

## Review

## Solving Immunology?

Yoram Vodovotz,<sup>1</sup> Ashley Xia,<sup>2,17</sup> Elizabeth L. Read,<sup>3</sup>  
 Josep Bassaganya-Riera,<sup>4</sup> David A. Hafler,<sup>5</sup> Eduardo Sontag,<sup>6</sup>  
 Jin Wang,<sup>7,8</sup> John S. Tsang,<sup>9</sup> Judy D. Day,<sup>10</sup>  
 Steven H. Kleinstein,<sup>11,12</sup> Atul J. Butte,<sup>13</sup> Matthew C. Altman,<sup>14</sup>  
 Ross Hammond,<sup>15</sup> and Stuart C. Sealfon<sup>16,\*</sup>

**Emergent responses of the immune system result from the integration of molecular and cellular networks over time and across multiple organs. High-content and high-throughput analysis technologies, concomitantly with data-driven and mechanistic modeling, hold promise for the systematic interrogation of these complex pathways. However, connecting genetic variation and molecular mechanisms to individual phenotypes and health outcomes has proven elusive. Gaps remain in data, and disagreements persist about the value of mechanistic modeling for immunology. Here, we present the perspectives that emerged from the National Institute of Allergy and Infectious Disease (NIAID) workshop ‘Complex Systems Science, Modeling and Immunity’ and subsequent discussions regarding the potential synergy of high-throughput data acquisition, data-driven modeling, and mechanistic modeling to define new mechanisms of immunological disease and to accelerate the translation of these insights into therapies.**

### The Complexities of Immune Dysregulation

Dysregulation of immune and inflammatory pathways is at the core of many diseases facing citizens of the 21st century. These diseases are multifactorial, exhibit great patient-to-patient variability, and are often intractable to both traditional therapy and reductionist insight. Immune regulation in healthy individuals is orchestrated by highly complex, massively and dynamically-interacting gene networks, immune signaling pathways, cellular networks, host–pathogen interactions, and nutrition–microbiota–host interplay. Immune-mediated diseases, such as those that stem from the response to viral or bacterial infection, as well as trauma-induced inflammation, Crohn's disease (CD), type 1 diabetes mellitus (T1DM), or multiple sclerosis (MS), result when these homeostatic immunoregulatory mechanisms deviate or fail.

Recent decades have brought tremendous progress in the application of ‘omics’ and associated bioinformatics methods to several immunological diseases. Genome and disease-mapping consortia have helped elucidate how genetic variation influences disease risk, and consortia over the past 15 years have provided progressively finer resolution of immune disease-associated genetic variation [1,2]. Multidimensional and high-throughput technologies are facilitating a comprehensive and systematic interrogation of complex immune pathways and the identification of relations among their components at both the cellular and molecular levels, in both healthy individuals and those with infections or immune-mediated diseases. Moreover, computational analyses in genome-wide association studies (GWAS) are identifying increasing numbers of candidate immune disease-modulating genetic elements. For example, over 60% of the genetic variation predicting MS onset has been elucidated; these genetic variants largely overlap with other autoimmune diseases [3].

### Trends

Unbiased high-throughput data-driven analysis approaches are identifying molecular processes associated with immunological function and immune-related diseases, as well as candidate immune disease-modulating gene elements.

Hypothesis-based mechanistic modeling tests whether our understanding of how a system works provides a possible explanation for the existing immunological data, and helps focus subsequent research.

A variety of immunology-friendly computational modeling tools have been developed.

Mechanistic modeling and simulations are serving increasingly important roles in drug development and clinical trials.

<sup>1</sup>Departments of Surgery, Immunology, Computational and Systems Biology, Clinical and Translational Science, and Communication Science and Disorders, and McGowan Institute for Regenerative Medicine, University of Pittsburgh, Pittsburgh, PA 15213, USA

<sup>2</sup>Division of Allergy, Immunology and Transplantation, National Institute of Allergy and Infectious Diseases National Institutes of Health, 5601 Fishers Lane, Room 7A78, MSC 9828, Rockville, MD 20852, USA

<sup>3</sup>Department of Chemical Engineering & Materials Science, Department of Molecular Biology & Biochemistry, University of California, Irvine, CA 92697, USA

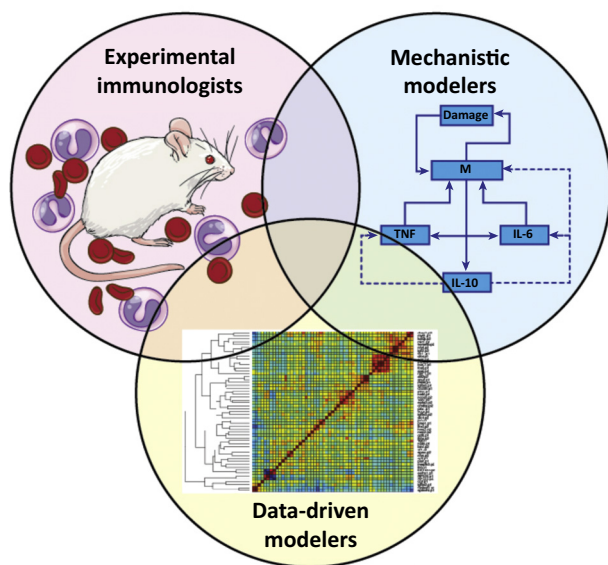
<sup>4</sup>Biocomplexity Institute of Virginia Tech, 1015 Life Sciences Circle, Blacksburg, VA 24060, USA

<sup>5</sup>Departments of Neurology and Immunobiology, Yale School of

A fundamental challenge for immunology is how to extend this rapidly growing, individualized data acquisition. More specifically: what is the optimal way to leverage focused, hypothesis-driven immune research to understand how immune system behavior in health and disease emerges from molecular, genetic, epigenetic, cellular, and environmental modulatory elements? The NIAID convened a workshop of computational and experimental immunologists in October 2015 to address this challenge. At the workshop and in subsequent discussions, different viewpoints separated experimental immunologists, data-driven modelers, and mechanistic modelers (Figure 1). Although researchers from divergent backgrounds approach fundamental questions in immunology from different and sometimes clashing perspectives, they share a common interest in reducing the burden of immune disease by improving the predictability of immunological responses.

A further common view shared by experimental and computational immunology researchers is that, despite the explosion of ‘omics’ and related techniques, such as mass cytometry (CyTOF), in experimental immunology, a dearth of key data is a major impediment to an improved mechanistic and predictive understanding of immunology. The ability to infer novel interactions, principal drivers, and interconnected networks using data-driven modeling, including approaches such as bioinformatics, machine-learning classifiers, or mathematical modeling of signaling networks or cell dynamics, is now well accepted [4]. However, mechanistic mathematical modeling has been incorporated less well into experimental immunology research. Some researchers are skeptical that mechanistic models can provide predictive power, especially in the absence of key quantitative data to anchor the model fully. However, modelers point out that even relatively simple mathematical models can improve insight from existing data, and that better integration of mechanistic modeling approaches into immunology research can optimize the choice of what and when to measure, sharpen hypothesis testing, and help make the resulting data sets more informative (Box 1).

The initial process of formally developing a simple and perhaps initially incorrect model that includes only a few entities often proves invaluable in revealing hidden assumptions, formalizing hypotheses, and identifying key missing measurements and experiments. Importantly, data mining and mechanistic modeling are synergistic approaches that are not intended to replace



**Figure 1. Different Perspectives of Immunology Researchers.** Generally, immunologists seek to understand the role of immune system in health and disease, focusing on how genetic, epigenetic, and environmental variations affect immune function and influence individual disease expression and response to treatment. Data-driven modelers seek to map predictive relations between multiple data variables and disease in individuals, reconstructing and examining how molecular and cellular networks change in health and disease. Mechanistic modelers seek to define how a defined set of molecular and cellular interactions can lead to complex outcomes in health and immunological disease. This latter group utilizes abstractions of biological mechanisms encoded into computational models, and seeks to explain how diverse, complex health and disease phenotypes emerge from a defined set of immune interactions, set in motion by specific initial biological conditions.

Medicine, New Haven, CT 06520, USA

<sup>6</sup>Department of Mathematics, Institute for Quantitative Biology, and Cancer Institute of NJ, Rutgers University, New Brunswick, NJ 08903, USA

<sup>7</sup>State Key Laboratory of Electroanalytical Chemistry, Changchun Institute of Applied Chemistry of Chinese Academy of Sciences, Changchun 130022, China

<sup>8</sup>Department of Chemistry and Physics, Stony Brook University, Stony Brook, NY 11794-3400, USA

<sup>9</sup>Systems Genomics and Bioinformatics Unit, Laboratory of Systems Biology, National Institute of Allergy and Infectious Diseases, Trans-NIH Center for Human Immunology, National Institutes of Health, Bethesda, MD 20892, USA

<sup>10</sup>Departments of Mathematics and Electrical Engineering & Computer Science, National Institute for Mathematical and Biological Synthesis, University of Tennessee, Knoxville, TN 37996, USA

<sup>11</sup>Department of Pathology and Department of Immunobiology, Yale School of Medicine, New Haven, CT 06511, USA

<sup>12</sup>Interdepartmental Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT 06511, USA

<sup>13</sup>Institute for Computational Health Sciences, University of California, San Francisco, Mission Hall 550 16th Street, 4th Floor, Box 0110, San Francisco, CA 94158-2549, USA

<sup>14</sup>Department of Allergy and Infectious Diseases, University of Washington, Seattle, WA 98109, USA

<sup>15</sup>Center on Social Dynamics & Policy, Economic Studies Program, The Brookings Institution, Washington DC 20036, USA

<sup>16</sup>Center for Advanced Research on Diagnostic Assays, Immunology Institute, Friedman Brain Institute, and Departments of Neurology, Neuroscience, Pharmacology, and Systems Therapeutics, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

<sup>17</sup>Current address: Molecular Transducers of Physical Activities (MoTrPAC) Division of Diabetes, Endocrinology, & Metabolic Diseases (DEM), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institutes of Health (NIH), 6707 Democracy Blvd, RM 6074, Bethesda, MD 20817, USA

\*Correspondence: [stuart.sealfon@mssm.edu](mailto:stuart.sealfon@mssm.edu) (S.C. Sealfon).

**Box 1. Small versus Big Models**

Small or reduced models help select among alternative mechanistic hypotheses and provide qualitative predictions. By contrast, highly detailed models can lead to specific, quantitative predictions of behavior over time. Model complexity can be a double-edged sword. An increase in the number of model parameters increases both model resolution and the precise representation of detailed mechanisms while also increasing the risk of overfitting the experimental data.

**Small Models—Qualitative Insights**

As an example, the host innate immune response to bacterial sepsis was simulated using a reduced model [79]. In this study, a mechanistic model comprising three ordinary differential equations accounted for a bacterial pathogen population, an early proinflammatory mediator derived from the initial activation of macrophages and/or neutrophils, and the later production of damage-associated molecular pattern molecules (also a surrogate for the damage sustained by the host in the course of infection and inflammation). Each equation represented the rate of change of each of these components (pathogen, composite early immune response, composite damage response) relative to the levels of the other components and simulation of the model provided the predicted level of each component over time. This highly abstracted model reproduced both a healthy outcome and diverse negative outcomes that are observed in patients with sepsis, depending on initial conditions and parameter values. Simulation supported the hypothesis that the clinical condition of sepsis can arise from several distinct physiological states, each of which requires a different treatment approach.

**Large Models—Quantitative Insights**

Large, detailed models can integrate data, procedural knowledge, and theory into information-processing representations of innate and adaptive immune responses to simulate complex immunological systems. An example is an inflammatory bowel disease (IBD) model [27]. The modeling platform used incorporated mathematical modeling and machine-learning methods to allow clinical trial data integration, including subjects' CD activity index (CDAI), endoscopic measures of mucosal healing, data from electronic medical records, preclinical data from animal models, and cellular and molecular signatures. Simulating such models can guide clinical plans, predict results of clinical trials, and suggest new therapeutics. Through this modeling and experimental validation, a novel Lanthionine Synthetase C-like 2 (LANCL2)-based oral therapeutic for CD was identified and is being developed [80]. Model simulations will be used to help guide the design of Phase 2 and 3 clinical studies [71].

experimentation, but rather to each serve a complementary role in an iterative cycle of data exploration, hypothesis exploration, and experiment. In this regard, one motivator for the NIAID conference was the concern that academic immunology research is lagging industry and other medical research fields in incorporating modeling approaches. In adapting to the high rate of failure and the more than US\$2.6 billion total research and development (R&D) costs involved in the approval of a new drug or biologic [5], industry has turned increasingly to computational modeling at all levels, from modeling drug–receptor interactions to pharmacokinetic and pharmacodynamic modeling, to *in silico* clinical trials [6,7]. The Critical Path initiative of the US Food and Drug Administration (FDA) has further stimulated the incorporation of mathematical modeling in drug discovery [8]. Recent NIH requests for applications in many areas, such as cancer research, specifically call for the integration of modeling approaches. These trends are applicable to immunology.

We suggest that immunology research needs to evolve more rapidly from a paradigm of experiment, data analysis, and interpretation, to a more integrative approach that incorporates formal computational modeling and prediction. The fruitful application of this type of complementary, iterative approach to the question of how the T cell response is scaled to the level of the antigenic signal has recently been reviewed [9]. Here, we explain the complementary role and limitations of both data-driven and mechanistic modeling, and review successes of computational immunology in both basic and translational settings. Finally, we lay out a roadmap for bridging the immunologist–modeler divide to accelerate insight into emergent mechanisms of immune system function and the translation of data into individualized treatment for infectious and immune-mediated diseases.

**Modeling Approaches and Applications in Immunology**

One likely reason for the narrow adoption of mathematical modeling by the immunology community, similar to the broader biomedical research community, is the sheer diversity of

methods and approaches broadly encompassed within this discipline, as well as the lack of formal computational, mathematical, and statistical training of most immunologists. Below, we review briefly major types of computational modeling and their application to immunology.

Various data-driven modeling or data-mining approaches are an essential element of large-scale systems immunology research. These methods share a grounding in rich experimental data sets, providing quantitative measurements of molecular concentrations, biochemical reactions, or cellular and tissue-level parameters. These data are then combined with topologies and injected into computational approaches that seek to determine the multivariate relations among objects in the networks, and how the system as a whole responds to perturbation [10]. This type of modeling iterates between prediction, experimental testing via perturbation, and refinement of the model based on the new data [4].

Unsupervised analysis approaches, such as hierarchical clustering, principal component analysis, and more classical statistical analyses, are considered data-driven modeling, as are various network inference algorithms that can provide information about the correlation and, in some instances, causal structure among variables [11,12]. These approaches have been used in modeling and simulation of the mono- (e.g., folding) or bimolecular interactions [e.g., MHC-T cell receptor (TCR)-Ag complexes], of signal transduction cascades, of CyTOF data representing T cell phenotypes, of genetic regulatory networks, or of cellular behavior (e.g., [13]). Some studies use regression techniques that build models predictive within the conditions of the data upon which they were trained [14], as well as predictive classifiers [15]. These methods can provide insight into the molecular circuits underlying immunological phenomena, such as switching between immune cell subsets [16]. As an adjunct to this type of endeavor, resources such as a user-friendly, web-accessible, 38 000 public experiment-based resource to predict immune-system and disease relevant gene–gene relations supports big data mining and hypothesis development by immunology researchers without computational training [17]. A potential limitation of data-driven modeling is that, at least in some cases, predictions are based on correlations, making generalizable mechanisms difficult to infer from the resulting model.

Mechanistic modeling (sometimes called ‘hypothesis-based modeling’ [18]) is based on specific mechanistic hypotheses, and uses mathematical equations and computer simulations to describe abstractions of complex biological systems that are, in essence, hypotheses about the core functions or processes that govern the system [4,19–21]. Data-driven and mechanistic modeling have commonalities and can complement each other. However, whereas data-driven modeling is based on obtaining data and inferring relations among entities and/or differences among groups, mechanistic modeling typically starts with an initial picture of how things work that, after abstraction into mathematics, can then be played forward in time and compared against experimental or clinical results. Mechanistic modeling has provided non-intuitive mechanistic insight into many areas of immunology, ranging from molecular systems, such as how analog signals from TCR engagement become digital further downstream and the basis for nonlinearities in the system [22–25], to clinical responses in disease, as described below. A major rationale for pursuing modeling in immunology (and systems biology more generally) is that complex networks with cooperativity and feedback (which broadly describes biological systems at every scale) exhibit complex, nonlinear relations (a characteristic that can be inferred from, but not tested explicitly with, data-driven modeling [13,26]). Simply put, it is difficult to predict how a given perturbation at the molecular level will affect system behavior at higher scales over time or at homeostasis without the aid of a mechanistic model that can be ‘played forward in time’ under a variety of conditions that would be impractical to test only using experimentation.

Perhaps the most fundamental difficulty when pursuing mechanistic modeling concerns the appropriate level of abstraction and/or complexity for the model (e.g., do we study at the

### Box 2. Immunology-Friendly Computational Modeling Software

There is a variety of user-friendly platforms that facilitate immunological modeling without detailed understanding of the mathematical considerations ‘under the hood’. These include:

- Complex Pathway Simulator (COPASI): enables researchers to construct biochemical and immunological ODE models that help them understand how a system works. Model parameters can be adjusted as necessary to reproduce experimental results<sup>iii</sup>.
- ENteric Immunity Simulator (ENISI) MultiScale Modeling (MSM): a platform designed for modeling complex mucosal immune responses. ENISI MSM can simulate *in silico* experiments from signaling pathways to tissue level events, and *in silico* clinical trials. It integrates agent-based modeling (ABM), ordinary differential equations (ODE), and machine-learning methods<sup>iii</sup>.
- Simmune: provides intuitive visual interface-guided modeling of molecular interactions to provide cellular, spatial, and time-dependent simulations of immune processes<sup>v</sup>.
- Monte Carlo Cell (MCell): a program that uses spatially realistic 3D cellular models and specialized Monte Carlo algorithms to simulate the movements and reactions of molecules within and between cells, supported by extensive training tutorials<sup>v</sup>.
- RuleBender: this tool helps the collaborative development of diagram-based models, usable by immunologists who are not necessarily experts in mathematical modeling, while providing the deeper insights into the simulation process that are needed for tasks such as model and simulation debugging, and overall understanding of molecular systems<sup>vi</sup>.
- Virtual Cell (VCell): is a comprehensive platform for modeling cell biological systems that is built on a central database and disseminated as a web application<sup>vii</sup>.

molecular or cellular level, at the second-, minute-, or days-long timescale, how few or many components and relations should be included?) relative to the underlying biological system and the scientific questions being posed (Box 1). Defining the optimum balance of modeling simplicity versus complexity for the problem at hand benefits from interdisciplinary interactions between modeling experts and experimentalists. The development of user-friendly software tools (Box 2) facilitates both formal modeling by experimentalists and this interaction between experimental immunologists and modelers. Some platforms, such as ENISI MSM, integrate mechanistic modeling at the intracellular scale and data-driven modeling at tissue level and population levels [27] (Box 2).

In the following sections, we briefly highlight some of the major common approaches to mechanistic modeling, and the diverse areas of immunology to which they have been applied.

### Systems of Ordinary Differential Equations

Systems of ordinary differential equations (ODEs) remain one of the most flexible, widespread approaches for mechanistic modeling in immunology. Each distinct ‘species’ of interest in an ODE model, whether it be an organism, cell type, virus, bacterium, or biomolecule, is represented by a continuous time variable. These variables often interact with each other in a recursive manner that precludes ‘solving’ the system of equations in the manner familiar from traditional mathematics courses; rather, these systems of equations are simulated using a numerical integration algorithm, yielding predicted trajectories for each variable (i.e., a graph of their predicted levels over time). ODE models almost always describe system dynamics (i.e., behavior over time). Individual terms in the equations express how rates of change of each species depend on population sizes, interactions among species, and external factors. Even when time-course data are not available, ODE models can be useful for analyzing the steady-state behavior of a system.

Classical applications of simple ODE models simulated interacting virus and lymphocyte populations in HIV infection [28,29]. These relatively simple models involving few species revealed an unexpectedly dynamic interplay between the virus and immune system during latent infection. Quantitative estimates of staggeringly large daily rates of HIV-1 production helped identify the need for combinatorial antiretroviral therapy regimens and spurred their development and implementation. These studies also paved the way for detailed mathematical studies of lymphocyte dynamics in infection and homeostasis, in conjunction with experiments. A few recent

examples of immunological insights contributed by this type of ODE modeling are the differences in maintenance of peripheral naïve T cells between mice and humans [30], the phenotype-specific dynamics of antigen-stimulated CD8<sup>+</sup> T cells in SIV [31], and the balance between immunopathology and protection in T cell-based vaccines [32].

At the subcellular scale, ODE models have been successfully applied to elucidate dynamic control mechanisms governing signaling and gene regulation in the immune system. An example of iterative experiments and model development leading to new insights is found in the NF- $\kappa$ B signaling system [33]. In another example, a detailed model involving approximately 50 biomolecular species and hundreds of reactions was used to study control of dendritic cell (DC) maturation [34]. The model enabled zeroing-in on cell-specific regulation of the RelB protein in mouse embryonic fibroblasts (MEF) and DCs. Differences in two biochemical rate constants were identified as being able to account for qualitatively different, cell type-specific responses *in silico*. These predictions were confirmed by altering these rates through genetic engineering of MEFs, producing DC-like control. Emerging advanced applications of ODE modeling include landscape and flux theory [35], which may hold the potential to define and quantify homeostatic versus pathological immune system states, as well as the barriers to transition between states.

#### Agent-Based Models

Whereas an ODE model typically predicts average population sizes or concentrations over time, an agent-based model (ABM) explicitly represents each individual model entity (often cells with distinct attributes) in a computer simulation. The ‘rules of engagement’ are defined in advance, and events occur probabilistically. ABMs are typically spatial (meaning, for example, that cells must be in proximity to interact), and are useful for exploring how local interactions influence larger-scale behavior [20]. In immunology, ABMs were introduced to study lymphocyte antigen-specific recognition and activation [36]. Representations of receptor–ligand interactions that represent unoccupied and/or occupied sites as zeros-ones bit strings has enabled ABMs to go some way towards modeling the combinatorial diversity of immune repertoires. ABMs have been applied to the study of: granuloma formation in *Mycobacterium tuberculosis* [37,38]; germinal center formation [39]; thymic selection [40]; multiscale modeling of mucosal immune responses to irritable bowel disease (IBD) and gastroenteric infection using ENISI, which recapitulate the relation between mucosal immune responses and inflammatory lesions in the gastrointestinal (GI) tract [41,42]; and spatial dynamics of infection [43]. Folcik *et al.* also generated a multiscale ABM that recapitulates the interactions between the cells of the innate and adaptive immune system, ultimately resulting in immune memory for an antigen following infection [44].

#### Logical Models

Logical models allow researchers to model network dynamics in the absence of detailed information about parameters governing interactions (the general paucity of this information is a perennial challenge for many types of mathematical modeling). In a logical model, network species are categorized according to a small number of discrete states (e.g., ‘active’ or ‘inactive’) and interactions such as ‘X activates Y’ govern updates to the system state at each time step. As such, only qualitative information about interactions, rather than quantitative rate parameters, is required in the building of the model [45]. Examples of applications of logical models include TCR signaling [46,47] and T cell specification [48,49].

#### Models Can Either Be Deterministic or Stochastic

A deterministic model, such as an ordinary differential equation described above, will always give the same result each time it is simulated with the same parameters. However, when studying responses at the single cell level, the small number of molecules involved may lead to large differences in behavior that are truly random. In stochastic models, the behavior of a system is affected by random uncertainty. ‘Stochastic modeling’ is an umbrella term: several different

mathematical frameworks and simulation methods are available for stochastic modeling, including stochastic differential equations, diffusion-type equations, stochastic biochemical network models, reaction-diffusion models, and probabilistic Boolean networks, among others. Agent-based models are also typically stochastic. By accounting for individual molecules and biochemical reactions, stochastic models have revealed how lymphocytes respond sensitively to small amounts of antigen. For example, a synergistic modeling and experimental study showed that feedback regulation in a particular TCR signalling module gives rise to a digital ('on' or 'off') activation response at the single-cell level [25]. Other examples of stochastic modeling applied to immunology include: studies of stochastic cytokine expression [46,50]; dynamics of lymphocyte differentiation and expansion [51]; cellular response to infection [51–54]; lymphocyte repertoire development [55]; and affinity maturation [56,57].

All models, at some level, aim to describe and predict complex systems. Mechanistic models, particularly ODE-based deterministic models, lend themselves to further practical applications, such as the dynamic control of a given complex system [58]. This application is at the heart of a field known as Control Theory, which has been applied in the context of immunology for nearly four decades [59,60]. The immune system can be viewed as comprising interlocked and finely tuned homeostatic and regulatory mechanisms, which must achieve an appropriate, proportional, and dynamic defense against infection and autoimmunity. Mathematical control theory is central to the design of engineered systems, and core concepts, such as feedback control, are clearly at work in the immune system [61]. Increased application of control theory to immunology will shed light on how natural immune responses work and how therapies can be designed to tailor immune responses in specific disease contexts. A complementary role of control theory is in the rigorous formulation and solution of optimal immunotherapy drug infusion regimens [62], as well as in identifying patient-specific therapeutic strategies for the acute inflammatory response to severe infection [63].

### Clinical Trial Simulations in Immune-Mediated and Inflammatory Diseases

*In silico* clinical trials, in which mechanistic models are carried out over a broad range of parameter values and initial conditions, using advanced machine-learning methods or stochastic simulation methods, such as ABMs, have been applied for over a decade in the setting of acute inflammation and immunity [64–69,63,70]. We highlight two examples of *in silico* clinical trials in the settings of CD and trauma-induced inflammation.

ABM and ODE models, in combination with supervised machine-learning methods, have been used to create virtual patients with CD [27,42]. In a recent study [71], 10 000 virtual patients with CD were created using information from clinical trials where both individual response to treatment and aggregated data were combined with expert knowledge with the goal of comparing current and investigational treatments. Virtual treatment strategies were designed based on changes in immunological parameters that drive response to treatment. Such virtual patients provide a diversified population to generate novel non-intuitive hypotheses regarding treatment response to novel therapeutics (Box 1). Another example of virtual patients based on mechanistic modeling involves inflammation and critical illness caused by traumatic injury. A recent study described a multicompartiment ODE model that represents a virtual trauma patient [68]. This mathematical model was used to create both individual-specific variants as well as a cohort of 10 000 virtual trauma patients [68]. The model was able to predict trauma-induced mortality as well as a non-intuitive role for interleukin-6, which was validated with single nucleotide polymorphisms (SNP) data [68].

These examples indicate that the outcome and process of clinical trials in complex immune and inflammatory diseases are becoming computationally tractable. *In silico* clinical trials can also have pitfalls [72]. They carry the risk of deprioritizing therapies that might have been valuable, or

of being unsuccessful because the pathophysiological knowledge that underlies the model remains incomplete (and, hence, the model is overfitting to the sparse data that are typically available). For instance, in the case of T1DM, the identity of effector or inhibitory molecules and cells, the impact of beta-cell regeneration, and the role of viruses or other environmental triggers that conspire to provoke autoimmune diabetes are poorly understood, let alone their quantitative or dynamic parameters (and there are likely to be several patient subtypes and instigating causes). While a plausible quantitative model of T1DM [73] could readily be programmed to simulate immunomodulatory trials of T1DM in thousands of patients, its results would only be as good as the assumptions underlying the model.

### **Integrating Modeling and Experimentation for 21st-Century Immunology**

The successes of immunology-focused modeling suggest that computational, mechanistic modeling should be at the heart of interdisciplinary research to optimize the translational impact of the immunology research portfolio, from highly focused, hypothesis-driven research, to large-scale team science, to bedside applications. Our core contention is that rather than a relatively linear process of hypothesis→experiment→data, a better approach to the complexity of immunology might be data→model→experiment→validation→refined model, leading to both novel hypotheses, data and/or theory integration, streamlined experimentation, and prediction of translation to humans and therapeutic impact of fundamental biological mechanisms. Both data-driven and mechanistic modeling would serve the immunology community in this revised workflow. Data-driven modeling would allow for inference of novel interrelations, feedback architectures, and principal characteristics; this type of modeling is already a mainstay of large, collaborative consortia focused on ‘omics’ studies. Mechanistic modeling could be a major ‘equalizer’ for many labs in an environment of limited grant funding because it would help target key experiments in a resource-constrained environment. Alternatively, mechanistic modeling could help delineate key data sets and help with design and interpretation in large, systems-oriented studies.

A major role for modeling that needs to be better utilized by the immunology research community is as a customized data analysis tool for the extraction of new, and perhaps non-intuitive, mechanistic hypotheses from experimental data. Fitting an experimental data set that provides dynamics (i.e., measurements obtained over time) using a simple model often highlights features in the data that fit poorly to the predictions of the model. This leads to thoughtful exploration of the cause of the discrepancy between data and simulation: are there experimental errors? measurement noise? a mechanism not included in the model? This process can lead to model revision to test whether a novel mechanism is a plausible explanation for the discrepancy, and to further experiments to test the hypothesis. Thus, modeling allows us to better take advantage of otherwise discarded information in the experimental data that reflect important novel mechanisms, which in turn can have a major influence on subsequent experimental work or therapeutic development. This type of modeling, illustrated by the classical HIV work by Perelson and colleagues described above [28,29], is most useful when it is anchored not only in the data, but also in the biology and immunology that underlies the assumptions of the model. Mechanistic modeling is in many ways an art, and it is important to start the process by crafting equations that are based on plausible immunological mechanisms. Indeed, one may think of mechanistic modeling as ‘the art of the possible’: if one includes well-vetted mechanisms and/or agreed-upon hypotheses, does the system as a whole behave as everyone thinks it should (i.e., is the behavior ‘possible’ or ‘impossible’)? If so, then the model suggests (but does not prove) that the underlying hypotheses are correct to some level of precision. If not, then either the underlying hypotheses are incorrect, or there is something incorrect or missing in the model. Either way, knowledge is gained in a rational fashion. To practice this, there is a need for closer collaborations between modelers and experimentalists, as well as for better training of modelers in immunology and of experimental immunologists in modeling, as described further below.



More success in predictive analysis will benefit from additional data and the release of more experimental and clinical trials data to the public. The NIAID-supported Immunology Database and Analysis Portal (ImmPort) system [74] is a repository for these types of data. More than 200 data sets are now available at ImmPort (with the number rapidly growing), and many of these reflect the raw data released upon the completion of an interventional clinical study. Such raw data can be reanalyzed to defend the reproducibility of statistical testing from the original publication, and the aggregation of multiple studies through meta-analysis can be used to increase power to detect subtle signals. They could also be useful for new scientific discoveries, such as finding patient subsets that best respond to an intervention [75], or predicting why drugs might be succeeding or failing in clinical trials.

The NIAID meeting participants noted that interdisciplinary collaborations are key to realizing the iterative cycle of modeling→data→modeling. Many obstacles can derail interdisciplinary collaboration; thus, it is important to consider communication and cultural roadblocks between immunology experimentalists and computational modelers, and how to overcome them.

### Recommendations on Establishing Integrative Collaborations

The first obstacle to overcome is the identification of appropriate and suitable collaborators. The panel's recommendations focus on collaborations in which the biological questions determine the computational and mathematical tools utilized. From the immunologist's point of view, it is crucial first to identify the key immunological questions of interest and the biological scales encompassed by the questions. Finding the 'right' fit within the multivariate space of experimentalists, modelers, and biological questions will undoubtedly require multiple conversations, and the entire search process may take several iterative cycles.

When beginning the conversation, both experimentalists and modelers should be prepared for a 'culture shock' of sorts, as each attempts to describe the aspects of their field relevant to the question to be studied. Even seemingly innocuous vocabulary can become a stumbling block, perhaps best exemplified by terms such as 'model': one party may be thinking of a mouse model, while the other is thinking in terms of mathematical equations or a set of rules. In a similar vein, the limitations of each collaborator's field should be freely discussed, to avoid overestimating what either experiments or modeling can achieve.

Both experimentalists and modelers need to keep in mind that their collaborative process is necessarily a feed-forward, feed-back endeavor that originates with, and repeatedly returns to, addressing important immunological questions and leveraging critical experimental data sets as well as novel computational technologies. These data can then be fed into the development of mathematical models that represent the known biological processes and interactions of various biological entities or into the implementation of computational methods. These computational tools can then feed-back into the biological side by raising novel hypotheses and suggesting additional data to collect or to generate testable hypotheses to guide further experimental work. Additional experimental findings can then feed-forward again to refine further and calibrate mechanistic models.

### Concluding Remarks

At its core, the integration of modeling in mechanistic, preclinical, translation, and clinical immunology is centered on rigor and a foundational 'back-to-basics' approach. The lack of rigor and reproducibility of experimental studies is an area of widespread concern in both preclinical and clinical research [76,77]. When quantitatively predictive modeling is built on data, the integration of modeling into the research pipeline provides an additional quality-control mechanism that improves the rigor and reproducibility of immune studies. Furthermore, mechanistic modeling might be crucial when attempting to circumvent the difficulty in extrapolating

### Outstanding Questions

How should the various approaches to modeling be standardized to improve model sharing and integration?

Can model formatting and annotation be standardized without restricting the applicability of these approaches to divergent goals and questions?

How can model simulation platforms be refined to facilitate their use by immunology researchers not trained in modeling, without compromising scientific rigor?

How can data-driven analysis approaches and hypothesis-based mechanistic modeling be best integrated?

from reductionist mechanisms to system-level outcomes in individuals and populations with immune-mediated diseases [78].

The remarkable advances in computational modeling described above have occurred in an environment of relatively stable NIH funding opportunities for systems biology and 'omics' research, although most of the rapid progress has related to the adoption of bioinformatics techniques by the immunology research community. While 21% (38 of 178) NIH 'omics' and 'systems biology' Funding Opportunity Announcements (FOA) identified at the time of writing this article are standing requests that are open over multiple years, the rare mechanistic immune modeling FOAs have been overwhelmingly one-time opportunities. If the broader immunology research enterprise is to reap the synergies of modeling, key changes are required to achieve a cohesive, growing, and efficient interdisciplinary effort. Mechanisms to encourage progress on outstanding sections, such as the further development of software platforms and interchange standards that improve model reuse and hypothesis testing, especially by researchers with limited mathematical and/or computational training, will assist the beneficial integration of modeling into immunology research (see Outstanding Questions). Improved understanding of the value and limitations of modeling by the experimental immunology research community is key; experiential learning could be encouraged by student and fellowship training, workshops, and visiting scientist programs. Finally, stable large-scale (P01, U19, U24, BAA) and smaller (R01, R21, SBIR/STTR) funding mechanisms are needed to encourage the integration of simulation and experimental immunology, ultimately leading to the improved therapeutic development and clinical outcomes. Changing the culture, language, tools, and research enterprise to improve the synergy of data mining, mechanistic modeling, and data-generating immunology research techniques should accelerate the translation into improved diagnosis and treatment of infectious, inflammatory and immune-mediated diseases.

### Acknowledgments

We gratefully acknowledge the NIAID for sponsoring this workshop and the efforts of the organizing committee, including: Timothy Gondre-Lewis, Katarzyna Bourcier, Joseph Breen, Wendy Davidson, Alison Augustine-Deckhut, Mark Robien, and Charles Hackett, as well as Isaac Rodriguez-Chavez from NIDCR for his valuable suggestions during the meeting. We thank the many attendees and participants in the workshop for their input and Jill Gregory for creating the illustration. We also thank Waj Mehal, Raquel Hontecillas, Vida Abedi, and Christophe Benoist for critical comments and/or textual suggestions. We apologize for the absence of the discussion of many approaches and of many important references due to space and citation constraints.

### Resources

<sup>i</sup> <http://grants.nih.gov/grants/guide/rfa-files/RFA-CA-15-014.html>

<sup>ii</sup> <https://nimml.org/tools/copasi-suite>

<sup>iii</sup> [www.nimml.org/tools/enisi-suite](http://www.nimml.org/tools/enisi-suite)

<sup>iv</sup> [www.niaid.nih.gov/research/simmune-software](http://www.niaid.nih.gov/research/simmune-software)

<sup>v</sup> <http://mcell.org>

<sup>vi</sup> <http://visualizlab.org>

<sup>vii</sup> <http://vcell.org>

### Supplemental Information

Supplemental information associated with this article can be found online at [doi:10.1016/j.it.2016.11.006](https://doi.org/10.1016/j.it.2016.11.006).

### References

1. The 1000 Genomes Project Consortium (2015) A global reference for human genetic variation. *Nature* 526, 68–74
2. Lek, M. *et al.* (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285–291
3. Farh, K.K. *et al.* (2015) Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* 518, 337–343
4. Germain, R.N. *et al.* (2011) Systems biology in immunology: a computational modeling perspective. *Annu. Rev. Immunol.* 29, 527–585
5. DiMasi, J.A. *et al.* (2016) Innovation in the pharmaceutical industry: new estimates of R&D costs. *J. Health Econ.* 47, 20–33
6. Terstappen, G.C. and Reggiani, A. (2001) *In silico* research in drug discovery. *Trends Pharmacol. Sci.* 22, 23–26

7. Holford, N. *et al.* (2010) Clinical trial simulation: a review. *Clin. Pharmacol. Ther.* 88, 166–182
8. FDA (2004) *Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products*, FDA
9. Mayya, V. and Dustin, M.L. (2016) What scales the T cell response? *Trends Immunol.* 37, 513–522
10. Tsang, J.S. (2015) Utilizing population variation, vaccination, and systems biology to study human immunology. *Trends Immunol.* 36, 479–493
11. Schadt, E.E. *et al.* (2005) An integrative genomics approach to infer causal associations between gene expression and disease. *Nat. Genet.* 37, 710–717
12. Janes, K.A. and Yaffe, M.B. (2006) Data-driven modelling of signal-transduction networks. *Nat. Rev. Mol. Cell Biol.* 7, 820–828
13. Newell, E.W. *et al.* (2012) Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8<sup>+</sup> T cell phenotypes. *Immunity* 36, 142–152
14. Janes, K.A. and Lauffenburger, D.A. (2006) A biological approach to computational models of proteomic networks. *Curr. Opin. Chem. Biol.* 10, 73–80
15. Zeevi, D. *et al.* (2015) Personalized nutrition by prediction of glycemic responses. *Cell* 163, 1079–1094
16. Yosef, N. *et al.* (2013) Dynamic regulatory network controlling TH17 cell differentiation. *Nature* 496, 461–468
17. Gorenshsteyn, D. *et al.* (2015) Interactive big data resource to elucidate human immune pathways and diseases. *Immunity* 43, 605–614
18. Arazi, A. *et al.* (2013) Human systems immunology: hypothesis-based modeling and unbiased data-driven approaches. *Semin. Immunol.* 25, 193–200
19. Callard, R.E. and Yates, A.J. (2005) Immunology and mathematics: crossing the divide. *Immunology* 115, 21–33
20. Forrest, S. and Beauchemin, C. (2007) Computer immunology. *Immunol. Rev.* 216, 176–197
21. Narang, V. *et al.* (2012) Systems immunology: a survey of modeling formalisms, applications and simulation tools. *Immunol. Res.* 53, 251–265
22. Altan-Bonnet, G. and Germain, R.N. (2005) Modeling T cell antigen discrimination based on feedback control of digital ERK responses. *PLoS Biol.* 3, e356
23. Artomov, M. *et al.* (2010) Only signaling modules that discriminate sharply between stimulatory and nonstimulatory inputs require basal signaling for fast cellular responses. *J. Chem. Phys.* 133, 105101
24. Chakraborty, A.K. *et al.* (2009) Molecular origin and functional consequences of digital signaling and hysteresis during Ras activation in lymphocytes. *Sci. Signal.* 2, pt2
25. Das, J. *et al.* (2009) Digital signaling and hysteresis characterize ras activation in lymphoid cells. *Cell* 136, 337–351
26. Abboud, A. *et al.* (2016) Computational analysis supports an early, type 17 cell-associated divergence of blunt trauma survival and mortality. *Crit. Care Med.* 44, e1074–e1081
27. Abedi, V. *et al.* (2015) ENISI multiscale modeling of mucosal immune responses driven by high performance computing. In *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 680–684, IEEE
28. Wei, X. *et al.* (1995) Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* 373, 117–122
29. Ho, D.D. *et al.* (1995) Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* 373, 123–126
30. den Braber, I. *et al.* (2012) Maintenance of peripheral naive T cells is sustained by thymus output in mice but not humans. *Immunity* 36, 288–297
31. Petrovas, C. *et al.* (2013) High production rates sustain *in vivo* levels of PD-1 high simian immunodeficiency virus-specific CD8 T cells in the face of rapid clearance. *J. Virol.* 87, 9836–9844
32. Johnson, P.L. *et al.* (2011) Vaccination alters the balance between protective immunity, exhaustion, escape, and death in chronic infections. *J. Virol.* 85, 5565–5570
33. Basak, S. *et al.* (2012) Lessons from mathematically modeling the NF-kappaB pathway. *Immunol. Rev.* 246, 221–238
34. Shih, V.F. *et al.* (2012) Control of RelB during dendritic cell activation integrates canonical and noncanonical NF-kappaB pathways. *Nat. Immunol.* 13, 1162–1170
35. Wang, J. (2015) Landscape and flux theory of non-equilibrium dynamical systems with application to biology. *Adv. Physics* 64, 1–137
36. Seiden, P.E. and Celada, F. (1992) A model for simulating cognate recognition and response in the immune system. *J. Theor. Biol.* 158, 329–357
37. Ray, J.C.J. *et al.* (2009) Synergy between individual TNF-dependent functions determines granuloma performance for controlling *Mycobacterium tuberculosis* infection. *J. Immunol.* 182, 3706–3717
38. Segovia-Juarez, J.L. *et al.* (2004) Identifying control mechanisms of granuloma formation during *M. tuberculosis* infection using an agent-based model. *J. Theor. Biol.* 231, 357–376
39. Keşmir, C. and De Boer, R.J. (2003) A spatial model of germinal center reactions: cellular adhesion based sorting of B cells results in efficient affinity maturation. *J. Theor. Biol.* 222, 9–22
40. Eroni, S. *et al.* (2007) Emergent dynamics of thymocyte development and lineage determination. *PLoS Comput. Biol.* 3, e13
41. Carbo, A. *et al.* (2013) Predictive computational modeling of the mucosal immune responses during *Helicobacter pylori* infection. *PLoS One* 8, e73365
42. Mei, Y. *et al.* (2015) Multiscale modeling of mucosal immune responses. *BMC Bioinform.* 16 (Suppl 12), S2
43. Beauchemin, C. *et al.* (2005) A simple cellular automaton model for influenza A viral infections. *J. Theor. Biol.* 232, 223–234
44. Folcik, V.A. *et al.* (2007) The Basic Immune Simulator: an agent-based model to study the interactions between innate and adaptive immunity. *Theor. Biol. Med. Model.* 4, 39
45. Karlebach, G. and Shamir, R. (2008) Modelling and analysis of gene regulatory networks. *Nat. Rev. Mol. Cell Biol.* 9, 770–780
46. Miskov-Zivanov, N. *et al.* (2013) The duration of T cell stimulation is a critical determinant of cell fate and plasticity. *Sci. Signal.* 6, ra97
47. Saez-Rodriguez, J. *et al.* (2007) A logical model provides insights into T cell receptor signaling. *PLoS Comput. Biol.* 3, e163
48. Georgescu, C. *et al.* (2008) A gene regulatory network armature for T lymphocyte specification. *Proc. Natl. Acad. Sci. U. S. A.* 105, 20100–20105
49. Naldi, A. *et al.* (2010) Diversity and plasticity of Th cell types predicted from regulatory network modelling. *PLoS Comput. Biol.* 6, e1000912
50. Mariani, L. *et al.* (2010) Short-term memory in gene induction reveals the regulatory principle behind stochastic IL-4 expression. *Mol. Syst. Biol.* 6, 359
51. Buchholz, V.R. *et al.* (2013) Disparate individual fates compose robust CD8<sup>+</sup> T cell immunity. *Science* 340, 630–635
52. Althaus, C.L. and Boer, R.J.D. (2008) Dynamics of Immune escape during HIV/SIV infection. *PLoS Comput. Biol.* 4, e1000103
53. Read, E.L. *et al.* (2012) Stochastic effects are important in intrahost HIV evolution even when viral loads are high. *Proc. Natl. Acad. Sci. U. S. A.* 109, 19727–19732
54. Schiffer, J.T. *et al.* (2009) Frequent release of low amounts of herpes simplex virus from neurons: results of a mathematical model. *Sci. Transl. Med.* 1, 7ra16
55. Mayer, A. *et al.* (2015) How a well-adapted immune system is organized. *Proc. Natl. Acad. Sci. U. S. A.* 112, 5950–5955
56. Kleinstein, S.H. and Singh, J.P. (2003) Why are there so few key mutant clones? The influence of stochastic selection and blocking on affinity maturation in the germinal center. *Int. Immunol.* 15, 871–884
57. Wang, S. *et al.* (2015) Manipulating the selection forces during affinity maturation to generate cross-reactive HIV antibodies. *Cell* 160, 785–797
58. Sontag, E.D. (1998) *Mathematical Control Theory. Deterministic Finite-Dimensional Systems* (2nd edn), Springer-Verlag
59. Perelson, A.S. (1978) Applications of optimal control theory to immunology. In *Recent Developments in Variable Structure Systems, Economics and Biology* (Mohler, R.R. and Ruberti, A., eds), pp. 272–287, Springer-Verlag

60. Stengel, R.F. and Ghigliazza, R. (2004) Stochastic optimal therapy for enhanced immune response. *Math. Biosci.* 191, 123–142
61. Germain, R.N. (2001) The art of the probable: system control in the adaptive immune system. *Science* 293, 240–245
62. Kepler, T.B. and Perelson, A.S. (1993) Somatic hypermutation in B cells: an optimal control treatment. *J. Theor. Biol.* 164, 37–64
63. Day, J. *et al.* (2010) Using nonlinear model predictive control to find optimal therapeutic strategies to modulate inflammation. *Math. Biosci. Eng.* 7, 739–763
64. Clermont, G. *et al.* (2004) *In silico* design of clinical trials: a method coming of age. *Crit. Care Med.* 32, 2061–2070
65. An, G. (2004) *In-silico* experiments of existing and hypothetical cytokine-directed clinical trials using agent based modeling. *Crit. Care Med.* 32, 2050–2060
66. Mi, Q. *et al.* (2007) Agent-based model of inflammation and wound healing: insights into diabetic foot ulcer pathology and the role of transforming growth factor- $\beta$ 1. *Wound Rep. Reg.* 15, 617–682
67. Kumar, R. *et al.* (2008) A mathematical simulation of the inflammatory response to anthrax infection. *Shock* 29, 104–111
68. Brown, D. *et al.* (2015) Trauma *in silico*: individual-specific mathematical models and virtual clinical populations. *Sci. Transl. Med.* 7, 285ra61
69. Ziraldo, C. *et al.* (2015) A computational, tissue-realistic model of pressure ulcer formation in individuals with spinal cord injury. *PLoS Comput. Biol.* 11, e1004309
70. Marino, S. *et al.* (2007) Differences in reactivation of tuberculosis induced from anti-TNF treatments are based on bioavailability in granulomatous tissue. *PLoS Comput. Biol.* 3, 1909–1924
71. Abedi, V. *et al.* (2016) Phase III Placebo-controlled, randomized clinical trial with synthetic Crohn's disease patients to evaluate treatment response. In *Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology – Systems & Applications* (2nd edn) (Arabnia, H. and Tran, Q-N., eds), pp. 411–428, Elsevier/MK
72. Marshall, J.C. (2004) Through a glass darkly: the brave new world of *in silico* modeling. *Crit. Care Med.* 32, 2157–2159
73. Marinkovic, T. and Oresic, M. (2016) Modeling strategies to study metabolic pathways in progression to type 1 diabetes – challenges and opportunities. *Arch. Biochem. Biophys.* 589, 131–137
74. Bhattacharya, S. *et al.* (2014) ImmPort: disseminating data to the public for the future of immunology. *Immunol. Res.* 58, 234–239
75. Nasrallah, M. *et al.* (2015) Reanalysis of the Rituximab in ANCA-Associated Vasculitis trial identifies granulocyte subsets as a novel early marker of successful treatment. *Arthritis Res. Ther.* 17, 262
76. Begley, C.G. and Ellis, L.M. (2012) Drug development: raise standards for preclinical cancer research. *Nature* 483, 531–533
77. Goldacre, B. (2016) Make journals report clinical trials properly. *Nature* 530, 7
78. An, G.V.Y. (2014) *Translational Systems Biology: Concepts and Practice for the Future of Biomedical Research*, Elsevier
79. Kumar, R. *et al.* (2004) The dynamics of acute inflammation. *J. Theoretical Biol.* 230, 145–155
80. Bissel, P. *et al.* (2016) Exploratory studies with BT-11: a proposed orally active therapeutic for Crohn's disease. *Int. J. Toxicol.* 35, 521–529