SPECIAL ISSUE REVIEW

Editorial: Mathematical modelling of infectious diseases

ANDY FENTON*

Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK

(*Received 25 January 2016; accepted 26 January 2016; first published online 30 March 2016*)

Key words: Ecology, evolution, data integration, model testing, population dynamics, theory, model complexity.

The field of disease ecology – the study of the spread and impact of parasites and pathogens within their host populations and communities – has a long history of using mathematical models. Dating back over 100 years, researchers have used mathematics to describe the spread of disease-causing agents, understand the relationship between host density and transmission and plan control strategies. The use of mathematical modelling in disease ecology exploded in the late 1970s and early 1980s through the work of Anderson and May (Anderson and May, 1978, 1981, 1992; May and Anderson, 1978), who developed the fundamental frameworks for studying microparasite (e.g. viruses, bacteria and protozoa) and macroparasite (e.g. helminth) dynamics, emphasizing the importance of understanding features such as the parasite's basic reproduction number (R_0) and critical community size that form the basis of disease ecology research to this day. Since the initial models of disease population dynamics, which primarily focused on human diseases, theoretical disease research has expanded hugely to encompass livestock and wildlife disease systems, and also to explore evolutionary questions such as the evolution of parasite virulence or drug resistance. More recently there have been efforts to broaden the field still further, to move beyond the standard 'one-host-one-parasite' paradigm of the original models, to incorporate many aspects of complexity of natural systems, including multiple potential host species and interactions among multiple parasite species.

Given this rich history in development of fundamental theory, the applied aspect of the questions being asked, and the extensive data collected on many disease systems (particularly human and livestock, but increasingly wildlife), the field of disease ecology represents one of the richest areas of research at the interface of pure ecological theory and data. However, as with any field, there is the danger that those working in one particular area, or using one particular technique, will end up communicating primarily with themselves, resulting in rather distinct and discrete silos of research, with little cross-talk between sub-disciplines. This would be particularly counter-productive in the case of theory, where increasingly abstract theory could be developed with little connection to 'real world' processes. For such an applied topic as disease control this could lead to theory that is irrelevant, or unable to be connected to the kind of data available to empiricists working on natural disease systems. For these reasons, this Special Issue was put together with the aims of: (1) presenting a snapshot (although by no means exhaustive) of the current state of theoretical disease ecology and evolutionary research; (2) assessing how closely connected are the theoretical and empirical research areas on each topic; and (3) providing ideas for how those theoretical and empirical research areas can be brought closer together. To achieve this, each set of authors were asked to write personal reviews of their relevant fields, emphasizing existing or potential links between models and data as appropriate, and outlining areas for future closer integration of theory and data. As such, although the authors of these various papers are primarily theoreticians, the intention is that these papers should be accessible and relevant to both theoreticians and empiricists, with a combined interest of advancing each area of research in the broader field of disease ecology.

Of the 10 papers in this Special Issue, eight were broadly ecological (de Leo *et al*. 2016; Garnier *et al*. 2016; Magpantay *et al*. 2016; McCallum, 2016; Norman *et al*. 2016; Park *et al*. 2016; Viana *et al*. 2016; Wearing *et al*. 2016) with two focusing primarily on evolutionary aspects (Cressler *et al*. 2016; Greischar *et al*. 2016). These papers covered a range of topics, from highly conceptual (e.g. de Leo *et al*. 2016) to applied work on various specific systems, covering human diseases (Magpantay *et al*. 2016), livestock (Garnier *et al*. 2016) and wildlife conservation (McCallum, 2016). McCallum (2016) reviews the practical application of models to guide alternative management strategies (e.g. vaccination or culling) of wildlife diseases for conservation.

* Corresponding author. Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK. E-mail: a.fenton@liverpool.ac.uk

Parasitology (2016), **143**, 801–804.
doi:10.1017/S0031182016000214

He emphasizes that models can play a valuable role, both in exploring alternative hypotheses about the impact of disease, but also in the exploration of alternative management strategies. In particular, McCallum (2016) describes how models have allowed a range of vaccination and culling strategies to be explored, both in terms of general feasibility, and for setting thresholds for disease eradication or suppression. A key point is that typically wildlife disease systems are highly data-poor, in comparison to human or livestock systems. Often even broad estimates of basic demographic rates of the host (e.g. fecundity and age-specific mortality) are unknown, and problems of observational error are considerable (e.g. the correspondence between the observed prevalence of infection in a sample of hosts and the true prevalence in the host population as a whole is often unknown). This lack of information makes it essential to rigorously explore model sensitivity to variations or uncertainty in those parameters. Viana *et al*. (2016) continue this theme, emphasizing that it is usually impossible to directly observe the fundamental process of disease systems: transmission. Therefore, transmission rates have to be inferred from changes in (for example) seroprevalence, which may be an unreliable indicator of 'true' infection status or the timing of infection. For these reasons Viana *et al*. (2016) advocate using a combination of data types to infer transmission, in particular combining serological data, which is commonly available, with finer resolution type-specific genetic data, which is typically sparser. They show how such data can be combined within a Bayesian framework to infer the contributions of different potential host species to overall transmission, and human infection risk, of *Brucella* bacteria, the causative agents of brucellosis. A key point to emerge is that it is crucial to consider how uncertainties at each stage of the transmission process propagate, and to use an analytical framework that handles those uncertainties appropriately; standard GLMs aimed at identifying sources of human infection from observed prevalences in other species performed very poorly, likely due to inappropriate handling of uncertainty. Relatedly, Magpantay *et al*. (2016) applied recently developed statistical inference techniques (Ionides *et al*. 2015; King *et al*. 2015) to time series data of pertussis incidence to explore alternative hypotheses relating to observed occurrences of vaccine failure. By fitting different stochastic models that corresponded to different hypothesized mechanisms of vaccine failure, Magpantay *et al*. (2016) found strong evidence that vaccinated individuals were just as transmissible as unvaccinated individuals, although susceptibility and symptoms were reduced. This work shows how the combination of data and alternative mathematical models, again within an appropriate analytical framework, can reveal; otherwise hidden patterns within the data, enabling rigorous evaluation of alternative hypotheses into the underlying mechanisms driving those patterns.

Three of the papers specifically explored issues relating to vector-borne parasites (Norman *et al*. 2016; Park *et al*. 2016; Wearing *et al*. 2016). Such systems involve many important diseases of both humans (e.g. malaria, chikungunya and Lyme disease) and livestock (e.g. bluetongue and louping ill). They are typically also highly complex, involving the parasite, the vector and often multiple potential host species; as such mathematical models are essential in helping to understand and manage the dynamics of these systems. Norman *et al*. (2016) review models of tick-borne diseases (TBDs), and emphasise the distinction between studies that model the tick dynamics alone, primarily focusing on environmental drivers of those dynamics (e.g. climatic effects, seasonality, etc.), and those studies that explicitly model both the tick and the tick-borne pathogen dynamics. These latter studies have tended to explore how host community composition, particularly the complex interplay of pathogen and tick competent and non-competent hosts, affects TBD dynamics. These models show that overall TBD occurrence depends on the combination of separate tick and pathogen thresholds for persistence [effectively, both $R_{0,\text{ticks}} > 1$ and $R_{0,\text{pathogen}} > 1$ (Gilbert *et al*. 2001)]. As such, non-intuitive changes in pathogen infection risk can occur from changes in relative or absolute abundance of the different host species in the community. Wearing *et al*. (2016) emphasize that many models consider just the endemic phase of diseases, ignoring the initial process of emergence and expansion into naïve populations. They show how models may be used to contrast the emergence, spread and persistence of two mosquito-borne viruses of current concern: chikungunya and dengue. Wearing *et al*. (2016) highlight that vector-borne pathogen occurrence depends on a series of requirements: the presence of a suitable vector, the introduction of the pathogen and the right ecological conditions for onward local transmission. As in Norman *et al*. (2016), Wearing *et al*. (2016) show how different modelling approaches have been used to explore the dynamics of these stages separately. By separating these stages out and considering, for example, how climatic factors, human activity and habitat alteration differently affects each one, the modelling approaches described by Wearing *et al*. (2016) potentially provide an appropriate framework for risk assessment and prediction of future emergence events. Park *et al*. (2016) expands some of these ideas by considering an overlooked level of complexity; the role of multiple co-occurring vector species in driving and maintaining parasite transmission. They highlight that the composition of vector communities can vary spatially, and emphasize the importance of the functional diversity of vector

https://doi.org/10.1017/S0031182016000214 Published online by Cambridge University Press

community, rather than just abundance of the vector species, in driving transmission dynamics.

Relating to community effects, de Leo *et al*. (2016) extend classical one-host-one-parasite models to incorporate multiple macroparasites (helminths), and use these to assess the rules of parasite community assembly; under what conditions can one parasite invade and persist in presence of another? Specifically they argue that bioenergetic scaling laws provide a 'quantitative unifying theory of biological organization', and assess this by assuming parasite functional traits (fecundity, developmental rate, pathogenicity, etc.) scale with parasite body size. de Leo *et al*. (2016) show how host body size constrains the range of viable body sizes, mediated by a trade-off between the fecundity benefits to parasites of large size and the costs, due to the increased likelihood of host mortality. Garnier *et al*. (2016) also explore host–helminth interactions, seeking an appropriate level of model complexity to describe, in a data-friendly way, within-host helminth–immune system interaction dynamics. Many models of this interaction are either highly detailed, explicitly describing many components of the immune response, or very simplified, using highly phenomenological functions to describe the broad dynamics of the immune response. Garnier *et al*. (2016) use a model of intermediate complexity – one that is 'biologically grounded, yet parsimonious', that explicitly models both the mean and variation in worm burden and immune activation. They show the model provides a good match to data and, as in Magpantay *et al*. (2016) and Viana *et al*. (2016) the process of model fitting provides insight into underlying processes (specifically, dose-dependency in various immune-related parameters) that would not be detectable otherwise.

Finally, two of the papers explicitly consider the application of mathematical models to evolutionary questions of host–parasite relationships. Greischar *et al*. (2016) show how mathematical modelling can help interpret data and observations of within-host parasite dynamics, and help refine analytical and statistical methods of inference. As with many of the other papers in this Special Issue, they emphasize how the close integration of mathematical models and data can allow the testing of alternative hypotheses, the identification of factors underlying complex dynamics and aid the development of appropriate experiments to refine our understanding. They also emphasize the value of models getting things wrong – predicting things we don't see – which tells us we either need to change our understanding, or allows exploration of alternative 'what if' scenarios to determine under what conditions those predictions would be realized. Cressler *et al*. (2016) consider the specific issue of the evolution of parasite virulence, reviewing the state of theory, and assessing the current level of empirical support for those predictions. They show that, a few notable exceptions aside, that there is generally a poor connection between theory on the evolution of virulence and currently available data. This they argue is primarily due to each field focussing on different aspects: the theory rarely models traits that are measurable, and empiricists rarely collect data on traits predicted by theory. For example, just the definition of 'virulence' varies widely both within and between the empirical and theoretical fields, hampering connection between them.

Despite the diversity of topics and systems covered by these papers, there are a number of clear, recurring points that were consistently raised by them. In particular, the appropriate level of model complexity, or the type of model to adopt, was raised a number of times. As stated by McCallum (2016), 'The simpler a model can be whilst still answering the questions asked of it, the better'. Models based on ordinary differential equations (ODEs) are often the default option for many models, but they certainly can't answer every question, particularly when there are considerable heterogeneties present in the system [e.g. spatial or contact structure, lags in development, host population age structure, etc. (Magpantay *et al*. 2016; McCallum, 2016; Viana *et al*. 2016; Wearing *et al*. 2016)]. These complexities are often ignored by ODEs, where average rates are assumed, potentially ignoring important, and informative, sources of variability. One particular complexity raised by a number of authors was the need to explicitly consider spatial aspects of disease systems. The spatial or social connectance between individuals, heterogeneities in dispersal/movement rates, different patch types and differences in community structure between patches can all alter disease dynamics, both ecologically (e.g. McCallum, 2016; Norman *et al*. 2016; Park *et al*. 2016; Viana *et al*. 2016; Wearing *et al*. 2016) and evolutionarily (Cressler *et al*. 2016). Hence, determining the relevant spatial scale, and adopting an appropriate modelling approach (e.g. connected ODEs, metapopulation models, spatially explicit individual-based models) are major considerations.

Perhaps the most important recurring point is that there is a general need for closer integration of theoreticians and empiricists. As described in several of these papers, there are now a number of excellent analytical methods for bringing empirical data and mathematical models closer together than has previously been possible (Magpantay *et al*. 2016; McCallum, 2016; Viana *et al*. 2016). However, to make the most of these advances it is essential that both empiricists and theoreticians work closely at all stages of a project to avoid situations where, for example, a modeller is brought in at the end of an empirical study to 'do some modelling', or where a modeller scours the literature for some data to fit

https://doi.org/10.1017/S0031182016000214 Published online by Cambridge University Press

their model. As such, the 'model-guided fieldwork framework' of Restif *et al*. (2012) has much to recommend it, by which modellers and empiricists together embark on an iterative process of study design, data collection, model development, and project refinement to ensure the theory is relevant and the data collected are interpretable. Clearly the onus is on both the theoreticians and empiricists, but also on funding agencies, to enable such truly integrated interdisciplinary work, and a willingness to seek and support longer-term projects as needed.

REFERENCES

Anderson, R. M. and May, R. M. (1978). Regulation and stability of host-parasite population interactions. I. Regulatory processes. *Journal of Animal Ecology* **47**, 219–247.

Anderson, R. M. and May, R. M. (1981). The population dynamics of microparasites and their invertebrate hosts. *Philosophical Transactions of the Royal Society of London, Series B* **291**, 451–524.

Anderson, R. M. and May, R. M. (1992). *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, Oxford.

Cressler, C., McLeod, D., Rozins, C., van den Hoogen, J. and Day, T. (2016). The adaptive evolution of virulence: a review of theoretical predictions and empirical tests. *Parasitology*. doi:10.1017/S003118201500092X.

de Leo, G., Dobson, A. and Gatto, M. (2016). Body size and meta-community structure: the allometric scaling of parasitic worm communities in their mammalian hosts. *Parasitology*. doi:10.1017/S0031182015001444.

Garnier, R., Grenfell, B., Nisbet, A., Matthews, J. and Graham, A. (2016). Integrating immune mechanisms to model nematode worm burden: an example in sheep. *Parasitology*. doi:10.1017/S0031182015000992.

Gilbert, L., Norman, R., Laurenson, K. M., Reid, H. W. and Hudson, P. J. (2001). Disease persistence and apparent competition in a three-host community: an empirical and analytical study of large scale wild populations. *Journal of Animal Ecology* **70**, 1053–1061.

Greischar, M., Reece, S. and Mideo, N. (2016). The role of models in translating within-host dynamics to parasite evolution. *Parasitology*. doi:10.1017/S0031182015000815.

Ionides, E. L., Dao, N., Atchade, Y., Stoev, S. and King, A. A. (2015). Inference for dynamic and latent variable models via iterated, perturbed Bayes maps. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 719–724.

King, A. A., Ionides, E. L., Breto, C. M., Ellner, S., Kendall, B., Wearing, H., Ferrari, M. J., Lavine, M. and Reuman, D. C. (2015). Pomp: Statistical inference for partially observed Markov processes (R package). http://pomp.r-forge.r-project.org.

Magpantay, F. M., Domenech de Celles, M., Rohani, P. and King, A. (2016). Pertussis immunity and epidemiology: mode and duration of vaccine-induced immunity. *Parasitology*. doi:10.1017/S0031182015000979.

May, R. M. and Anderson, R. M. (1978). Regulation and stability of host-parasite population interactions. II. Destabilizing processes. *Journal of Animal Ecology* **47**, 249–267.

McCallum, H. (2016). Models for managing wildlife disease. *Parasitology*. doi:10.1017/S0031182015000980.

Norman, R., Worton, A. and Gilbert, L. (2016). Past and future perspectives on mathematical models of tick-borne pathogens. *Parasitology*. doi:10.1017/S0031182015001523.

Park, A., Cleavland, C., Dallas, T. and Corn, J. (2016). Vector species richness increases haemorrhagic disease prevalence through functional diversity modulating the duration of seasonal transmission. *Parasitology*. doi:10.1017/S0031182015000578.

Restif, O., Hayman, D. T. S., Pulliam, J. R. C., Plowright, R. K., George, D. B., Luis, A. D., Cunningham, A. A., Bowen, R. A., Fooks, A. R., O'Shea, T. J., Wood, J. L. N. and Webb, C. T. (2012). Model-guided fieldwork: practical guidelines for multidisciplinary research on wildlife ecological and epidemiological dynamics. *Ecology Letters* **15**, 1083–1094.

Viana, M., Shirima, G., Kunda, J., Fitzpatrick, J., Kazwala, R., Buza, J., Cleaveland, S., Haydon, D. and Halliday, J. (2016). Integrating serological and genetic data to quantify cross-species transmission: brucellosis as a case study. *Parasitology*. doi:10.1017/S0031182016000044.

Wearing, H., Robert, M. and Chrsitofferson, R. (2016). Dengue and chikungunya: modeling the expansion of mosquito-borne viruses into naïve populations. *Parasitology*. doi:10.1017/S0031182016000421.