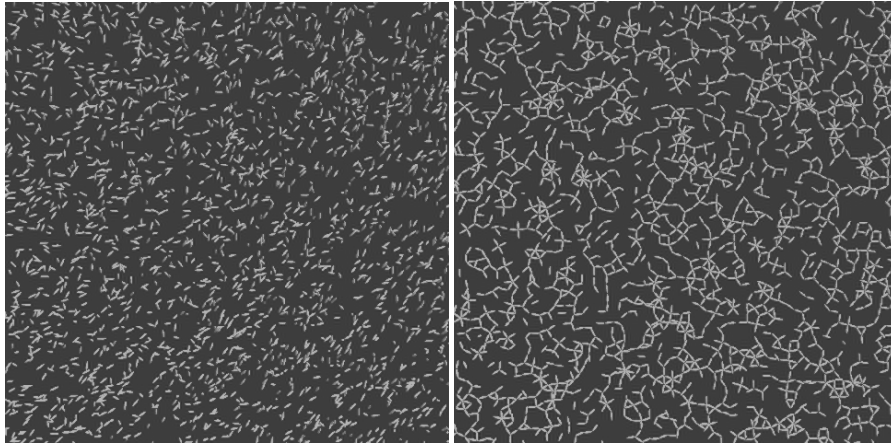


Fig. 8 Visualisation showing agent positions and orientations at the start (left) and end (right) of the simulation; “stable” pattern

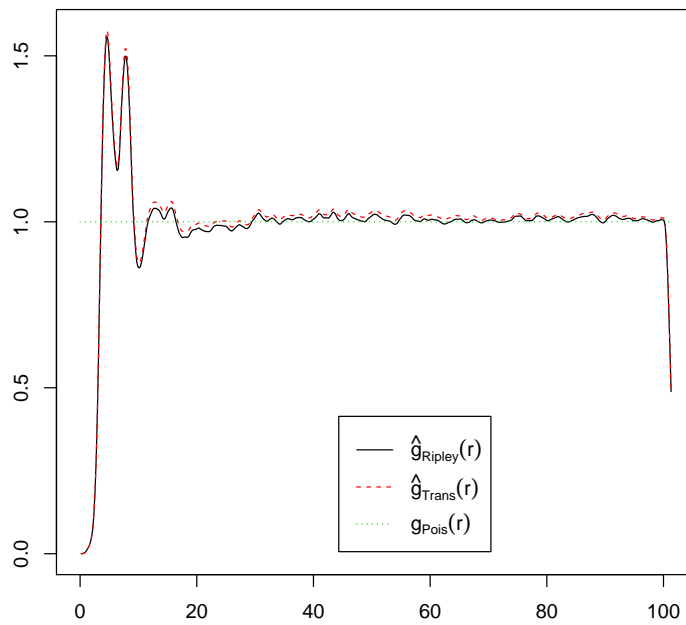


Fig. 9 Radial distribution of agents in Fig. 8 (right); X axis is distance between agents in simulation units, and Y axis is normalised probability of finding another agent at that distance

in Model 2, with the broader code structure noted in Sect. 2.5 that supports reuse with minimum changes.

2.7 Summary

Following the outline of the CoSMoS process, we have characterised Model 1 in terms of a set of assumptions that may affect reuse of the model. The assumptions (with their associated CoSMoS modelling phases) are:

- Only contact force, adhesion force and resistance force are significant (Domain).
- Cell size and shape do not change during the experiment (Domain Model).
- Cells do not divide or die during the timeframe of our experiment (Domain Model).
- Cells can be implemented as ellipsoids (Platform Model).
- Matter is evenly distributed within a cell (Platform Model).
- Contact force can be computed using the Perram-Wertheim approach (Platform Model).
- Adhesion force can be modelled as a step function on distance (Platform Model).
- Contact and adhesion forces balance at a defined point when cells are in contact, and the strengths of the forces can be calibrated based on this (Platform Model).
- Cells move at very low speed so their acceleration approaches zero and the forces upon them are balanced (Platform Model).
- Resistance force can be computed using Stokes' law, and the known properties of the fluid medium (Platform Model).
- Interactions with the substrate can be modelled as interactions with mirrored agents (Platform Model).

3 Model 2: Towards Spheroid Growth

Model 2 is also presented via the CoSMoS process. Clearly the research context is different, and so differences will cascade through all of the modelling phases. Despite this, both research contexts consider mammalian cells interacting in space over time. This commonality may offer an opportunity for code reuse from Model 1, and we use CoSMoS to help us structure our thinking on the reuse of code.

3.1 Research Context

As with the vascular development model (Model 1), our objective is to relate lower-level physical interactions to higher-level structural behaviours: specifically, we want to explore the effects of

- certain cancer treatment drugs,
- hypoxia (low concentrations of oxygen), and
- different cell lines (types of cell grown for experimental purposes)

upon the growth of cancer cells in general and tumours in particular. This work forms part of a wider programme of activity developing techniques for cancer drug discovery and development (Bown et al, 2012). Our domain experts are cancer researchers who are interested in making use of models and simulations to direct experimentation.

Tumours develop distinctive patterns of cells, which can be classified by domain experts either manually or using automated image processing. Indeed medical pathology relies heavily on this patterning as an indicator of disease progression and prognosis. It is specifically these kinds of spatial patterns that we are interested in reproducing within a simulation.

Our existing physical model (Model 1) has already demonstrated the ability to reproduce spatial patterns of cell growth resulting from physical interactions within an agent-based simulation, and we have existing tools to visualise and analyse output from the model. We would like to reuse as much of this infrastructure – both the model and the simulation code – as possible to reduce development time, but to do this we must identify the changes that need to be made by re-evaluating our original assumptions within the new research context.

In addition, we must identify what information necessary for reengineering and calibrating the model needs to be obtained by wet-lab experimentation. We aim to maximise the value obtained from this experimentation.

3.2 Domain

In the real-world domain, cancer cells develop and grow into tumours within surrounding tissue (Brú et al, 2003). In laboratory conditions, growth experiments are limited to popu-

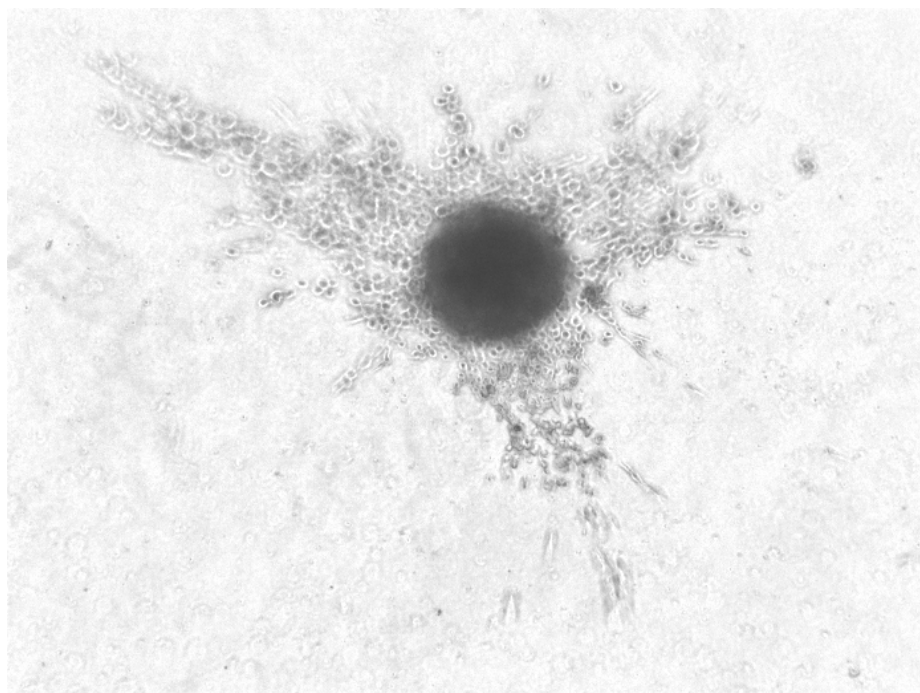


Fig. 10 2D side-view image of a three-dimensional spheroid growing within a gel medium

lations in a Petri dish – in which case cells can grow as a monolayer of cells on a substrate – or be suspended in a volume of gel, in which case spheroid structures can form (Fig. 10). Such spheroid systems provide 3D systems of cell cultures that are closer to the real system than 2D Petri dishes in terms of cellular response to drug action – see Kimlin et al (2013) for review. We have worked on 3D systems ourselves (Savage et al, 2013) and recognise both the value of 3D systems and the difficulties working with these systems. As noted by Kimlin et al (2013) an *in vitro* three-dimensional tissue model is not yet the standard but will play an increasing role in the drug discovery process in due course.

Petri dish experiments are the current standard since they are easier to configure, control and collect data from, since images can be taken non-destructively; spheroids must be sectioned before imaging in order to obtain data at a cellular resolution. A typical Petri-dish experiment may contain around 10^3 cells; a spheroid contains on the order of 10^6 cells. A single section through a spheroid is comparable in size to a Petri-dish experiment.

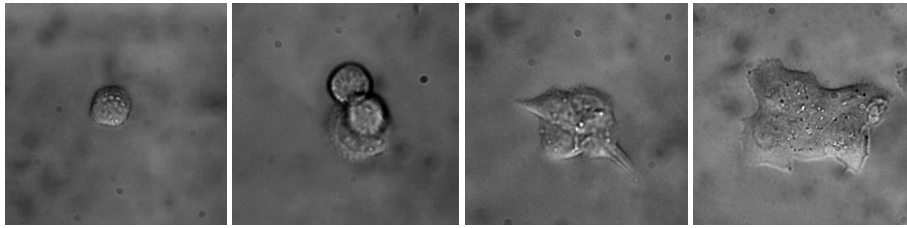


Fig. 11 HCT-116 (p53+/+) cells, imaged at 0 h, 6 h, 12 h and 18 h of growth on a glass plate. The diameter of the initial cell is 10 μm .

As a first step we are modelling Petri dish experiments although our ultimate goal is to work with spheroid structures. Our Petri dish experimental system has three important features that serve as a platform for spheroid modelling: we are using cells (cell line HCT116) that will cluster into a spheroid; we are using a therapeutically relevant drug; we are exploring the impact of hypoxia, i.e. oxygen stress and this is experienced by cells inside the spheroid.

The shape and volume of cells varies as they progress through their developmental cycle (Fig. 11) and the rates at which the cycle progresses vary somewhat among cell lines. The HCT-116 cells we are using (human colorectal cancer cells; see Sect. 4.1 for details) typically have diameter 10 μm immediately after division. Our experimental setup allows time-lapse imaging of cell growth and division at 10 min intervals. The maximum length of an experiment is constrained by cells only remaining healthy under experimental conditions for a limited period of time.

Some cancer drugs limit cell growth by arresting the cell cycle at a particular stage (Morgan, 2007). The progression of the cell cycle within the individual cells is therefore important when understanding the effects of drugs upon a tumour: if the cell cycles are synchronised (as can happen under *in vitro* and *in silico* conditions), then a drug can arrest many cells simultaneously, whereas cells at a mix of developmental stages will be less strongly affected.

For *in vitro* spheroid structures and *in vivo* tumours, we are particularly interested in the effects of hypoxia, which can have a suppressive effect on cell growth (Kaida and Miura, 2012). The high density of cells within a 3D tissue structure means that cells become increasingly hypoxic towards the centre of that structure.

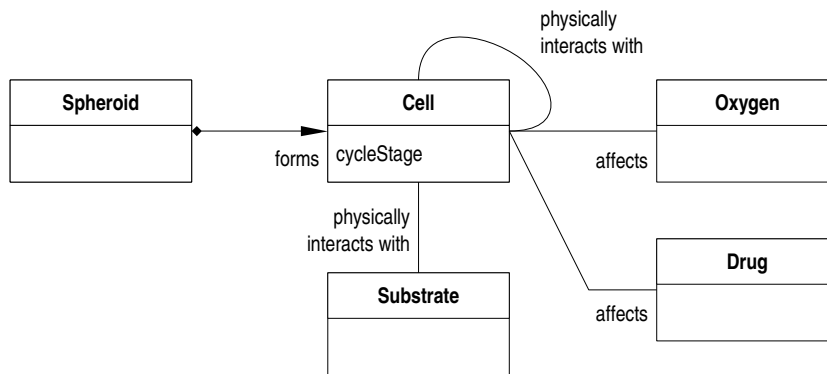


Fig. 12 Domain model: the entities of concern in the domain, shown as a class diagram

3.3 Domain Model

Fig. 12 shows the entities within the domain model. While the domain is substantially different from the previous one, the way cells are modelled retains a level of similarity, because the emergent behaviour of interest still results from physical interactions among cells. However, the physical properties of the cells themselves are somewhat different from our previous model – in particular, the cells’ properties are known to change over time, and we are interested in the effects of this on the emergent properties.

The main difference from Model 1 is that in order to capture the effects of drug action and hypoxia we must introduce a cell cycle. The cell cycle is the process by which the cell grows, replicates through division and dies. It comprises several phases that happens in a specific order. Following cell division cells are in G_1 phase in which a cell increases its size. During phase G_1 the cell may switch to and from phase G_0 – a resting phase. To progress beyond G_1 , the cell is required to pass a G_1 *check point* before entering the next phase in the cell cycle. The next phase after the G_1 *check point* is the synthesis phase S in which DNA replication occurs. Then the cell enters phase G_2 and continues to grow. The cell is next required to pass a G_2 *check point* before entering the mitosis (nuclear division) phase M , in which cell growth is stopped. A check point in the middle of phase M ensures that the cell is ready for division. The cell divides right after phase M into two daughter cells that

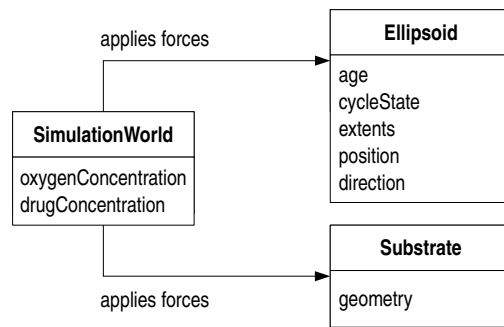


Fig. 13 Platform model: the entities of concern in the simulation, shown as a class diagram

contains roughly equal cellular components. The cell apoptoses if it fails to finish the cell cycle within a limit amount of time (Karamysheva, 2008).

The drugs and hypoxia conditions are added to our environmental conditions. They each impact the cell cycle progression differently. For the two-dimensional experiment, we still need to consider the substrate. But for the spheroid experiment, as the tumour cells grow in 3D in agar gel, we will no longer consider the substrate.

3.4 Platform Model

We established in the domain model that the individual development of the agents – e.g. the growth of cells over time – will be important to the behaviours we are trying to replicate, and must be taken into account in the simulation. As a result, we have chosen again to use an agent-based modelling approach. Fig. 13 shows the entities within the simulation platform. Comparing to the platform model of Model 1, vascular formation, it is clear that the age, cell cycle state and cell extents are newly added into our ellipsoidal agent, and the oxygen concentration and drug concentration are newly added to the simulation world, while the growth factor is removed from the simulation world.

Fig. 14 shows the state machine that models a simplified cell cycle and drives the behaviour of the simulated agent. This represents the observed behavioural modes of the cell – growth, reproduction, apoptosis – rather than the biological markers that would normally be used to describe cell cycle stages. We model the resting phase G_0 as the Idle state, from

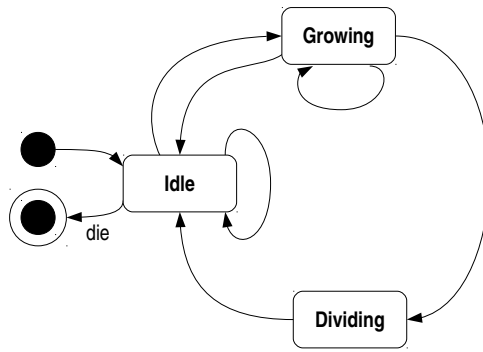


Fig. 14 Platform model: simplified cell cycle, shown as a state diagram

which the cycle of the agent starts. The growth phases G_1 , S and G_2 are modelled as the Growing state, during which the agent doubles its size. As we hope to focus on the spatial structure formed by the agents, we simplify the cell growth dynamics and will consider the change of agent size and shape only.

As introduced in Sect. 3.2, a cell may switch between phase G_0 and G_1 repeatedly. This behaviour is modelled as unlimited switching between the Idle and Growing state. The phase is not retained in our model, because the size of the agent is considered as constant in this phase. We model all check points together as one single state – the Dividing state, in which the agent decides whether to divide into two identical agents or to die. An agent enters the Dividing state when it doubles its volume, or is forced to by environmental conditions. The probability that one agent successfully divides is determined by the type of cell the agent represents and the environmental conditions. The process of death of agent is also simplified. If an agent dies it will be moved to the Idle state and is then deleted.

We introduce typical age which is the time period that an agent may exist and this represents the length of cell cycle (of a real cell). The total time an agent exists is recorded, and when it reaches the typical age, the agent will undergo the process of death. The value of typical age will be calibrated using *in vitro* data since this varies across cell types. The mechanism of action of our perturbations of drug intervention and hypoxia will take effect on these simplified states. An agent not affected by any environmental condition will start from the Idle state, then switch to the Growing state and remain there until it doubles its

volume, then either divide to two identical agents or die. However, a perturbation may alter the probabilities of transitions between states.

In the construction of the physical aspects of this platform model, we aim to reuse, as far as possible, our previous approach. In order to evaluate whether this is appropriate, we must reconsider our previous assumptions, listed in Sect. 2.7, based on our knowledge about the new research context.

While the physical properties of the cells and medium are somewhat different, we propose that most assumptions remain valid (see Sect. 3.7 for a summary). Two assumptions, however, are no longer reasonable: cell size and shape *do* change during the simulation. This requires changes to how the interaction potentials and their resulting forces and torques are computed, since these must now take changes to cell size and shape into account.

We must also ensure that we have sufficient experimental data to allow calibration of the physical parameters. We no longer just need the typical size and shape of a cell: we need a profile of how cell size and shape can change as the cell cycle progresses. This information will need to be obtained by time-series imaging under the experimental conditions we wish to simulate, as in Fig. 11. We will then give each simulated cell an interpolated cell growth curve relating cell size to time based on the measured points. The other information we need for calibration is available in the literature (e.g. the dynamic viscosity of the medium).

The choice of timestep size (i.e. unit time in the model) is a concern. The timestep must be short enough to obtain results at a comparable temporal resolution to the *in vitro* experimental data. However, smaller timesteps require more calculation steps to simulate the same length of real-world time; the 1 s timesteps used in the previous simulation would result in *in silico* experiments taking an impractically long time to run with typical simulation sizes of 10^3 – 10^6 cells.

In the *in vitro* experiment described in Sect. 4, the typical treatment time is 48 h to 72 h, and the sampling rate varies with different phases of the experiment. The total number of cells is measured every 24 h and a time-lapse image is taken every 10 min for the first 20 h of treatment. The simulation timestep does not therefore need to be any less than 10 min, and 1 h is assumed to be reasonable. The other model constants will need to be adjusted to suit

this time step – for example, this results in the simulation’s unit length being considerably smaller (which does not affect the outcome of the simulation).

3.5 Simulation Platform

The Simulation World initialises the agents at program start up. As discussed in Sect. 3.4, each agent is assigned a typical age, which is estimated by observing the length of the *in vitro* cell cycle in time-lapse images. However the initial simulation produced a staircase-shape growth curve which is unrealistic. To get a smooth group growth curve, a variance of 10% is introduced to the typical age, which means each individual agent has a different typical age that obeys a probabilistic distribution. After then initialisation, the Simulation World enters the loop for the desired number of time steps. In each loop iteration, the simulation not only maintains the physical aspects of agents, but also governs agents in switching their cell cycle state according to age. Also if an agent is in the Growing state, the Simulation World increases the size of the agent in each loop.

We introduce a probability of cell division to the simulation, which is a real number between 0 and 1. The probability of cell division represents the chance that one single agent divides at the end of its cycle: the maximum value is which means the agent will certainly divide; the minimum value is which means the agent will certainly die. The higher the value, the more likely it is that the agent will divide. In the model we use the probability of cell division to control the modelled cell fate such that each time an agent reaches its typical age, a random real number between 0.0 and 1.0 is generated: if this value is smaller than the probability of cell division value, the agent will be duplicated to represent cell division; otherwise the agent will undergo programmed cell death. If an agent is in a dividing state, the Simulation World duplicates it and resets the size and age of both agents; if an agent dies, the Simulation World removes that agent.

Similar to Model 1, the Simulation World also writes out the geometry information of all the agents regularly; it also records summary information including the agent population, for ease of visualisation.

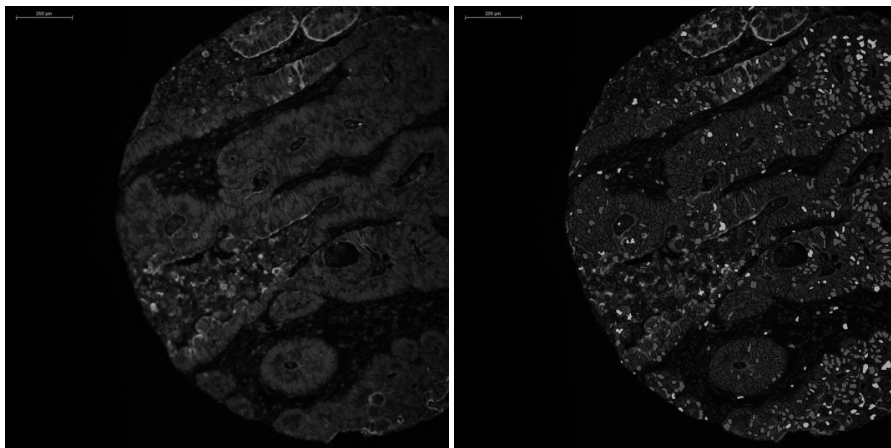


Fig. 15 Left: Ki67 expression in colorectal carcinoma tissue microarray data. Right: cell outlines and activity levels automatically identified from the previous image using Definiens.

3.6 Results Model

In our goal system of spheroids, we are primarily interested in the shape of spatial structures that tumour cells can form. Fig. 15 shows the complexity of spatial structures in real tissue and spheroid systems seek to replicate some of this complexity. We can expect cell density changes across a slice through the spheroid. Typically as hypoxic effects are stronger towards the centre of spheroids, the cell density in the centre part should be lower than the cell density near the surface of the spheroid.

We are also interested in the overall shape of the spheroid. There are existing image analysis tools that can be used for this. We will use them to extract information from wet-lab experimental imagery, including the position and direction of all the cells. We will then have directly compatible data from both wet-lab experiments and our simulation that can be analysed using a consistent approach. With this information we may use methods such as fractal geometry (Savage et al, 2013) to analyse the overall structure of both experimental and simulation data. Based on our experience with our physical model, we expect to be able to obtain reasonably good correspondence for those data obtained from growth experiments on substrate – but we suspect that extending the spatial interactions into three dimensions will require further elaboration of the model.

As a preliminary step to modelling spheroid dynamics we have developed an *in vitro* Petri dish system that allows us to monitor cell populations almost continuously. We use this system here to study population dynamics of cells under different experimental regimes: control, therapeutic drug action, hypoxic conditions (simulated with additional drug action) and therapeutic drug action in hypoxic conditions. The Petri dish lacks some of the realism of the spheroid structure but allows us to calibrate our model against a more substantial data set than we could feasibly derive from a spheroid. In subsequent work we will then re-fit this calibrated model to account for spheroid conditions and especially hypoxia; our studies in artificially introduced hypoxic conditions will make a valuable contribution to the model fitting we must ultimately undertake with limited data.

Initially we consider the broad-scale dynamics of the cell population, i.e. population growth. The change of cell population can be represented by a population growth curve, with time on the x-axis and number of cells on the y-axis. By comparing the shape of two growth curves, we are able to compare population dynamics in *in vitro* experiments with *in silico* results. This allows us to focus on our cell cycle model parameters alone, i.e. the state transition probabilities. In our next analysis, and in a subsequent publication, we will carry out analysis of the spatial patterning of cells over time in order to determine parameter values for the strength of cell-cell interactions.

3.7 Summary

To summarise, a number of assumptions from Model 1 still hold:

- Only contact force, adhesion force and resistance force are significant (Domain).
- Cells can be implemented as ellipsoids (Platform Model).
- Matter is evenly distributed within a cell (Platform Model).
- Contact force can be computed using the Perram-Wertheim approach (Platform Model).
- Adhesion force can be modelled as a step function on distance (i.e. the growth factor gradient does not have a significant effect on attractive force) (Platform Model).
- Contact and adhesion forces balance at a defined point when cells are in contact, and the strengths of the forces can be calibrated based on this (Platform Model).

- Cells move at very low speed so their acceleration approaches zero and the forces upon them are balanced (Platform Model).
- Resistance force can be computed using Stokes' law, and the known properties of the fluid medium (Platform Model) (Platform Model).
- Interactions with the substrate can be modelled as interactions with mirrored agents (Platform Model).

Other assumptions no longer hold:

- Cell size and shape *do* change during the experiment (Domain Model). In our *in vitro* data it is clear that cell size and shape changes as cells progress through the cell cycle – we must accommodate this.
- Cells *may* divide or die during the timeframe of our experiment (Domain Model). We are modelling behaviour over much longer time frames and so need to account for the cell life cycle and cell death.

In addition to these two changes, the only other substantial change is the need to amend the implementation of the detailed calculations of interaction potentials and their resulting forces and torques, since these must now take changes to cell size and shape into account.

4 2D Growth Experiment and Model Calibration

Now we have a framework in place that allows us to reuse the physical model within new research contexts, the next step is to use it along with wet-lab experimentation to calibrate it. In the following sections, we report on experimental set up and on model calibration, and then discuss the problems arising.

4.1 Cell Line and Experiment

We designed and implemented four Petri-dish growth experiments using the HCT-116 cell line. HCT-116 cells are human colorectal cancer cells carrying the wildtype *TP53* gene and with intrinsic metastatic capacity (Céspedes et al, 2007); they are very suitable for 2D

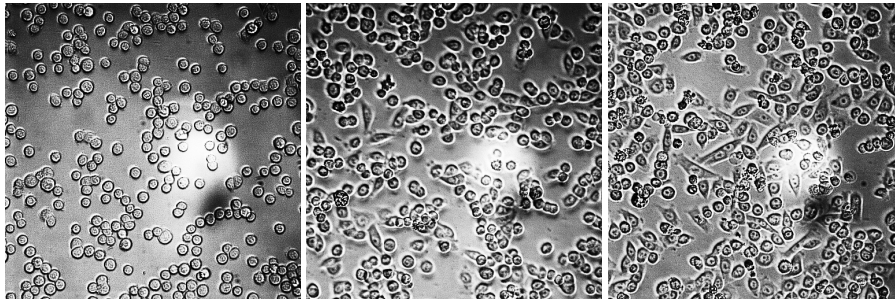


Fig. 16 Time-lapse images of HCT-116 cell growth under only 5-FU treatment conditions (Sect. 4.1), taken at starting point, 10 h and 20 h after planting

(monolayer) and 3D (spheroids) culture models. The cell line also harbours wildtype homozygous *thymidylate synthetase* gene (a key enzyme in DNA synthesis) but with reduced stability of both the transcriptional and translational products that may impair the response of HCT-116 cells to 5-FU. Further, the HCT-116 cells have wildtype *thymidine phosphorylase* (one of the enzymes system for the anabolic activation of 5-FU) and a low level expression mutant *dihydropyrimidine dehydrogenase* gene (a major enzyme for the catabolic inactivation of 5-FU).

The cells are planted in a glass bottom-flat well, 96-well tissue culture plate and left to grow for 120 h, with no treatment, 5-FU, hypoxia, and the combination of 5-FU and hypoxia, respectively. Each experiment starts with 20,000 cells, and the total number of cells in the population is measured every 24 h. In addition to the population count, we also took time-lapse photographs of the same area of the well every 10 min for the first 20 h (Fig. 16).

As a first step towards modelling the cells' spatial distribution, we need to simulate the cell population dynamics, i.e. the shape of the population growth curve. Several initial experiments were conducted to determine appropriate drug doses for experimentation. A large-scale *in vitro* growth experiment was then repeated three times under each of the four conditions; the results are shown in Fig. 17. This mix of different experimental conditions and replicates is sufficient to demonstrate plausible model calibration for our purposes.

Fig. 17 shows that in all four experiments the cell population decreases in the first 24 h; then for the control group and 5-FU group the cell population increases quickly. In the hypoxic condition, the cell population continues to decrease until 48 h after treatment, then

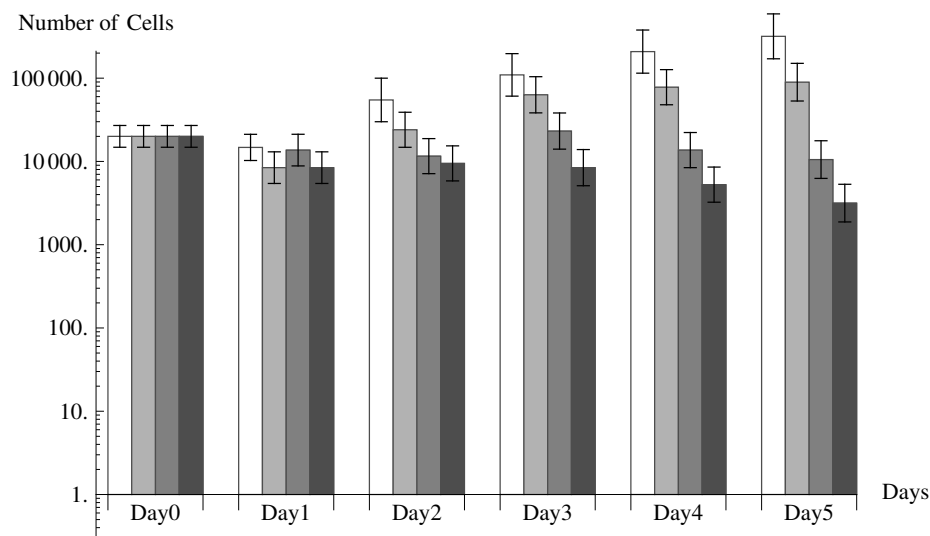


Fig. 17 Population count of HCT-116 cells in *in vitro* growth experiment (Sect. 4.1), measured at 24 h intervals under four different treatment conditions. Bars from left to right: with no treatment; with only 5-FU; with only hypoxia; with both 5-FU and hypoxia. (Note logarithmic Y scale.)

increases until 72 h but decreases thereafter. With the combined 5-FU and hypoxic regime, the cell population decreases for the first 24 h, then slightly increases for another 24 h and then decreases after that.

The time-lapse images are processed by CellProfiler (Carpenter et al, 2006) to find the outlines of each cell, giving not just the population over time, but also the positions and computed velocities of all the cells for the first 20 h. Based upon this data, we can estimate the typical time it takes for a cell to divide in this experimental setting, i.e. the cell cycle length, to be approximately 12 h.

4.2 Control Experiment Population Growth Curve

Our first step is to build a model that predicts the growth curve with neither 5-FU nor hypoxia. Note that during the experiment we sample the cell population five times at an interval of 24 h and take time-lapse photographs every 10 min; these time-lapse images cover only a small fraction of the cell population and so cannot be used to estimate total number of cells. As we do not know the changes in the total number of cells within each 24 h interval, we assume the population growth rate within every 24 h is constant. The population growth rate

represents the increase or decrease in cell population over a certain period of time: the value of the growth rate can be positive or negative.

Given the cell population at 0 h and 24 h, we can calculate the growth rate during this time with

$$N_0 e^{\mu t} = N_t \quad (2)$$

in which N_0 is the cell population at the starting time, N_t is cell population after time t , t is doubling time, and μ is the growth rate.

For a single cell, as stated in Sect. 3.4, we introduce the probability of cell division which is a real number between 0.0 (do not divide and so die) and 1.0 (always divide). From the experimental data we can calculate the population growth rate which is in the range $[-1, 1]$; for the model we require the probability of cell division. Thus a conversion is required: r represents the probability of cell division, and μ represents the growth rate:

$$r = \frac{\mu + 1}{2} \quad (3)$$

To reflect the growth condition between the sampling points, we calculate five sets of growth parameters, shown in Table 1, and the corresponding growth curve in Fig. 18.

4.3 Growth Curve Calibration

From Fig. 18, it is clear that the simulated population grows much more slowly than the experimental population. One possible source of error in our fitting is that our calculation depends on a constant doubling time – however, the doubling time may be affected by the environment, and changes over time in the experiment. From further simulation, we also know that simply increasing the maximum cell age parameter in the model cannot solve the problem.

In order to solve it, we have to avoid using a doubling time. From our definition of growth rate, we know that we only need cell population to be measured at two time points to find the average growth rate during the interval, which is independent of doubling time. Assuming the cell cycle is 12 h (as estimated from the time-lapse photography), which means

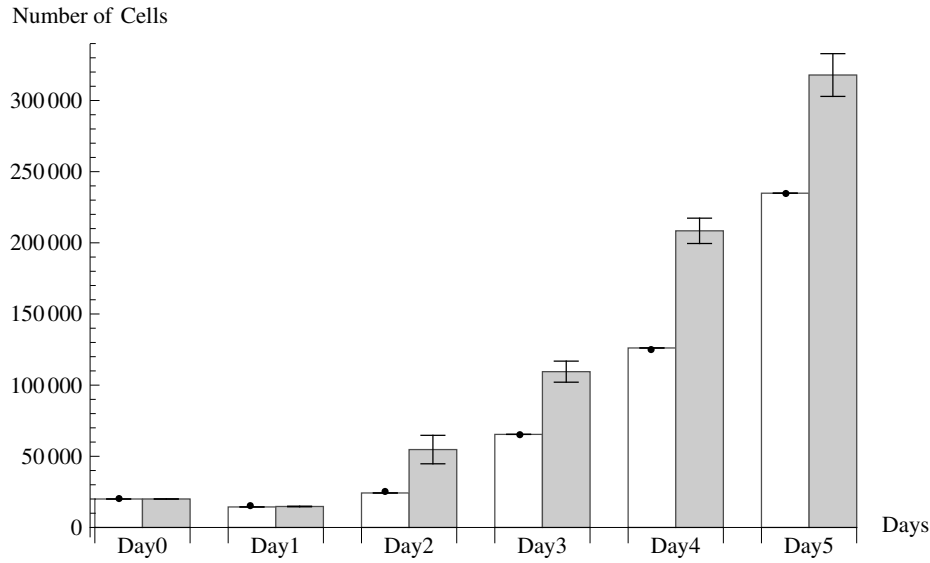


Fig. 18 Control group growth curve comparison before calibration, showing cell population measured at 24 h intervals. White bars: simulated cells. Grey bars: HCT-116 cells in *in vitro* growth experiment (Sect. 4.1), with no treatment.

that each measured value from the experiment contains two cycles, let x be the average growth rate during every 24 h, N_0 be the starting number, and N_t be the number at arbitrary time. Between time N_0 and N_1 there should be a N'_0 that satisfies

$$\begin{cases} N_0 \cdot x \cdot 2 = N'_0 \\ N'_0 \cdot x \cdot 2 = N_1 \end{cases} \quad (4)$$

Solving the equation we find:

$$x = \sqrt{\frac{N_1}{4N_0}} \quad (5)$$

In this way we find a new set of growth rate and probability of cell division, which are shown in Table 2 – note that they are larger values than in Table 1. They produce the growth curve in Fig. 19.

Fig. 19 shows that the simulated cell population curve is a good fit, both in values and shape, to the experimental results – although the simulated results are always slightly higher than the experimental results. This baseline model can now be extended for further *in silico* experimentation.

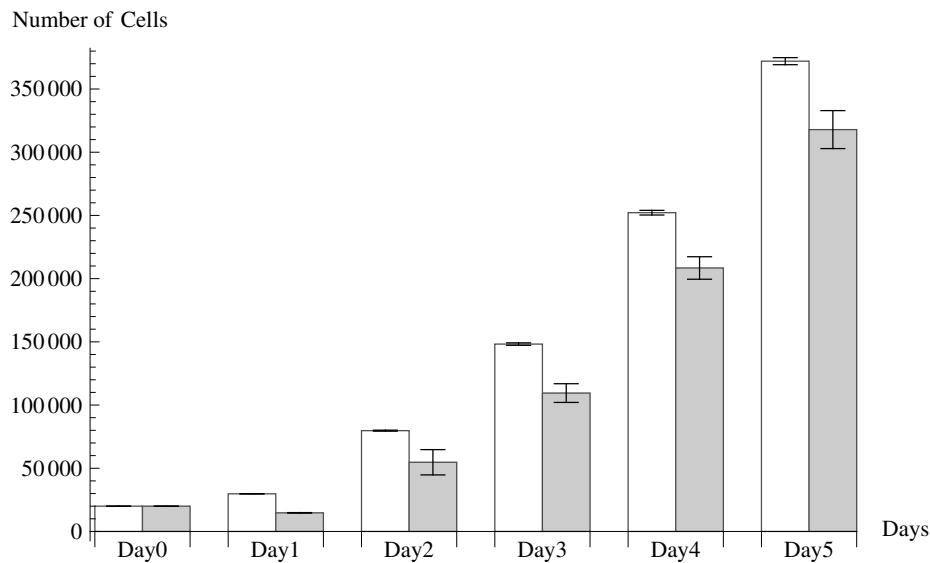


Fig. 19 Control group growth curve comparison after calibration, showing cell population measured at 24 h intervals. White bars: simulated cells. Grey bars: HCT-116 cells in *in vitro* growth experiment (Sect. 4.1), with no treatment.

4.4 Growth Curve with 5-FU

Instead of calculating a new value for probability of cell division from the 5-FU experiment data, we have chosen instead to model 5-FU action by the introduction of a plausible mechanism for 5-FU action – cell cycle arrest (Choudhary et al, 2012) – while otherwise keeping the model parameters the same as for the control group.

As we focus on the total number of the cells, we mainly study 5-FU's effect on cell growth rate. We add a new state to the state machine previously shown in Fig. 14: Arrested, as Fig. 20 shows. We do not change the growth rate or probability of cell division directly; instead, in the fate-determination phase, those cells that decide to divide are given a probability of entering the Arrested state. Arrested cells will not divide; instead, they die after a certain time. If the probability of arrest is too high, the total cell population will decrease, and if it is too low the growth curve will look like the control growth curve. We determined a good fit when the probability of arrest is set to the constant value of 5% (Fig. 21).

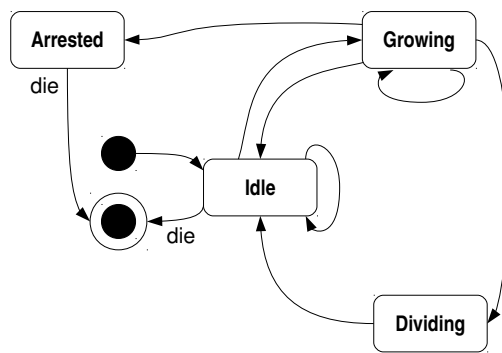


Fig. 20 5-FU model: cell cycle, shown as a state diagram

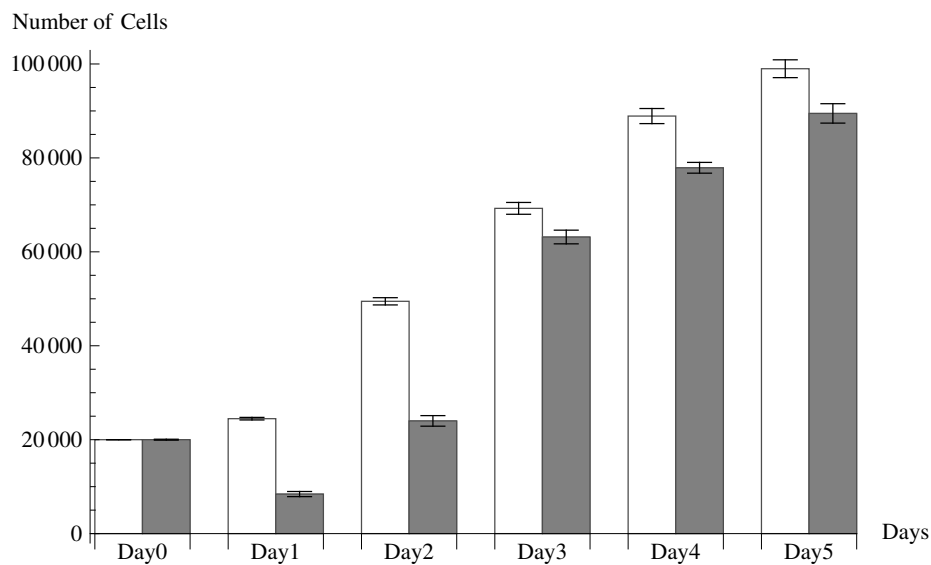


Fig. 21 5-FU group growth curve comparison after calibration, showing cell population measured at 24 h intervals. White bars: simulated cells. Grey bars: HCT-116 cells in *in vitro* growth experiment (Sect. 4.1), with only 5-FU treatment.

4.5 Growth Curve with Hypoxia

Next we consider hypoxia. The establishment, development, maintenance and progression of tumours creates unfavourable tumour microenvironments where the supply of oxygen and nutrients to cancer cells becomes very limiting and challenging. In response cancer cells learn to adapt (intrinsic and acquired) to these forms of environmental stress by putting in place signals and triggered responses to sustain hypoxic life (hypoxia). Tumour hypoxia

is now a recognised major driving factor for malignant progression and adverse prognostic marker among cancer patients. In either preclinical or clinical studies hypoxia has been implicated with poor response to both chemotherapy and radiotherapy. The hypoxia response pathway is governed by a family of hypoxia-inducible transcription factors (HIFs) that coordinate the adaptive response to hypoxia. HIF-1 is the most characterised family member, where two amino acid residues (Pro402 and Pro564) in the HIF-1 α subunit are subject to oxygen-dependent hydroxylation in the cytosol by prolyl hydroxylase (PHD). This modification renders HIF-1 α amenable to outright ubiquitination by the von Hippel Lindau (pVHL) tumour suppressor protein and degradation by the 26S proteasome. Under hypoxic conditions, however, PHD function is compromised and HIF-1 α escapes ubiquitination, accumulates in the cell and translocates to the nucleus, where it heterodimerises with HIF-1 β and occupies hypoxia response elements (HREs) of the key genes in cellular proliferation, survival, energy metabolism, angiogenesis, invasion and metastasis to activate and up-regulate their transcription. Thus considerable efforts have lately focused on the modulation of HIF-1 α to strategically inhibit tumour growth and development (including invasion and metastasis) and also to sensitise tumours to chemotherapeutic agents like 5-FU (Yoshikawa et al, 2001; C  spedes et al, 2007; Choudhary et al, 2012).

We consider the effect of hypoxia as reduced probability of cell division. Thus to introduce a hypoxic effect in the model we need to reduce the probability of cell division in all five phases. We assume that cells are affected equally during the whole experiment; therefore we have to amend the probability of cell division in the same way. As described in Sect. 4.2, the increase in cell population requires that the probability of cell division be higher than 0.5, and decrease of cell population requires that it be lower than 0.5. From Fig. 17, we know that the value by which we may decrease the probability of cell division is in the range $[0.1899, 0.20711]$. We choose the value 0.2 for simplicity and conciseness, giving the new probabilities of cell division for hypoxia affected cells shown in Table 3.

Even without simulation, we can see from the probability of cell division that the new growth curve cannot match the experiment during 24–48 h, because the probability is higher than 0.5 and will cause an increasing cell population, while the cell population decreases

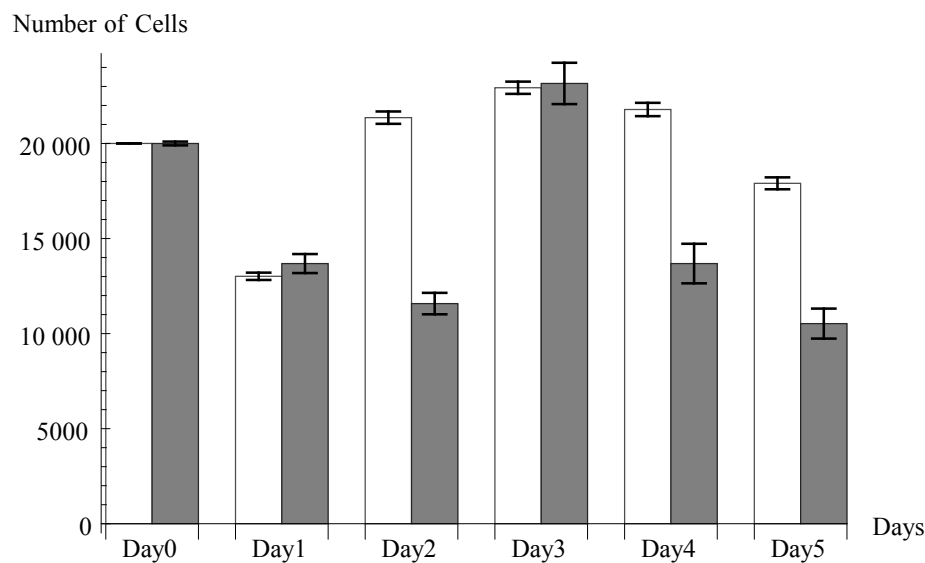


Fig. 22 Hypoxia group growth curve comparison after calibration, showing cell population measured at 24 h intervals. White bars: simulated cells, with the cell cycle length set to 12 h for Type 1 cells and 24 h for Type 2 cells. Grey bars: HCT-116 cells in *in vitro* growth experiment (Sect. 4.1), with only hypoxia treatment.

during this time in the experiment. We cannot account for this with a homogeneous population; we need to introduce a degree of heterogeneity into the model, and cell heterogeneity leading to differential responses to drug action and environment conditions is a recognised phenomenon (Bown et al, 2012).

We introduce heterogeneity in a minimal way, by assuming that the hypoxia also affects the cell cycle: we introduce a second type of cell into the model. The original cell type (*Type 1*) maintains the existing cycle length, but has an increased probability of cell division; the new cell type (*Type 2*) has a longer cell cycle but uses the original probability of cell division, shown in Table 2. By tuning the cell cycle length of Type 2, and adjusting the proportion of these two types, we are able to delay the time at which the cell population begins to increase. Table 4 shows the probability of cell division settings we are using; Fig. 22 shows the resulting growth curve.

We can see from Fig. 22 that, although for most time points the simulated cell population level is close to the experimental results, it is still much larger than the experimental observation at the 48 h (Day 2) point. As all the probability of cell division values must be

altered equally, to solve the problem all the values should be decreased until the value for 24–48 h is lower than 0.5, which will make the probability of cell division for 0–24 h lower than zero and out of its possible range, and this means that this problem cannot be solved by tuning the probability of cell division.

A possible solution to this fitting problem is to tune the cycle length of the new cell type, until it remains in its first three cycles between 24 and 48 h. As the old cell type is in 4th–6th cycle during that time, the combination of probability of cell division of both types may fit the simulation growth curve to the experimental one. This requires further work to determine the proper cycle length and proportion of the two types.

4.6 Growth Curve with Combination of 5-FU and Hypoxia

To independently test our modelling approach used to represent 5-FU and hypoxia mechanisms separately, we consider them acting together in the model and compare the simulation output with the experiment. This means that we keep the two types of cells and their cell cycle and probability of cell division setting with which we simulate the effect of hypoxia, combined with the probability of cell cycle arrest which is used in the 5-FU affected simulation. There is no new mechanism added and no data from experiment of combination of 5-FU and hypoxia is used for model calibration.

Fig. 23 shows good agreement between simulated and experimental results, where with no other model adjustment we are able to predict periods of population growth and decline. We note some differences in the rates of decline and increase in the earlier stages of the simulation. According to previous simulation results, this difference is most likely driven by the hypoxia effect, because the 5-FU simulation tends to have similar dynamics over all time points, while the hypoxia simulation has higher variability during the period of most difference in Fig. 23. We can see that the combined effects show a reasonably good fit to the experimental results, and we have reason to believe that, by improving the fit of the hypoxia simulation, we are able to improve the combined simulation at the same time.

Table 1 Growth parameters determined for untreated cells

Hours	0–24	24–48	48–72	72–96	96–120
Growth rate (μ)	-0.1525	0.656	0.3465	0.322	0.211
Probability of cell division (r)	0.42375	0.828	0.67325	0.661	0.6055

Table 2 Corrected growth parameters after calibration

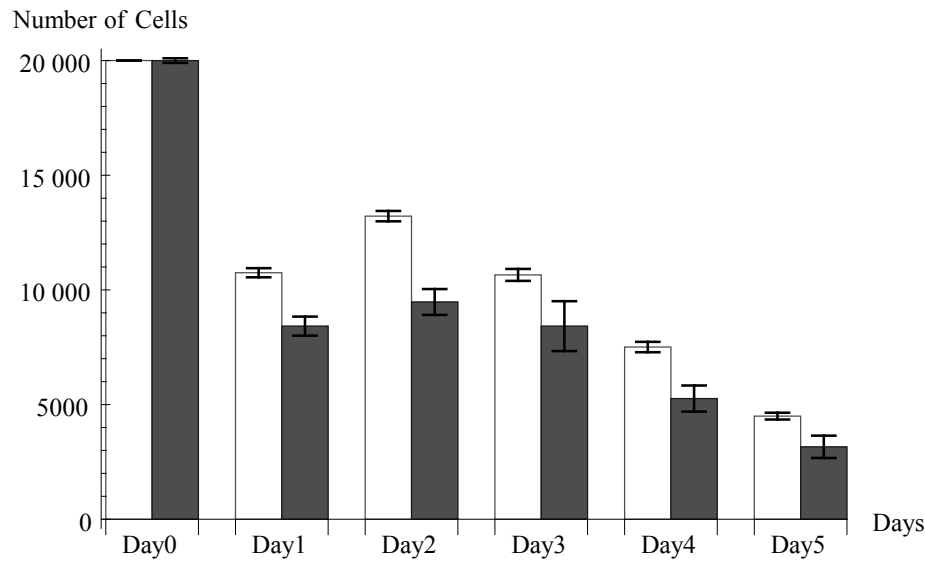
Hours	0–24	24–48	48–72	72–96	96–120
Growth rate (μ)	-0.1416	0.92724	0.41422	0.3798	0.23502
Probability of cell division (r)	0.42920	0.96362	0.70711	0.68990	0.61751

Table 3 Cell division probabilities for hypoxic cells

Hours	0–24	24–48	48–72	72–96	96–120
Original probability of cell division (r)	0.42920	0.96362	0.70711	0.68990	0.61751
New probability of cell division (r')	0.22920	0.76362	0.50711	0.48990	0.41751

Table 4 Cell division probabilities for two-population hypoxia model

Hours	0–24	24–48	48–72	72–96	96–120
Probability of Type 1 cell division (r_1)	0.22920	0.76362	0.50711	0.48990	0.41751
Probability of Type 2 cell division (r_2)	0.32920	0.86362	0.60711	0.58990	0.51751

**Fig. 23** Combined 5-FU/hypoxia group growth curve comparison after calibration, showing cell population measured at 24 h intervals. White bars: simulated cells. Grey bars: HCT-116 cells in *in vitro* growth experiment (Sect. 4.1), with combined 5-FU and hypoxia treatment.

5 Conclusion

Through initial analysis guided by the CoSMoS project structure, we have identified that the physical aspects of the new domain do indeed have considerable similarities with those of the original domain – so reuse of the physical model should be appropriate, provided that the assumptions in the model – which we have explicitly identified – are re-evaluated appropriately within the new context.

Other aspects of the two domains are substantially different; for example, the cancer simulation requires an implementation of the cell cycle in order to accurately simulate the effects of different treatments and environmental conditions, whereas this was unnecessary in the vascular simulation owing to the initial stage limitation.

We have also identified the gaps in our knowledge about the domain, necessary to appropriately calibrate the physical model, which are now being filled by wet-lab experimentation. This gives us confidence that we are asking the right questions when conducting experiments in support of calibration.

What we have ended up with is emphatically *not* “a CoSMoS project” – we have simply made use of a few aspects of CoSMoS to structure our thinking about model reuse. Stepney (2012) considers CoSMoS as based upon a pattern language that should be tailored to meet the needs of the project to which it is being applied. Moreover, Stepney (2012) notes that CoSMoS may be applied to some phases of a project, provided that the impact of the assumptions made during other key phases is properly understood. In line with this view, we have only made use of some of the large-scale CoSMoS patterns that describe concepts such as “domain model”, and that in a rather sketchy and informal way – but we feel that even this first step towards CoSMoS has been valuable in terms of forcing us to think in a principled way about our existing work. As this work continues, we intend to make increased use of CoSMoS techniques; for example, to more effectively structure our interactions with domain experts during the calibration of the spheroid model. We feel, in general, that the ability to adopt patterns as appropriate is a significant strength of the CoSMoS approach in terms of adoption by existing projects – as it is for other pattern languages.

While it is important to emphasise that this project is still work in progress, we feel that we have achieved a satisfactory degree of model and software reuse – and, more importantly, we are confident that this reuse has been achieved in a way that is *appropriate* and *useful* within our new research context. In the future, we would like to consider strategies and patterns for this kind of reuse within the CoSMoS process – in particular, how a validity argument might be constructed and updated as a model is reused.

In addition, by documenting this process, we have a framework in place that would allow us to reuse the physical model within new research contexts in future projects. This framework has been used in conjunction with experimental results that are based on a clinically relevant cell line – HCT-116 – and the cell line’s response to both 5-FU drug action and hypoxia. We fitted the model to data from a control condition, and then added in biologically plausible mechanisms for both drug action and hypoxia. We tested the model against independently measured data from the combination of 5-FU and hypoxia. In all cases we had good agreement between experimental and simulated growth population curves.

These results are a first step to modelling cell responses to environmental perturbations. Our modelling approach represents cells as individual agents with a cell cycle and each cell may have its own transition probabilities between the states in that cell cycle. Fitting to 5-FU required us to amend the cell cycle probabilities. Fitting to hypoxia required us to construct heterogeneous populations of cells. In each case we were able to interpret model changes in terms of the underlying biology and in this way propose mechanisms of action. Such mechanisms could ultimately be tested through further experimentation.

We also believe that the model is ready to be used for further study. Our next step will be to attempt to reproduce, in our 2D model, the spatial distribution of cells seen in experiments, using for calibration the data derived from time-lapse images of experiments. The radial distribution function mentioned in Sect. 2.6 will continue to be used as a quantitative measure when comparing experimental and simulated results. During adaptation of our model, we have identified four key parameters that affect the shape of the radial distribution function curve, and will calibrate these to match the experimental results. The calibrated model will then be used to improve our understanding of the effects of drugs and hypoxic

conditions upon cancer cells through the implementation of hypothesised mechanisms for these effects within the low-level agents in the model.

Acknowledgements

The authors would like to thank their colleagues who kindly provided resources for this paper. David Harrison and Peter Mullen (University of St. Andrews) provided the HCT-116 cells used for the experiments in Sect. 4. Fig. 10 appears courtesy of Simon Langdon (University of Edinburgh). Fig. 15 appears courtesy of Peter Caie (University of Edinburgh). James Bown acknowledges support from the Northwood Trust.

References

- Andrews PS, Polack FAC, Sampson AT, Stepney S, Timmis J (2010) The CoSMoS process version 0.1: A process for the modelling and simulation of complex systems. Tech. Rep. YCS-2010-453, Department of Computer Science, University of York
- Batchelor GK (2000) An Introduction to Fluid Dynamics. Cambridge University Press
- Bown J, Andrews PS, Deeni Y, Goltsov A, Idowu M, Polack FAC, Sampson AT, Shovman M, Stepney S (2012) Engineering simulations for cancer systems biology. *Current Drug Targets* 13(14)
- Brú A, Albertos S, Subiza JL, García-Asenjo JL, Brú I (2003) The universal dynamics of tumor growth. *Biophysical Journal* 85(5):2948–2961
- Carpenter A, Jones T, Lamprecht M, Clarke C, Kang I, Friman O, Guertin D, Chang J, Lindquist R, Moffat J, Golland P, Sabatini D (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biology* 7(10):R100
- Cavalli F, Gamba A, Naldi G, Semplice M, Valdembri D, Serini G (2007) 3D simulations of early blood vessel formation. *Journal of Computational Physics* 225(2):2283–2300
- Céspedes MV, Espina C, García-Cabezas MA, Trias M, Boluda A, del Pulgar MTG, Sancho FJ, Nistal M, Lacal JC, Manges R (2007) Orthotopic microinjection of human colon

- cancer cells in nude mice induces tumor foci in all clinically relevant metastatic sites. *Am J Pathol* 170(3):1077–1085
- Choudhary B, Hanski M, Zeitz M, Hanski C (2012) Proliferation rate but not mismatch repair affects the long-term response of colon carcinoma cells to 5FU treatment. *Cancer Lett* 320(1):56–64
- Étienne M (2006) Companion modelling: a tool for dialogue and concertation. In: *Biodiversity and stakeholders: concertation itineraries*, Editions Quae, pp 44–52
- Folkman J, Haudenschild C (1980) Angiogenesis in vitro. *Nature* (288):551–556
- Gamba A, Ambrosi D, Coniglio A, de Candia A, Di Talia S, Giraudo E, Serini G, Preziosi L, Bussolino F (2003) Percolation, morphogenesis, and burgers dynamics in blood vessels formation. *Phys Rev Lett* 90:118,101
- Goldstein H, Poole CP, Safko JL (2001) *Classical Mechanics* (3rd Edition), 3rd edn. Addison-Wesley
- Holmes MJ, Sleeman BD (2000) A mathematical model of tumour angiogenesis incorporating cellular traction and viscoelastic effects. *Journal of Theoretical Biology* (202):95–112
- Kaida A, Miura M (2012) Visualizing the effect of hypoxia on fluorescence kinetics in living HeLa cells using the fluorescent ubiquitination-based cell cycle indicator (Fucci). *Exp Cell Res* 318(3):288–297
- Karamysheva AF (2008) Mechanisms of angiogenesis. *Biochemistry (Moscow)* 73:751–762
- Kimlin LC, Casagrande G, Virador VM (2013) In vitro three-dimensional (3D) models in cancer research: An update. *Molecular Carcinogenesis* 52(3):167–182
- Morgan DO (2007) *The cell cycle: principles of control*. New Science Press
- Niles L (2012) Human cultured neural stem cells.
<https://www.youtube.com/watch?v=x_e3PEJgrFY>
- Palsson E (2001) A three-dimensional model of cell movement in multicellular systems. *Future Gener Comput Syst* 17(7):835–852
- Papageorgis P, Cheng K, Ozturk S, Gong Y, Lambert AW, Abdolmaleky HM, Zhou JR, Thiagalingam S (2011) Smad4 inactivation promotes malignancy and drug resistance of colon cancer. *Cancer Res* 71(3):998–1008

- Perram JW, Rasmussen J, Præstgaard E, Lebowitz JL (1996) Ellipsoid contact potential: Theory and relation to overlap potentials. *Phys Rev E* 54:6565–6572
- Reinders J (2007) Intel Threading Building Blocks: Outfitting C++ for Multi-core Processor Parallelism. O'Reilly & Associates, Inc.
- Savage A, Katz E, Eberst A, Falconer RE, Houston A, Harrison DJ, Bown J (2013) Characterising the tumour morphological response to therapeutic intervention: an ex vivo model. *Disease Models & Mechanisms* 6(1):252–260
- Serini G, Ambrosi D, Giraudo E, Gamba A, Preziosi L, Bussolino F (2003) Modeling the early stages of vascular network assembly. *EMBO J* 22(8):1771–1779
- Stepney S (2012) A pattern language for scientific simulations. In: *Proceedings of the 2012 Workshop on Complex Systems Modelling and Simulation*, Orleans, France, September 2012, Luniver Press, pp 77–103
- Stepney S, et al (2015) Engineering simulations as scientific instruments. Springer, to appear
- Szabó A, Ünnep R, Méhes E, Twal WO, Argraves SW, Cao Y, Czirók A (2011) Collective cell motion in endothelial monolayers. *Phys Biol* 7(4):046,007
- Turner S, Sherratt JA (2002) Intercellular adhesion and cancer invasion: A discrete simulation using the extended Potts model. *Journal of Theoretical Biology* 216(1):85–100
- Welter M, Barthä K, Rieger H (2009) Vascular remodelling of an arterio-venous blood vessel network during solid tumour growth. *Journal of Theoretical Biology* 259(3):405–422
- Yoshikawa R, Kusunoki M, Yanagi H, Noda M, Furuyama Ji, Yamamura T, Hashimoto-Tamaoki T (2001) Dual antitumor effects of 5-fluorouracil on the cell cycle in colorectal carcinoma cells: A novel target mechanism concept for pharmacokinetic modulating chemotherapy. *Cancer Res* 61(3):1029–1037
- Zhu J, Coakley S, Holcombe M, MacNeil S, Smallwood R (2006) Individual cell-based simulation of 3D multicellular spheroid self-assembly. *Eur Cells Mater* 11(Suppl 3):31