Combining experiments with multi-cell agent-based modeling to study biological tissue patterning

Bryan C. Thorne, Alexander M. Bailey and Shayn M. Peirce

Submitted: 31st March 2007; Received (in revised form): 24th May 2007

Abstract

Agent-based modeling (ABM), also termed ‘Individual-based modeling (IBM)’, is a computational approach that simulates the interactions of autonomous entities (agents, or individual cells) with each other and their local environment to predict higher level emergent patterns. A literature-derived rule set governs the actions of each individual agent. While this technique has been widely used in the ecological and social sciences, it has only recently been applied in biomedical research. The purpose of this review is to provide an introduction to ABM as it has been used to study complex multi-cell biological phenomena, underscore the importance of coupling models with experimental work, and outline future challenges for the ABM field and its application to biomedical research. We highlight a number of published examples of ABM, focusing on work that has combined experimental with ABM analyses and how this pairing produces new understanding. We conclude with suggestions for moving forward with this parallel approach.

Keywords: agent-based modeling; tissue patterning; microcirculation; morphogenesis; regenerative medicine

INTRODUCTION

Although systems biology has traditionally focused on the study of gene and protein interactions at the single-cell level, investigations are increasingly being conducted at the multi-cell level using agent-based modeling (ABM), a technique that has established roots in ecology and sociology. While differential equation-based models have also been used to investigate multi-cellular and tissue level phenomena, ABM, derived from cellular automata, explicitly represents individual cells in space and time. This modeling approach provides information about how tissue patterns emerge as a result of cellular interactions within the framework of the tissue-level environment. Tissue patterning events, which have implications for both physiological and pathological function, arise from a cascade of complex processes and rely on interactions between cells, genomic information, and intra-cellular signaling. These interactions are important in processes as diverse as tumor formation [1], microvascular response to myocardial ischemia [2], and bone tissue regeneration [3]. ABM, which can compute cell behaviors while integrating cell signaling events, is useful not only for probing mechanisms underlying the complex biology of tissue patterning, but also for its ability to facilitate high-throughput in silico experiments in a cost- and time-efficient manner.

The underlying philosophy of ABM is that individual agents respond to a set of rules and

Corresponding author. Bryan C. Thorne, Department of Biomedical Engineering, University of Virginia, PO Box 800759, Charlottesville, VA 22908, USA. Tel: 434-243-9892; Fax: 434-982-3870; E-mail: bthorne@virginia.edu
Bryan C. Thorne is an MD/PhD student in the Department of Biomedical Engineering at the University of Virginia. He is studying the role of inflammation in atherosclerosis using a combined agent-based modeling and experimental approach.
Alexander M. Bailey is a PhD student in the Department of Biomedical Engineering at the University of Virginia where he is studying the trafficking of adult human-derived stem cells in tissue ischemia using a combined agent-based modeling and experimental approach.
Shayn M. Peirce is an Assistant Professor in the Department of Biomedical Engineering at the University of Virginia. She received her BS in Biomedical Engineering from the Johns Hopkins University in 1997 and her PhD in Biomedical Engineering from the University of Virginia in 2002. Dr Peirce’s research focuses on developing in vivo and computational techniques to study tissue patterning processes.
make ‘decisions’ that lead to a particular behavior, and the aggregate of agent behaviors over time produces ‘emergent’ phenomena. Simulation of the interactions of a collection of agents responding to local environmental conditions and neighboring agents results in the production of higher-level patterns, and inclusion of multiple types of agents allows explorations of more complex behaviors. In this way, ABMs behave in a similar manner to biological tissues, whose functions are essentially the output of cell behaviors that collectively generate tissue–level patterns built from the bottom up (e.g. embryogenesis). Because the outputs of an ABM are highly dependent on empirical rules, it is necessary to couple models with laboratory experimentation at all stages of model development and to validate an ABM’s rule set by performing iterative in silico and in vivo/in vitro experimentation.

The purpose of this review is to provide an introduction to ABM as it has been applied to study complex multi-cell biological phenomena, to underscore the importance of coupling models with experimental work, and to outline future challenges for the agent-based modeling field and applications in biomedical research. Toward this end, we highlight a number of published examples of ABMs, focusing on work that has combined experimental and ABM analyses, and conclude with suggestions for moving forward with this parallel approach. Because it is a difficult task to completely communicate an ABM in a single experimental paper, it would be impossible to fully explore the depth of the models discussed herein. Therefore, we refer the interested reader back to the primary literature for more detailed information.

Overview of ABM

Since ABM has not achieved widespread use in biomedical research, we first present the underlying philosophy and describe the discrete nature of ABM and its fundamental components: agents, rules, simulation spaces and boundary conditions. Agents are autonomous entities that interact with their environment and other agents according to a set of literature-based rules. These local interactions give rise to complex global patterns of agent behaviors, termed ‘emergent phenomena’. The rule set governs each decision made by individual agents, and as agents interact with one another and with their environment, their state may change. Agent states are recorded as local variables, and these variables can be monitored in time to describe the history of states exhibited by any agent or set of agents. States can include any type of cell-level information, such as protein expression levels, cell activation states or cell–cell contact status (Figure 1). Agents are mobile within the simulation space and record not only their own states, but also their spatial positions over the course of the simulation; this allows investigation of how the system evolves and reorganizes over time.

The rules governing agent behaviors are generally procured from the literature or derived experimentally, although where published data or experimental techniques are unavailable they can be systematically parameterized. Each rule describes an agent’s reaction to a set of specific environmental conditions, its neighbor’s state, and/or its own internal state. Because, for a cellular-level agent, internal cellular machinery is usually ignored, and literature from which rules are derived is not always explicit, rules often incorporate a measure of stochasticity. For example: ‘If an agent of type “Endothelial Cell” encounters a threshold concentration of “TNF-Alpha”, it has a 60% chance of responding by up-regulating “VCAM-1” surface expression.’ [4].

In biomedical research, it is most common to represent an individual cell as a single agent in two-dimensions [4–9], although models at the scale of protein or even organ system interaction have also been developed [10]. The simulation space represents the cell’s external environment, which contains relevant non-cellular parameters for the experimental question being addressed. For example, it may include concentrations of diffusible factors such as chemokines or cytokines, or nondiffusible factors such as the extracellular matrix [4, 8]. When using a ‘cells-as-agents’ paradigm, where one agent represents one cell, agents can be programmed to exhibit a range of biologically relevant behaviors, including proliferation (mitosis), apoptosis (cell death), migration (cell movements), and differentiation (expression of a set of proteins characteristic of a different phenotype) [8, 11] (Figure 1). The simulation space in which the agents move can be given closed, open or periodic boundaries. With closed boundaries, agents cannot move past the boundaries, and proteins cannot diffuse beyond boundaries [6]. In this case, the simulation area must be sufficiently large relative to the area of agent interactions to avoid edge effects. Conversely, open conditions remove cells or protein from the simulation entirely when they pass beyond the boundaries (i.e. total flux of agents is not a priori
set to zero) [4]. Periodic (also known as toroidal) conditions are used when modeling a representative section of a larger tissue, wrapping flow across the boundaries to the opposite side so that an agent exiting the simulation through the left-hand border would automatically re-enter the simulation space from the right-hand border [9]. Boundary conditions must be carefully chosen depending on the goals of the simulation, and can differ between agents and other quantities within the same space. For example, quantities of proteins may be assumed to approach a background tissue level as they approach the simulation boundaries (Dirichlet or open boundaries), while periodic boundary conditions could be used for agents when modeling a random walk pattern of migration in a portion of a larger tissue—it is a reasonable assumption that if cells are migrating randomly, and one cell leaves the simulation space, another would enter from adjacent tissue [8].

The time-scale of an ABM simulation needs to be tailored to the specific biological hypothesis addressed. Simulations move through time in a series of discrete time steps that can be anywhere on the order of milliseconds [10] to hours [8, 12], and simulations can encompass years [12]. At each time step, a core set of processes is repeated, enabling agents to survey the environment, make ‘decisions’ that are dictated by the rule set, and exhibit ‘behaviors’ accordingly. The choice of how each time step is scaled is an important decision, and drives the type of questions an individual model can answer.

**ABMs in biomedical research**

While ABM has been embraced for a long time by ecologists and social scientists (for review, see Grimm and Railsback’s book ‘Individual-Based models in Ecology’ [13] or Elliott and Kiel, ‘Agent-based modeling in the social and behavioral sciences’ [14]), it has only recently been applied to biomedical problems. The usefulness of ABMs for investigation of higher level patterns originating from local agent-environment interactions has made them ideal for investigating tissue patterning events. Some of the areas in which ABMs have been applied are summarized in Table 1.

One of the key areas in biomedical research that has employed the use of the ABM approach has been cancer biology. Tumorigenesis is a prime example of an emergent tissue patterning phenomenon reliant on many cell types, both aberrant and normal cell behaviors, and atypical gene expression; ABM is well
**Table 1:** Summary of some published biomedical agent-based models. Each summarized model is classified by biological system or disease and presented with the motivation for its construction and key results

<table>
<thead>
<tr>
<th>Category</th>
<th>Model</th>
<th>Motivation</th>
<th>Key Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>CancerSim, an ABM implementing cell phenotypes corresponding to the ‘Hallmarks of Cancer’ [30]</td>
<td>Investigation of how heterogeneous cell populations acquire the hallmarks of cancer, which mutation pathway could provide the shortest pathway to a cancer phenotype</td>
<td>Compared to an ODE model of the same process, the ABM allows more realistic and in depth investigation.</td>
</tr>
<tr>
<td></td>
<td>Expanded multi-scale ABM from [40]. Simulated different expression levels of EGFR</td>
<td>Examine the influence of experimentally measured EGFR expression values on protein expression, cell interactions and emergent whole tumor dynamics</td>
<td>Model suggests that both proliferative and migratory phenotypes are necessary for rapid tumor expansion, underscores the importance of posttranslational regulation of protein expression. Both are experimentally testable.</td>
</tr>
<tr>
<td></td>
<td>Multi-scale model: integration of prior tumor ABM [44] with sub-cellular network model of EGFR signaling</td>
<td>Proof of concept, investigate the potential mechanisms behind cancer cell phenotype switching from proliferation to migration</td>
<td>Able to model proliferation versus migration decision based solely on intracellular gene network. Coupling model with experiments will allow hypothesis generation and testing.</td>
</tr>
<tr>
<td>Immunology</td>
<td>SimTriplex, a model to test different protocols for vaccination against cancer</td>
<td>Experimental testing of all possible protocols too time-consuming and expensive</td>
<td>Maximal protection could be obtained with 40% fewer vaccinations.</td>
</tr>
<tr>
<td></td>
<td>Used evolutionary game theory in a previously developed brain tumor ABM [44] to examine feedback effects between tumors and environment</td>
<td>Increased understanding of the biological system for hypothesis generation, save time and money by model-driven targeted experimentation</td>
<td>Highly malignant phenotypes can exist without detectable structure changes. Based on this tumor heterogeneity, the authors recommend multiple biopsies to characterize a tumor’s true state.</td>
</tr>
<tr>
<td></td>
<td>ABM of T-Cell/Antigen presenting cell junction to examine the influence of different mixtures of nonagonist peptides presented by APCs on T-Cell Receptor activation</td>
<td>Complexity inherent in possible mixtures of hundreds of peptides not experimentally tractable, and precise mixtures of peptides presented by APCs not controllable</td>
<td>New hypotheses: Small changes in peptide background may significantly change response, self peptides may contain agonists and antagonist peptides which modulate response to foreign peptides.</td>
</tr>
<tr>
<td>Immunology</td>
<td>ABM of granuloma formation in the lung during M. Tuberculosis infection</td>
<td>Interested in examining possible mechanisms controlling formation of stable versus necrotic granulomas</td>
<td>Model predicts importance of T-Cell spatial location, bacterial growth rate, chemokine diffusion rates, and migration rates of leukocytes.</td>
</tr>
<tr>
<td></td>
<td>Initial ABM of germinal center zone formation</td>
<td>Tests several hypotheses regarding the formation of light and dark zones in lymph node germinal centers</td>
<td>Predicted possible scenarios for germinal center zone formation and the function of the different zones.</td>
</tr>
<tr>
<td></td>
<td>Builds on previous ABMs of B-Cell germinal center reactions, tests possible mechanisms of germinal center deregulation.</td>
<td>Investigate several hypotheses about deregulation of germinal center reactions and the possible contribution of each to lymphoma formation</td>
<td>Deregulation of T-Cell induction of B-Cell differentiation implicated as a possible mechanism.</td>
</tr>
<tr>
<td></td>
<td>Extended an ABM describing possible mechanisms of B-Cell affinity maturation in lymph node germinal centers</td>
<td>Investigate the utility of several hypotheses about affinity maturation</td>
<td>Model suggests that a centrocyte refractory time between surveying dendritic cells and competition for T-Cell help are mechanisms that would produce experimentally observed patterns.</td>
</tr>
<tr>
<td>Clinical</td>
<td>ABM of innate immune response (IIR) including endothelial damage from bacterial infection and inflammation.</td>
<td>Introduce ABM to medical community, demonstrate utility for investigation of IIR and development of treatments for SIRS/MOF</td>
<td>Validated model by predicting results of cytokine-based clinical trials, and predicted results of hypothetical cytokine treatment protocols for SIRS/MOF</td>
</tr>
</tbody>
</table>

(Continued)
### Table 1: Continued

<table>
<thead>
<tr>
<th>Category</th>
<th>Model</th>
<th>Motivation</th>
<th>Key Results</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development</td>
<td>ABM of Xenopus Laevis epithelial cell morphogenesis and blastocoel roof thinning</td>
<td>Test a hypothesis regarding fibronectin fibril assembly and differential cell adhesion during blastocoel roof thinning</td>
<td>Predicted a pattern generated by the hypothesis; the same experimental manipulation performed in the laboratory produced a similar pattern.</td>
<td>[6]</td>
</tr>
<tr>
<td>Vascular</td>
<td>Modeled arteriogenesis and angiogenesis in response to vascular network pressure changes or exogenous VEGF application</td>
<td>Study how collections of different known growth factors impact microvascular remodeling in two different settings: a disease setting (ischemia) and a therapeutic setting (exogenous growth factor treatment)</td>
<td>Identified a functional module of three different growth factors capable of producing microvascular tissue patterning changes observed in independent experiments</td>
<td>[8]</td>
</tr>
<tr>
<td>Bone</td>
<td>ABM of monocyte trafficking in skeletal muscle microvasculature, integrated with network blood flow simulation</td>
<td>Proof of concept for integration of ABM with network blood flow model, design a preliminary model to examine leukocyte capture, adhesion and extravasation</td>
<td>Successful integration of fluid flow based analysis with ABM, recapitulated prior in vivo observations</td>
<td>[4]</td>
</tr>
<tr>
<td>In Vitro</td>
<td>Simulated signal propagation in osteocyte networks in response to loading induced stresses</td>
<td>Explore possible mechanisms behind recent in vivo observation that rest periods inserted between loading cycles enhance bone formation</td>
<td>Predicted that rest allows time for recovery of intracellular molecular stores, which increases the average signal</td>
<td>[27]</td>
</tr>
<tr>
<td></td>
<td>Discrete markov model validated with parallel in vitro studies using time-lapse microscopy</td>
<td>'Cell packing' phenomena was poorly understood, and understanding of process could explain dysplastic tissue architecture in cancers, cell shape, and morphogenesis</td>
<td>Predict that irregular epithelial cell packing patterns are due to cell proliferation, not 'efficient cell packing', and is not random</td>
<td>[19]</td>
</tr>
<tr>
<td></td>
<td>Very simple rule-set to investigate in vitro proliferation and differentiation of keratinocytes, given different culture conditions, although does not account for intra-cellular signaling</td>
<td>Explore the influence of calcium levels on cell growth and differentiation, and account for the observed differences in proliferative capacity between normal and transformed keratinocytes</td>
<td>A lack of the ability to differentiate cannot fully explain differing responses to changes in extracellular calcium levels</td>
<td>[20]</td>
</tr>
<tr>
<td></td>
<td>ABM of wounded epithelial cell monolayers accounting for cell cycle, proliferation, contact inhibition, migration, and calcium concentration</td>
<td>Predict the healing characteristics of scratch wounds under varying levels of extracellular calcium</td>
<td>Experimentally validated findings; propose hypotheses and future work to ascertain reason for differences observed in the lag-time associated with cell migration</td>
<td>[22]</td>
</tr>
<tr>
<td></td>
<td>Early generation ABM of wounded epithelial cell monolayers</td>
<td>Proof-of-concept model; investigate the effects of calcium on wound closure</td>
<td>Qualitative agreement between in silico predictions and parallel in vitro experimental outputs; suggests inadequate handling of proliferation via juxtacrine signaling</td>
<td>[23]</td>
</tr>
</tbody>
</table>

suited to capturing these interactions. Some tumor models explicitly include angiogenesis as contributing to the tumor [12], a process that has also been modeled separately [8].

Many processes in immunology are also highly dependent on cell interactions and their spatial locations. ABMs have been developed to examine granuloma formation [9], B-cell affinity maturation in lymph node germinal centers [7, 15, 16], and the initial stages of formation of an immune synapse between T-cells and antigen-presenting cells [10]. Thus, ABMs have been applied to the study of physiological and pathological events, and have produced new understanding with regard to how cells interact with one another to generate a unified tissue structure (Table 1).

The ABM approach is also ideal for studying the morphogenesis of tissues during embryonic development because these biological processes are highly dependent on both genetic and environmental events [17, 18]. ABMs are capable of capturing the spatio–temporal interplay between the two and thus interrogating the types of ‘chicken-or-the-egg’ questions that often arise in this field. Moreover, elegant fate-mapping techniques have provided detailed information about the anatomy and
structure of embryos, and this information has been used to generate well-informed initial conditions and boundary conditions for ABMs [6].

The appeal of ABMs for their ability to simplify complex physiological processes and allow interrogation of observed phenomena is growing. Their use and validation, however, has hardly been limited to the in vivo experimental setting. Modeling in vitro events generally offers the benefit of precisely controlled experimental and simulation parameters, highly confident rule-sets, and tractable experimental validation/hypothesis verification [19–23]. This benefits the modeler during ABM construction and model validation in that the process is already significantly simplified, and inputs, outputs and boundary conditions of both in vitro and in silico spaces are explicitly defined. In these ways, ABMs of in vitro phenomena are more easily paired with wet-lab experiments and are highly relevant in biomedical research (Table 1).

Integration with experiments

For a model to provide realistic or informative insights into the functioning of a complex biological system, it must, at some level, be paired with experimental work. The ways in which experiments and ABMs can be integrated with one another fall into four general categories (Figure 2). First, the rules for an ABM can be derived from published experimental literature or unpublished laboratory work. All of the models discussed so far fit into this category—many of the rules used in building the models come from experimental determinations of reaction kinetics, cell proliferation assays, determinations of the phenotypic characteristics of tumor cells and protein expression profiles. The types of data used for the basis of these rules often come from high throughput assays—gene and protein expression arrays, cell activation screens, and other in vitro assays, where the behaviors of individual cells can be easily monitored and independent parameters can be systematically adjusted. Bailey et al. [4] governed a model of monocyte trafficking in the microvasculature with rules for monocyte rolling dependent on chemokine levels, wall shear stress and adhesion molecule expression—all parameters that could be directly translated from in vitro experiments using parallel plate flow chambers or flow cytometry. These in vitro experiments allowed control of all the inputs (cell type, adhesion molecule presence, mechanical force presentation and soluble chemokine levels) and quantifiable high-resolution observations, something that is considerably more difficult to achieve in vivo. Similarly, Walker et al. [20] investigated the effects of calcium on normal and transformed keratinocytes and their subsequent proliferation and differentiation and relied on data obtained from parallel work in their own lab. Data from in vitro work helped construct a well-defined rule-set, and parallel experimental wet lab work allowed rapid and focused hypothesis testing—namely that the observed in vitro proliferative differences between normal and transformed keratinocytes is not due solely to differentiation capability [20].

Second, ABM rules can be tuned or parameterized so that the model output metrics fit a desired set of experimental data [7]. This involves holding the most certain literature-based parameters
steady, while systematically varying combinations of less well-defined parameters until a better match between model predictions and experimental results is achieved. For example, latin hypercube sampling, as implemented by Segovia-Juarez et al. [9], is a high dimensional sampling method that can ensure that the parameter space is fully explored.

Third, ABM outputs can be validated by comparison to a set of experimental results that were not used in the parameter fitting stage of model building. A model can be considered valid and potentially predictive if it reproduces results or patterns it was designed to reproduce, as well as others for which it was not specifically programmed. For example, validation can take the form of accepting model hypotheses that fit within a set of literature constraints: Meyer-Hermann tested a model of germinal center B-cell selection by rejecting as invalid all possible models that failed to reproduce minimum ratios of output centrocytes to centroblasts, and minimum increases in output plasma cells [7]. Validation can also involve more direct comparisons to simultaneous experimental work: Peirce et al. [8] validated a model of microvascular remodeling by performing matched in silico and in vivo experiments. Specifically, for experiments involving the generation of new blood vessels, rat mesenteric feeder vessels were ligated in order to change pressures or a dorsal skin window was implanted with a microbead containing vascular endothelial growth factor (VEGF), a potent angiogenic stimulus. In both cases, images were taken of the vasculature during the surgery, and those images were digitized and used as the initial conditions for the agent-based model. Experiments were run simultaneously and both in vivo and in silico outcomes were analyzed using the same metrics [8]. Similarly, for models designed to predict outcomes in medical practice, comparison to clinical trials can also be used for validation. A model of the innate immune response, where capillary endothelium could be damaged by inflammation and infection, was developed by An [5]. He validated this model by demonstrating that the outcomes of a number of failed cytokine clinical trials for systemic inflammatory response syndrome/multiple organ failure (SIRS/MOF) could have been predicted using an ABM [5].

Because ABMs at the multi-cell level produce outputs that typically take the form of tissue patterns or spatio–temporal arrangements of cells with defined features (e.g. juxtaposed sheets of cells, tubes of cells interconnecting in a network, clusters of cells with certain orientations), experimental techniques that allow visualization of tissue patterns are extremely useful for validation of ABM predictions. Confocal microscopy paired with immunohistochemistry and other methods of visualizing individual cells in tissues are often employed, but these usually allow only a single assessment (one analysis time point) per experiment. Techniques that allow cell lineage tracking and multiple time points, especially noninvasive techniques, are more useful for providing model validations because of the temporally dynamic nature of ABMs. Examples include intravital microscopy of thin tissues as a method of real-time visualization of blood vessels in a live animal [8], bone marrow transplantation for green fluorescent protein (GFP)-tagged cell tracking [24] and noninvasive IVIS tracking of luciferase-expressing cells [25].

Finally, predictions from a validated ABM can be used to drive future mechanistic wet lab work, which can, in turn, improve the realism of the model. Walker et al. [22, 23] have created several ABMs of in vitro epithelial cell monolayers to investigate the effects of contact inhibition, proliferation, cell cycle, migration and calcium concentration on the healing of scratch wounds. By intimately pairing their in silico simulations with parallel in vitro experimental work, they were able to identify differences in outcomes that reflected an incomplete understanding of the system and/or inadequate handling of the phenomena in the in silico rule space. This prompted further wet lab work to modify and restore a higher confidence in the initial rule set. Differences in healing times predicted from their model and observed in experimental work were used to suggest that juxtacrine signaling (e.g. cell–cell contact stimulates cell proliferation) may play a more significant role in wound healing than previously thought. There was evidence in the literature for this; although it was not included in early generation models, it was something the authors highlighted and proposed investigating further experimentally [23]. Walker et al. expanded their model of wounded epithelial monolayers and again identified multiple differences that highlighted the need for future wet-lab experimentation. For example, the computational model predicted that cell migration into a recently denuded scratch
wound under low calcium concentrations would occur immediately, while parallel wet lab work observed up to a 4 h time delay. This disparity led to another cycle of hypothesis generation and testing, and the authors outlined future necessary in vitro assays and offered new hypotheses [22].

**New insights from ABMs**

The results from the above studies and the motivations of the authors for undertaking them are instructive as to the kinds of insights that can be gained from the use of ABMs. One of the chief reasons cited for building computational models is that the corresponding experimental studies would be time- and cost-prohibitive [26]. Additionally, ABMs can provide insight into mechanisms responsible for complex observed phenomena that are not easily testable in the laboratory. Ausk *et al.* [27] built upon a previous model of signaling between osteocytes [28] to explore the potential causes behind a somewhat counterintuitive observation—that rest-inserted loading cycles lead to a greater increase in bone mass than continued loading cycles [29]. Using the model, they hypothesized that the increased total osteocyte network activity observed was a result of resting cycles allowing recovery of molecular stores, an observation that they intend to test in the future by expanding the model to incorporate signaling by specific molecules [27].

Other authors have employed this technique because the phenomena involved occur at timescales that are too short or with components too complex to be accurately measured by equipment available today. Casal *et al.* developed their model of T-cell recognition in an attempt to resolve ‘a continuing difficulty…[namely] that the characteristics of physiological peptide repertoires present on the cell surface remain poorly understood. This is due to the experimental difficulties inherent in manipulating and observing [peptide-Major Histocompatibility Complex] ligand mixtures of scale and complexity comparable to that on the surface of naturally occurring APCs [antigen presenting cells]’ [10]. They use the results of this model to propose three hypotheses about the effect of the sub-threshold background of peptides presented on APCs for T-cell activation, something not addressable with other techniques [10].

ABMs are ideally suited for testing and generating alternative hypotheses. Once a model is built, it becomes a simple task to ask questions such as ‘If leukocyte rolling functions according to this set of rules, will the cell flux or cell extravasation match what has been experimentally observed?’ [4]. While successful testing of a hypothesis in a model can never serve as proof of the validity of the hypothesis, the model can drive future experiments that would.

Meyer-Hermann *et al.* [7] used an ABM to compare five different hypotheses for B-Cell selection by follicular dendritic cells in lymph node germinal centers. By comparing the resultant patterns to previous experimental work, they concluded that two of the proposed hypotheses (a refractory time between centrocye and dendritic cell engagements, and competition for T-cell selection), either alone or in combination, could be responsible for experimentally observed B-cell maturation and warranted further experimental investigation [7].

Longo *et al.* [6] developed an ABM model of *Xenopus laevis* epithelial cell morphogenesis, consisting of ~300 agents to test a hypothesis regarding the relationship between fibronectin fibril assembly and differential cell adhesion during blastocoel roof thinning. After the ABM was validated by performing a series of identical in vivo and in silico manipulations and comparing the model prediction to the experimentally measured quantitative descriptors of the tissue patterning process, it was used to test the novel hypothesis. When the hypothesis was inserted as a series of rules in the main rule set, the ABM predicted a specific tissue patterning response, which was subsequently verified by an additional bench-top experiment [6]. In this way, the ABM was useful in providing quantitative evidence to support the feasibility of a novel hypothesis which might otherwise have been discarded.

For many biological problems, ABM can provide more information about the mechanisms of a process than other modeling techniques. Abbott *et al.* [12] built an ABM, CancerSim, of the interaction of cells exhibiting previously published ‘hallmarks of cancer’ [30] to investigate the mechanisms behind tumorgenesis and compared it to a previously published ordinary differential equation (ODE) model of the same cell behaviors [31]. The ABM and ODE models both identified a similar combination of phenotypes that would present the shortest path to cancer, but the ABM provided improved cell spreading resolution, was able to follow the fates of single cells, and could examine the interactions between cells within a heterogeneous population [12].
As steps in the modeling process, ABMs that fail to predict experimentally observed patterns can provide valuable information. Assuming that the rule set is correctly implemented in the model code, model failure suggests one of two things: that the cellular interactions included in the rule set are not sufficient to produce the observed patterns, or that assumptions made in determination of the rule set removed necessary complexity from those interactions. Sensitivity analysis of the model can also predict parameters or interactions on which the in vivo process is critically dependent. These interactions or molecules could be translated into new drug targets.

Challenges encountered using the ABM approach

**Determination of parameters**

One of the major challenges inherent in building an ABM is finding suitable and accurate empirical data on which to base a rule. For experimental data that is easily found in the literature, there is often conflict between sources. While basic expression profiles of many cell types can be determined from published literature or obtained through mRNA microarray experiments, protein expression levels are difficult to quantify in vivo. In Bailey et al. [4], the rule for P-selectin expression by microvascular endothelial cells (ECs) was derived from publications reporting that ~5% of microvessel ECs express this molecule [32, 33]. This observation was limited to post-capillary venules, and only qualitative observations concerning P-selectin expression across the arterial and capillary vascular beds were available. Furthermore, even when a publication reporting the desired data is available, the study may have been conducted in a different tissue or cell type, in which case those values may have to be extrapolated to the ABM environment. Values may be reported for dermal microvascular ECs or cardiac ECs but not for ECs in the tissue of interest, skeletal muscle [33]. A caveat in this case is that the expression profiles of ECs residing in skeletal muscle could be completely different than those residing in the dermal or cardiac microvasculature.

Secondly, experiments may have been published but in an obscure location and not as the main thrust of a paper. A thorough search of literature on topics related to the parameter of interest may generate useful results. The parameter may be able to be determined from the slope of a line in a figure (supplementary information of [4]), or in unpublished supporting data. In this case, it is possible that the authors of the relevant paper could be contacted to get the original data from the figure or solicited for additional data.

Another possibility is that parameters could be determined experimentally but the experiments may not have been performed or published. In this case, model builders have a few options—if they have the experimental capability, they could measure the parameters themselves, they could work with a collaborator to measure the parameters, or that parameter could be targeted for later parameterization and a sensitivity analysis. In Longo et al. [6], the migration rate of blastocoeol roof epithelial cells could not be found in the literature. Because collaborators were performing similar research, the parameter was measured as part of a subsequent experiment and incorporated into the ABM.

Finally, the desired parameter values could be impossible to determine with current experimental techniques. An example of this is the determination of gradients of VEGF concentration in vivo around an individual cell [34, 35]. Until experimental techniques become available to measure such parameters, the only option is to use a best estimate of the parameter and conduct a detailed analysis of the sensitivity of the model and robustness of results to variations in the estimated parameter. Thus, when constructing an ABM, one must interpret and devise ways to accommodate and incorporate the best available information.

**Level of abstraction of ABM**

Due to the complexity of the systems that are studied, it is often necessary to make a number of simplifying assumptions. In the construction of a model, one starts off with the simplest possible model, a ‘null model’, and gradually adds complexity [36]. Consistent with this principle, in building a model of granuloma formation, Segovia-Juarez et al. [9] initially simplified all chemokines involved in the immune response to one composite value, termed ‘chemokine’. This composite chemokine diffused through the simulation environment, and leukocytes were programmed to migrate up the chemokine gradient (toward higher levels of chemokine) [9]. While the implementation ignored the differential effects of many chemokines, it was still able to produce useful predictions that led to important conclusions about the system—namely that changes
in diffusion or degradation of chemokines could have strong influence on the size of granulomas [9]. As the ABMs are iteratively refined, these types of initial assumptions are revised, sometimes leading to loss of desired behaviors [37], at which point other potential hypotheses are examined.

**Computational power required**

The computational efficiency of ABMs declines as the number of agents (cells) increases and as the length of the rule set increases, making large and complex ABMs time consuming to compute. Because stochastic rules are often incorporated, simulations must be run repeatedly to obtain statistical significance, thereby increasing the run-time needed to complete an analysis. Most ABM software packages are not able to run across parallel processors, and thus models using pre-built modeling environments tend to be restricted to a single processor machine. In some cases, this can be addressed. For example, Casal et al. [10] initially modeled the entire surface of interaction between a T-cell and an APC. With the spatial resolution that they employed, this worked out to a simulation environment of a 700 by 700 grid, or 490 000 nodes [10]. Applying model rules to this entire area for the 60 000 time-step simulation limit for all the planned simulations would have been prohibitively expensive in terms of computational time. They were able to render their simulation tractable by reducing the size of the grid to 300 by 300, and showed that results were comparable to the full-scale simulation [10].

**Comparison of ABMs with differential equation models**

One of the major reasons for using ABMs is that they allow for tracking of individual cells and cell properties. This is something that cannot be done in most *in vivo* experimental settings and is tedious or impossible to do with partial differential equation (PDE) models. In larger ABMs, however, this technique can produce enormous amounts of data to analyze that may or may not be meaningful. Simulating the behavior of 1000 cells for 1500 time steps with 10 measured properties per cell would produce 15 million data points per model run. Additional data mining techniques must be employed to make productive use of this type of data. Another benefit of ABM is that some ABM modeling packages require less coding experience than PDE models, and code can be much more intuitive for a nonmodeler to understand. Even with the simplest types of descriptive code, however, all agents and subroutines must be thoroughly tested to ensure that emergent properties do not result from coding artifacts. The heterogeneous properties of tissue are also more easily represented in an ABM. One of the drawbacks to ABMs, as discussed earlier, is that at commonly used levels of abstraction, necessary parameters are not concrete—stochasticity added by a cell’s initial internal state can make determination of parameters more tedious. However, this stochasticity allows for emergent properties during ABM simulations. At common levels of abstraction for differential equations, parameters often involve reaction rates and binding coefficients, properties which can be definitively measured. In both ABMs and PDE models, a large number of simulations are required for adequate exploration of the parameter space. In PDE models, there are well-developed mathematical methods for this exploration, while these techniques are just beginning to be applied to ABMs. The speed of execution of ABMs may be more dependent on the skill of the programmer than for PDE models. That said, larger ABMs can monopolize computational resources, and few ABM suites allow for parallel processing to reduce computational time. Finally, ABMs require the assumption that all properties can be modeled discretely, while PDE models require a continuum approximation. The reality for biological systems may lie somewhere in the middle of these approaches—some cell variables may be continuous, while others may have discrete states. As shown here, each approach has its benefits and drawbacks, and should be selected carefully depending on the type of question to be addressed. New approaches are being developed to bridge ABM and PDE modeling techniques [38] in order to capitalize on beneficial attributes of these approaches and to compensate for their respective drawbacks.

**Communication of models**

A final issue often encountered in research that employs ABM is the logistical difficulty of communicating ABMs—which often have lengthy and detailed rule sets—within the page limits of journals, and in a way that is readable yet instructive enough for readers to reconstruct the ABM and fully understand the details of how the rules were obtained. A thorough description of an ABM involves a complete annotated list of all the possible
states of each agent type, descriptions of all parameters, rules, simplifying assumptions, time and spatial scales, handling of computational order of execution issues and any other pertinent issues [13].

In many cases, ABMs are presented along with experimental validation, meaning that two sets of materials and methods must be presented within the page limits of a publication, creating a substantial challenge for the author and editor. Indeed, with the addition of methods and results for testing and application of the model, the amount of space required for publishing an ABM quickly becomes overwhelming and intractable in a single article. Grimm and Railsback propose resolving this problem by publishing one paper solely about the design and construction of the model in a more theoretical journal, which can then be referenced with relevant changes as the model develops [13].

Two series of articles (Meyer-Hermann et al. [7, 15, 16, 39] and Athale et al. [40], Athale and Deisboeck [41], Mansury et al. [42–44], Zhang et al. [45]) demonstrate this technique and the evolution of their respective models well, from proof of concept models with a large number of simplifying assumptions [15, 44] to more complex models [16, 40, 43], application to developing new hypotheses [7, 39, 41, 42] and even the addition of multi-scale components [40, 41, 45].

The future of ABMs in biomedical research

As biomedical research advances and evolves to address increasingly complex multi-factorial problems, it approaches a point where intuition alone may not serve to develop plausible hypotheses. The need for computational simulation and prediction of multi-cellular events in space and time (as in ABM) then becomes apparent [46]. Outputs of ABMs can be visual and easily accessible to biologists (which facilitates interdisciplinary collaboration) and models can be built more quickly and at lower cost than laboratory experiments, freeing resources for a more informed exploration of the hypothesis-space.

Integration of cellular level ABMs with targeted subcellular gene and protein interaction network models—multi-scale modeling—as begun by Mansury et al., will greatly enhance the ability of this technique to accurately predict complex emergent phenomena. Carefully considered selection of subcellular networks for incorporation into ABMs can increase the resolution of an ABM without excessively increasing the computational resources required [38, 42].

In order for ABMs to gain widespread usage in biomedical research, the problem of how to effectively convey their construction, implementation, validation, and simulation in peer-reviewed journals needs to be addressed. To this end, if journal space constraints require publication of an abbreviated description of the model or subset of rules, a full description and annotated rule set should be published online, either as an official supplement to the article or made available on the authors’ website [13]. Because it can be difficult to replicate a model and its results from a published textual description, the source code or a compiled version of the model should also be made available so that other scientists can verify the model and perform additional independent validation [13].

As ABM continues to emerge and be recognized as a powerful tool, it is paramount that effective methods for development and testing be standardized. Grimm and Railsback address this in the ecological field [13]. Although the methods they describe are applicable to ABMs in general, consensus within the biomedical ABM community should be sought; this may require dialogue and collaborations with ecologists and epidemiologists who have used this technique previously. This lack of standardization remains a hurdle for the acceptance of ABM in the biomedical field as a whole. As such, we feel that the parallel experimental and modeling approach we describe in this review joins ABMs with more traditional science, and is essential to its acceptance. The coupling of ABM and experiment can help speed that acceptance by aiding in developing and testing hypotheses that explain nonintuitive experimental results. This approach is useful in studying a variety of tissue patterning events, including embryogenesis, physiological and pathological tissue remodeling, and the effects of therapeutic interventions. Thus, its applicability and utility span basic (e.g. identification of molecular mechanisms) and applied science (e.g. targeted drug discovery). ABM can be used to answer both in vitro and in vivo questions, and the application of this approach to new and different settings will only help to expedite a formalization for ABM construction, validation, and implementation in biomedical research.
Key Points

- Agent-based modeling is a computational technique that predicts complex patterns by simulating the interactions of a set of individual agents with each other and their environment according to a literature-derived rule set.
- Agent-based models provide different types of information about a process than can be derived from differential equation-based models. In ABM, each cell is represented explicitly, complex interactions between cells are easily probed, and a spatially and temporally heterogeneous tissue environment can be more thoroughly explored.
- The dependence of agent-based models on empirical rules makes validation of models with laboratory experimentation a necessity. There are many ways to pair this modeling approach with experimental analysis, and the synergy between the two expedites biological discovery (Figure 2).
- Agent-based models are especially useful for generation and initial profiling of hypotheses about how tissues undergo pattern changes by structurally remodeling, adapting, or growing.

References