Analysis with Mathematical Models Provides Insights into Infectious Diseases

Modeling can help address questions about pathogen turnover, immune system responses, and pathogen spread within populations

Olivier Restif, Juilee Thakar, and Eric T. Harvill

uch of microbiological research on infectious agents is increasingly specialized and reductionist, viewing the host as a virtually static environment and thus sharply narrowing the context of the disease under study. We advocate an alternative quantitative approach that enables investigators to go beyond their reliance on conventional and somewhat simplified qualitative terms for describing experimental outcomes—typically, "gene *x* is required to control infection *y*."

While those conventional approaches give rise to detailed knowledge of components involved in pathogenesis and host immunity, they leave important questions unanswered. Here we consider three. (i) What are the processes such as cell death and divisions that underlie variations in bacterial numbers in sets of experiments involving infections in animals? (ii) What is the temporal organization of immune response during an infection? (iii) How do dynamic responses within an infected individual relate to the spread of pathogens between individuals and within populations?

Examples of recent model-based approaches that address these three questions provide a representative sampling of what several research groups are doing to examine the responses of complex mammalian immune systems to infectious agents. These mathematical model-augmented studies build on the strengths of reductionist approaches with which molecular microbiologists are familiar. It is our hope that increasing numbers of microbiologists will recognize the value of

these approaches and also take advantage of opportunities to work with mathematical modelers when examining system-level aspects of infectious diseases. Based on personal experience, such collaborations can be stimulating and fun, while providing important insights into infectious diseases.

How Mathematical Models Can Help in Analyzing Microbial Findings

Applying mathematical models to results of microbial experiments allows concepts and techniques from different disciplines to be combined. This multidisciplinary approach can

Summary

- Applying mathematical models to results of microbial experiments enables us to combine concepts and techniques from different disciplines, translating biological models into equations or scenarios that computers render dynamic.
- Studies of *Salmonella enterica* in mice illustrate how experiments coupled with mathematical models provide insights into host and pathogens processes that could not be observed directly.
- A model accurately simulated different immune modulatory effects and outcomes from infections by two closely related *Bordetella* species, each of which produces similar but distinct virulence factors.
- Model analysis also helps to explain how the very similar pathogens, *B. pertussis* and *B. parapertussis*, manage to coexist—and the latter to persist—despite wide use of a vaccine to control whooping cough.

Royal Society Universitv Research Fellow in the Department of Veterinary Medicine, University of Cambridge, Cambridge, United Kingdom; Juilee Thakar is Post doctoral fellow in the Department of Physics. The Pennsylvania State University, University Park; and Eric Harvill is Associate Professor of Microbiology and Infectious Disease in the Department of Veterinary and Biomedical Science, The Pennsylvania State University, University Park.

Olivier Restif is a





Average (± standard deviation) bacterial loads (CFU) at 0.5, 6, 24, and 48 hours postinoculation in the livers of experimental (red bars) and virtual (gray bars) mice. The gray curve shows the reconstructed dynamics of the average bacterial load of bacteria in the liver. (C) Distributions of the number of WITS recovered in the livers of experimental (red bars) or virtual (gray bars) mice. In (B) and (C), experimental distributions (red) were obtained from 20 mice per time point, while theoretical distributions (gray) were obtained from 1000 replicate stochastic simulations (then scaled down to match experimental numbers in panel C). For more details see Grant et al., PLoS Biology **6**:e74, 2008).

provide insights into both the underlying mechanisms and the larger-scale implications of particular microbiological studies, with mathematical models being a framework that unifies both qualitative and quantitative information.

Traditionally, experimental results are analyzed with statistical models that typically establish quantitative relations between pieces of data. However, the mechanisms responsible for these relations cannot be inferred from statistical analyses. By contrast, mathematical modeling begins with specific assumptions about underlying processes, and then assesses how well those proposed mechanisms can explain the observed data.

Thus, a mathematical model translates a biological model into equations, which computers can put into motion and then compare to the experimental findings. In practice, mathematical models consist of collections of variables (quantities that vary across time and which are directly compared to experimental measures), processes (specific events that affect the variables and represent typically unobserved biological phenomena), and parameters (constant quantities that measure the rate at which processes occur).

Fitting a mathematical model consists of looking for the values of the parameters that minimize the difference between experimental measures and the values of the model variables. It is then possible to compare different models based on how well they match the data once fitted. Naturally, the fact that a given model reproduces the data perfectly well does not mean that it provides an accurate representation of the system—most models are simplistic and must be interpreted with caution. But this approach can be useful in at least two ways: first, it can allow one to reject, or refine, those models FEATURES

that do not provide a good fit to the data and, second, it offers the opportunity to make new predictions to guide the design of subsequent experiments.

Indeed, once a mathematical model is fitted to the data from an experiment, the model can be used to simulate different experiments, some of which can be conducted to validate or refine the model. Thus, mathematical models can be integrated at all stages of the research process to improve the overall efficiency of research in microbiology. We recently applied mathematic modeling techniques to help understand: (i) the in vivo population dynamics of bacteria during an acute infection, (ii) the complex interplay between different components of the immune system leading to clearance of an infection, and (iii) the potential implications of within-host patterns of cross-immunity on epidemiological dynamics in populations.

Applying Population Ecology Models To Study Infection Dynamics

Measuring pathogen loads in infected organisms provides one way to evaluate host-pathogen interactions. Is the host controlling or clearing the infection, or is the pathogen overcoming its host? While a decrease in bacterial load looks like improved control of the infection, the patterns tend to be more complicated.

For example, a particular treatment may lead to fluctuations in pathogen loads because of dynamic underlying processes driving the bacteria, such as cell divisions and deaths as well as movement among host cells, organs, and tissues. To a large extent, these are the same processes governing the growth of any population of living organisms. The key differences lie in the detailed mechanisms responsible for those processes during an infection, such as molecular interactions with the immune system in the case of pathogens. Hence, models from population ecology can be adapted to describe some aspects of infections.

Case Study: Acute Infection with *Salmonella* in Mice

The gram-negative bacterial pathogen Salmonella enterica comprises many serovars. S. enterica Typhi, for instance, causes typhoid fever in humans, which is an important public health threat in developing countries. Although this disease can be modeled by infecting mice with *S. enterica* serovar Typhimurium, little is known about the population dynamics of the bacteria within such animals. However, to overcome this deficit, Pietro Mastroeni at the University of Cambridge in the United Kingdom and his collaborators are using mathematical models along with sophisticated experimental techniques to decipher the in vivo dynamics of such infections.

One question that they addressed is how *S*. *enterica* bacteria spread within the organs of mice. Lesions, consisting mostly of pathogen-loaded macrophages, form in the liver and the spleen, increasing both in size and in number during infection. Confocal microscopy combined with multiple fluorescent markers reveals that, while the overall bacterial load grows exponentially, most infected macrophages contain only one or two bacteria, and small numbers of macrophage cells harbor very high numbers of bacteria, even in the late stages of infection.

What leads to this skewed distribution of intracellular bacteria? Is there intrinsic heterogeneity among macrophages or rapid spread of bacteria from cell to cell that keeps overall bacterial numbers per host cell so low? These and other scenarios were translated into a mathematical model that tracks numbers of bacteria per macrophage, allowing for intracellular division of bacteria and also lysis of infected macrophages followed by infections of other host cells. The best-fit model assumed that (i) all macrophages are identical, (ii) resources limit the intracellular division rate of bacteria, and (iii) lysis of infected macrophages is a stochastic process that is independent of bacterial numbers.

This modeling approach enabled the Cambridge group to test specific predictions experimentally. Moreover, their findings raised questions about *S. enterica* serovar Typhimurium. Although regarded as an exclusively intracellular pathogen, their findings suggested that it becomes extracellular every time that a host cell bursts, a view that is in line with independent reports that antibodies are required to clear attenuated strains of *Salmonella* in mice.

The Cambridge group then investigated the global dynamics of infection, measuring levels of pathogen within whole animals. Earlier studies indicated that individual bacteria cause lesions in organs, suggesting that the clonal expansion of just a few bacteria might account for



systemic infections. A related question is whether the mouse immune system actively kills bacteria or reduces their growth. In other words, what is the relative importance of bactericidal and bacteriostatic processes in the immune defense against *S. typhimurium*?

To address those questions, the investigators needed to estimate the division and death rates of these bacteria and then to monitor their movements between organs through the bloodstream of mice. Their novel approach consisted of first inoculating mice with a mixture of eight wild-type isogenic tagged strains (WITS) of *S. typhimurium*. Each of these strains has identical phenotypes and differs from any other by a short insert of 40 nucleotides in a noncoding region of the chromosome. Then with quantitative PCR, they recorded levels of these WITS in the organs and the blood of mice at different times, making it possible to interpret decreases in strain diversity as killing of bacteria.

With a mathematical model, they next estimated the rates of division, death, and migration of S. typhimurium in different organs in the first 72 h of infection. This analysis revealed that bacteria rapidly divide and die for only the first 7 h before reaching a bottleneck in the liver and spleen, yielding two independent subpopulations that continue to divide at low rates without further dving. Spillover into the bloodstream is delayed until the second day, and those cells mix with bacterial populations in the liver and the spleen. A host NADPH oxidase proves critical for both the early bactericidal and the later bacteriostatic stages of immunity, according to their similar experiment in which the host mice were missing gp91 phox.

These studies illustrate how experiments coupled with analysis based on mathematical models provide insights into processes that are difficult or impossible to observe directly. This approach also permits the investigators to interpret their experimental observations with a new kind of rigor—in this case, leading to findings that are at odds with previous explanations.

Dynamic Networks Used To Analyze Immune Responses to Pathogens

Of course, immune system responses are seldom singular as they affect the dynamics of a bacterial pathogen during an infection. Moving beyond analysis of the effects of particular component of immunity, such as the NADPH oxidase example, requires a different approach. One such alternative is called network modeling, exemplified by the work of Reka Albert and colleagues at Pennsylvania State University. It depends on mathematical models that integrate information, mainly qualitative, on the many cells and molecules that impinge on and regulate bacterial infections.

This approach entails compiling different sources of information to decipher causal relationships among these many components. These relationships are constructed into a network. Components within this network of infection include the pathogen and its virulence factors as well as host components such as immune cells, antibodies, and cytokines. They are linked in pairwise fashion, in each case specifying what member of the pair activates or inhibits the other. This static network is then incorporated into a dynamic model that tracks quantitative changes in these different actors during the course of an infection. Such models are especially intuitive because they entail converting a static description of a biological system, much as one might write on a blackboard, into a fluid mathematical model.

Case Study: Controlling a *Bordetella* Infection in Mice

Bordetellae are gram-negative coccobacilli that infect a range of mammalian species. *Bordetella pertussis*, which causes whooping cough in humans, apparently derived from the progenitor of a closely related species, *B. bronchiseptica*, with which it shares a variety of virulence factors. Although these two species modulate host immune responses, those effects involve different mechanism. In both cases, immunocompetent hosts can usually control and clear these pathogens.

Although various *Bordetella* species infect mice, none of these infections exactly models whooping cough in humans. However, they do provide a means for studying the infectious process and interactions between these bacterial pathogens and host immune responses, including time courses of bacterial loads and densities of immune cells and cytokines. For instance, various host cell types are recruited to *Bordetella*-infected lungs and can produce cytokines such as interferon- γ (IFN- γ).

To study how the different cell types and cytokines work in concert to contribute to the



Features and results from the network model of within-host immune interactions (reproduced from Thakar et al., PLoS Computational Biol. **3**:e109, 2007). Subnetworks depicting innate immune interactions between host and (a) *B. bronchiseptica* and (b) *B. pertussis* virulence factors. Network nodes denote immune components and virulence factors, and edges represent interactions and processes. The edges are classified into two regulatory effects, activation and inhibition, and are represented by incoming black arrows and incoming red blunt segments, respectively. Dynamic behavior of the network is simulated by developing species-specific Boolean models. The wild-type activation pattern of (a) *B. bronchiseptica* and (b) *B. pertussis* is shown. Each colored pattern is a square grid representing the state of the nodes on the y-axis versus time steps on the x-axis. The colored squares correspond to active nodes (having state 1) at the represented time step. One time step of the simulation corresponds to 1 d to 2 d of the real infection. Dashed lines represent three stages of infections discussed in the text. The "in silico" simulation depicting an earlier clearance of (e) *B. bronchiseptica* due to the adoptive transfer of antibodies is contrasted with the no effect on (f) *B. pertussis* clearance.

dynamics of the infection proves challenging. Hence, we have adapted network-modeling approaches to decipher interactions between and among immune components and virulence factors and also to analyze time-dependent contributions of specific immune components (Fig. 2). The networks, which represent a static picture of the immune response, were incorporated into a dynamic model of the sequential events in lungs during the infection by *B. bronchiseptica* or *B. pertussis*. The model helps to illuminate three stages of the infection: the first is the innate immune response, the second is when bacterial virulence factors modulate the immune response, and the third involves highly efficient immune responses that lead to clearance of the bacteria. The model delineates when different immune components contribute to the host responses, and it helps to portray what happens when the immune system is perturbed.

The model also depicts differences in the time course of infections for hosts that previously encountered the pathogens versus hosts that had not. For instance, the model predicted that convalescent hosts can rapidly clear both pathogens, but that antibody transfer would rapidly clear *B. bronchiseptica* but not *B. pertussis* infection. This prediction held up when tested experimentally. Thus, the model accurately simulated different immune modulatory effects and





Features and results from the mathematical model on asymmetric cross-immunity in bordetellae (reproduced from Restif et al., Parasitology **135**:1517–1529, 2008). (A) Structure of the model. The human population is divided into compartments representing the status of individuals: susceptible to both *B. pertussis* and *B. parapertussis* (S), vaccinated against *B. pertussis* and not yet infected (Vp), currently infected with *B. pertussis* (primary infection lp1, secondary infection lp2), *B. parapertussis* (lpp1 and lpp2) or both (lco), or recovered from infection and currently immune to *B. pertussis* only (Rp), *B. parapertussis* only (Rpp) or both (R). Arrows show possible transitions (thick black arrows: infection, thin black arrows: recovery, dotted gray arrows: loss of immunity, dashed black arrows: births). (B) Simulated population dynamics following introduction of anti-pertussis vaccination on year 0 (vaccine uptake gradually increases from 0 to 75% by year 20 and then remains constant. Black lines: joint prevalence of *B. pertussis* (solid line) and *B. parapertussis* (dashed line). Gray line: prevalence of *B. parapertussis*. Here vaccination and primary infection with *B. pertussis* confers stronger cross-protection than vaccination, resulting in an increase in *B. parapertussis* prevalence following introduction of vaccination.

outcomes from infections caused by two closely related bacterial pathogens, each of which produces similar but distinct virulence factors.

With additional nodes, the model can begin to account for how different components of the immune system, such as T helper cells, change the bacterial numbers during such infections. For example, dynamics of the network can be different due to the different kinetic parameters of various immune components, illustrating how much the immune response varies across different hostpathogen systems. Moreover, the model reveals the importance of local versus systemic interactions in the dynamics of immune responses. The dominance of one T helper cell type over another depends on local cytokines. The model also suggests that other immune components play a role in regulating T helper subtypes. This type of modeling thus provides a range of insights into virulence,

pathogenesis, and host adaptations to diseasecausing microorganisms.

Modeling Can Help in Understanding How Infections Spread

For both practical and basic reasons, studies of infectious diseases in animals tend to focus on what happens to the host, without paying much attention to how pathogens are transmitted between hosts. Hence, investigators unwittingly ignore an important part of the life cycle of pathogens in ways that can be misleading, particularly when considering the evolution of pathogens. While host immune responses create strong selective pressures, other factors may play critical roles enabling or blocking transmission of pathogens between hosts. It could prove valuable to analyze how the dynamics of an FEATURES

infection within a single host can affect the epidemiology of the disease at the population level. The challenge is to reconcile these two parts of infection cycles that typically are studied under very different conditions. In particular, care is needed when extrapolating from infected animals to human populations. Moreover, the analytic approaches in microbiology and epidemiology differ in fundamental ways. For example, microbiologists measure the efficacy of immunity as a reduction in pathogen loads in vivo, whereas epidemiologists look for drops in pathogens that are circulating through population groups.

Thus, translating pathogen loads within individuals into rates of transmission within populations is a matter of guesswork. However, mathematical models can prove helpful in addressing this issue. Where empirical information is missing, they enable investigators to evaluate a range of scenarios before comparing them to empirical data or determining what data to seek.

Case Study: Cross-Immunity among *Bordetella* Species

What happens when several *Bordetella* species infect an individual simultaneously? Do two similar pathogens compete directly, and is there cross-immunity? Such questions are matters of public health interest because *B. pertussis* and *B. parapertussis* circulate freely and cause disease naturally only in human populations, and they can coinfect individuals. Meanwhile, the closely related species, *B. bronchiseptica*, aparrently circulates only in animal populations.

However, these three bacterial species can infect mice under experimentally controlled conditions. Thus, for example, mice can be sequentially infected with *B. pertussis* and *B. para*- *pertussis*. When such animals are first infected with *B. parapertussis*, this exposure protects them from subsequent challenges with either type of pathogen. However, if the mice first are infected with *B. pertussis*, the animals develop only limited cross-protection against subsequent exposure to *B. parapertussis*.

Further, the O antigen of B. parapertussis enables it to evade immunity that the host develops to *B. pertussis*. This finding suggests that *B*. parapertussis could have a competitive advantage in human populations, raising questions as to how wide use of the vaccine that protects humans against B. pertussis affects circulating B. parapertussis. To address such questions, the group at Pennsylvania State University developed a simple epidemiological model that accounts for asymmetric cross-immunity in the human population, describing it in terms of an individual's differential susceptibility to re-infection with either pathogen. Because the reported incidence of B. parapertussis infections is low compared to B. pertussis, the model assumes that the former pathogen is less fit than the latter-that is, B. parapertussis has lower infectivity or a shorter infectious period, compared to *B. pertussis*.

Use of this model helps to explain how different combinations of benefits, such as evasion of cross-immunity, and costs, such as reduced fitness, allow *B. parapertussis* to coexist with, but not outcompete, *B. pertussis*. Moreover, the model suggests that vaccinating against *B. pertussis* has little effect on circulating *B. parapertussis* so long as it evades vaccine-induced immunity in the same way as natural immunity. This prediction holds true when tested in mice, providing useful feedback between mathematical models and experiments.

SUGGESTED READING

Brown, S. P., S. J. Cornell, M. Sheppard, A. J. Grant, D. J. Maskell, B. T. Grenfell, and P. Mastroeni. 2006. Intracellular demography and the dynamics of *Salmonella enterica* infections. PLoS Biology 4:e349.

Grant, A. J., O. Restif, T. J. McKinley, M. Sheppard, D. J. Maskell, and P. Mastroeni. 2008. Modeling within-host spatiotemporal dynamics of invasive bacterial disease. PLoS Biology 6:e74.

Restif, O., D. N. Wolfe, E. M. Goebel, O. N. Bjørnstad, and E. T. Harvill. 2008 Of mice and men: asymmetric interactions between *Bordetella* pathogen species. Parasitology 135:1517–1529.

Thakar, J., M. Pilione, G. S. Kirimanjeswara, E. T. Harvill, and R. Albert. 2007. Modeling systems-level regulation of host immune responses. PLoS Computational Biol. 3:e109.

Thakar, J., A. Saadatpour-Moghaddam, E. T. Harvill, and R. Albert. 2008. Constraint-based network model of pathogenimmune system interactions. J. R. Soc. Interface, doi:10.1098/rsif.2008.0363, in press.

Wolfe, D. N., E. M. Goebel, O. N. Bjørnstad, O. Restif, and E. T. Harvill. 2007 The O antigen enables *Bordetella* parapertussis to avoid *Bordetella pertussis*-induced immunity. Infect. Immun. 75: 4972–4979.