

Special Issue: Long-term ecological research

Never mind the length, feel the quality: the impact of long-term epidemiological data sets on theory, application and policy

Pejman Rohani^{1,2,4} and Aaron A. King^{1,2,3,4}

¹ Department of Ecology & Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA

² Center for the Study of Complex Systems, University of Michigan, Ann Arbor, MI 48109, USA

³ Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA

⁴ Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA

Infectious diseases have been a prime testing ground for ecological theory. However, the ecological perspective is increasingly recognized as essential in epidemiology. Long-term, spatially resolved reliable data on disease incidence and the ability to test them using mechanistic models have been critical in this cross-fertilization. Here, we review some of the key intellectual developments in epidemiology facilitated by long-term data. We identify research frontiers at the interface of ecology and epidemiology and their associated data needs.

Historical background

John Graunt, who pioneered the collection of vital statistics in the 17th century, was a seminal figure in epidemiology [1]. Because he was interested in an early warning system for the spread of bubonic plague, he focused on disease mortality, but his systematic analyses of vital statistics were highly influential and led to programs for regular documentation of demographic fluxes and causes of death [2]. One of Graunt's most substantial legacies is a wealth of data on infectious disease morbidity and mortality systematically collected from the 16th century. In England and Wales, for example, notifications of deaths attributable to several high-profile diseases (e.g. measles, whooping cough, diphtheria, scarlet fever, plague) have been recorded since 1836 [3]. Figure 1 shows spatially replicated data spanning different eras for weekly notifications of whooping cough deaths and incidence for the largest population centers in England and Wales from the first years of the 20th century [4]. Similarly, the US Public Health Service has published the Weekly Abstract of Sanitary Reports since 1878 [5] and comparable collections are available in many other countries.

Although long-term epidemiological data sets vary in reporting fidelity, frequency and duration, relative to most other ecological time series they tend to be long and highly resolved. Examples include excellent data sets on cholera and malaria mortality in the former British India [6],

recent incidence data on dengue hemorrhagic fever from Thailand [7], raccoon rabies data from the eastern USA [8], and bubonic plague in gerbils in Kazakhstan [9]. This wealth of ecological data is perhaps rivaled only by fisheries data [10], trapping data for small mammals [11], and data on forest insect outbreaks [12]. Another distinguishing feature of infectious disease data is the availability of parallel information, often including details on host demography, immunization practices, and societal and behavioral changes. These types of information have been invaluable in placing observed epidemiological patterns within their ecological context (see below).

Long-term data were instrumental in the development of epidemiological ideas in the late 19th and early 20th centuries, when a number of researchers explored the roles of seasonality, immunity and competition in infectious disease dynamics [13,14]. At the same time, fundamental theoretical insights led to formulation of the classic mathematical models that underpin modern epidemiological research, including Hamer's presentation of the so-called susceptible–infectious–recovered (SIR) model [15], the development of the first malaria transmission model by Ross [16] and the influential work of Kermack and McKendrick [17] on the threshold properties of the SIR system. In the mid-20th century, Bartlett's ground-breaking analyses of measles epidemics and their extinction frequency led to the important concept of the critical community size (the smallest host population size above which the pathogen persists [18]) and the dynamic impact of demographic noise in amplifying fluctuations and sustaining oscillations in SIR models [19]. Epidemiological theory was further boosted by the seminal contributions of Dietz [20] and Bailey [21]. In many ways, however, the true marriage of epidemiological theory and long-term data sets had to wait for the consummate work of Anderson and May. Starting with their compelling 1979 treatise, they drew attention to the important parallels between ecological theory (especially predator–prey systems) and infectious disease [22,23]. They subsequently published a series of elegant studies in which meaningful, policy-relevant

Corresponding author: Rohani, P. (rohani@umich.edu).

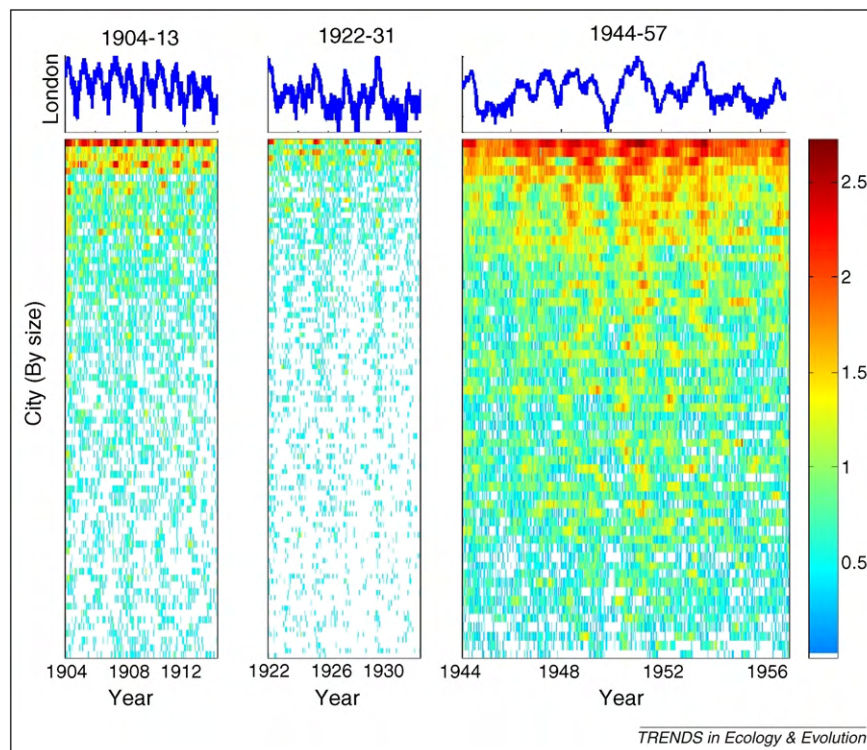


Figure 1. Long-term data on whooping cough (pertussis) in the largest cities of England and Wales [4]. Weekly notifications of pertussis deaths from the largest 95 towns and cities in England and Wales for (a) 1904–1913 and (b) 1922–1931. (c) Weekly case data for the 60 largest cities for 1944–1957. The line graph at the top of each panel shows data for London and the colored panels present spatial log-transformed data. The figures demonstrate changes in epidemiological patterns of pertussis through the decades. In London, for example, pertussis outbreaks during 1904–1913 were annual, gave way to biennial cycles in 1922–1931 and were characterized by a mixture of annual and multiennial oscillations after World War II. The color panels show that in between large outbreaks, pertussis became locally extinct in small populations, as illustrated by white regions.

conclusions were extracted from epidemiological data (summarized in [24]). Subsequently, infectious disease ecology has burgeoned as a field, becoming a prime testing ground for ecological concepts and theory [25,26]. Currently, the crosstalk between ecology and epidemiology is exciting and productive: examination of epidemiological data from an ecological perspective informs public health issues [27–30] and methodology developed for dealing with long-term epidemiological data sets are usefully applied in ecological contexts [e.g. 31,32]. The abundance of long-term data unquestionably continues to play a critical role in this blossoming.

In this opinion piece, we review the major epidemiological lessons learned from long-term data, outline some of the outstanding challenges to epidemiological theory, and identify an urgent need for new long-term data sets that differ in type and scale.

Lessons learned (so far)

To address issues of causality in natural systems, dynamical models are indispensable [33]. The most natural and rigorous means of evaluating such models is to apply them to long-term time-series data. In the ecology of many infectious diseases, two circumstances make such models relatively easy to formulate. The first is the pronounced separation between the generation times for micro-parasites (viruses, bacteria, protozoa) and those of their hosts. The second is the close ecological connection between many obligate specialist parasites and their hosts. Even given the relative simplicity of the ecology in such cases, it is

remarkable that the very simplest models have proved surprisingly efficient at explaining data. This is in stark contrast to the experience in ecology where, by and large, the simplest models are thought to be of limited use in explaining nature, with the notable exception of well-studied laboratory systems [34–36]. This degree of success probably stems from a combination of factors. First, at the macroscopic scale (e.g. long-term epidemic dynamics in a metropolitan centre), many infectious disease systems are characterized by well-understood biology and a reasonably simple natural history (host specificity, known durations of latency and infectiousness, and long-lasting immunity). Second, and intriguingly, many heterogeneities seem to average out in such infectious disease systems so that admittedly oversimplified descriptions often effectively capture prominent dynamical patterns [37].

Nonlinearity, seasonality and stochasticity

One of the earliest attempts to apply an epidemiological model to data was by Hamer [15], who noted the inconsistency between the constant prevalence predicted by the simplest transmission model and the violent, high-amplitude oscillations observed in notifications of measles case in London. Hamer speculated that a missing component of the model was rhythmic variation in the number of susceptible subjects. This topic was re-examined by Soper, whose exploration of Glasgow measles data led him to suspect seasonal variation in transmission rates that could be attributed to the opening and closing of schools [14]. This conclusion that epidemics of measles epidemics (and

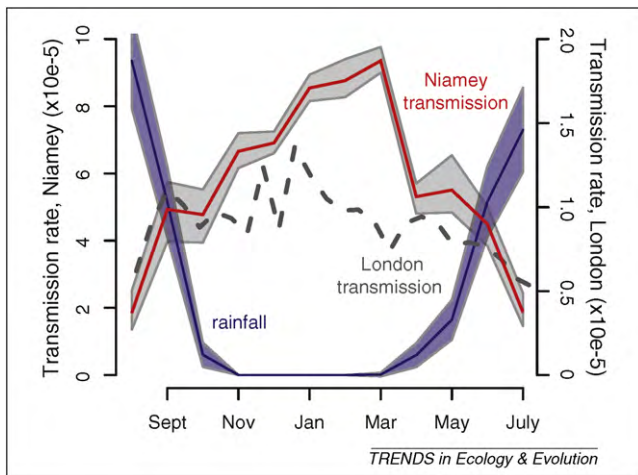


Figure 2. Dynamics of measles outbreaks in Niamey, Niger. Mean monthly rainfall from 1995 to 2004 (blue) are plotted together with 62 standard deviations (blue shading). The estimated seasonal transmission rate for Niamey is depicted in red, with the shaded grey regions representing the 95% Bayesian credible intervals. The dashed line depicts the seasonality (scaled for population size) for the pre-vaccine era (1950–1968) in London for comparison. Reproduced with permission from [40].

of other childhood diseases) in large populations are driven by school-term-driven seasonal changes in contact rates has since been confirmed [38,39]. Interestingly, a recent study of measles in sub-Saharan Africa has revisited the mechanism of transmission seasonality. Ferrari *et al.* [40] explored measles incidence in Niger and suggested that seasonal human migration associated with agricultural practices is a key driver there (Figure 2). In data for Niamey, the capital of Niger, the estimated amplitude of seasonality is much greater than it is in London (Figure 2). This difference is thought to be largely responsible for the unpredictable and perhaps chaotic oscillations in measles in Niger.

Compared with the pattern in Niger, measles data for England and Wales, Europe and the US show a striking temporal regularity (Figure 3). In England and Wales, from 1950 until the introduction of national pediatric immunization in 1968, measles epidemics in larger towns and cities exhibited a predictable 2-year cycle [24,41]. The fact that the most basic SIR transmission model with school-term forcing reproduces this and other qualitative features of measles epidemics in large populations [42] has led some to comment on the essentially deterministic nature of these data [37].

As Figure 3 shows, however, infectious diseases vary in their predictability and in the appearance of stochasticity and seasonality in their dynamics [43]. For example, epidemics of chickenpox (Figure 3a) are highly regular, with a constant inter-epidemic period. Similarly, although measles outbreaks (Figure 3b) exhibit distinct dynamical shifts, these are well explained by changes in birth rates (see below) [42]. By contrast, mumps (Figure 3c) and pertussis (Figures 3 and 1d) exhibit more unpredictable dynamics, an observation that presents both challenges and opportunities: challenges, inasmuch as increased noise levels obscure patterns that reveal the underlying ecology; and opportunities, because increasing variability broadens the dynamic range of the system, thereby potentially revealing more about the mechanisms shaping the

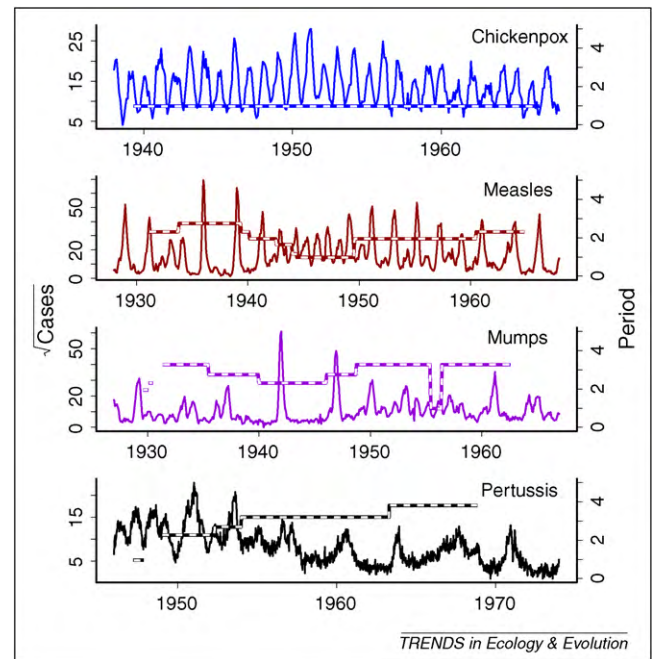


Figure 3. Depiction of long-term epidemiological data sets. Monthly case notifications from Copenhagen for (a) chickenpox, (b) measles and (c) rubella. (d) Weekly incidence of whooping cough in London (note that national immunization commenced in 1957 in England and Wales). The statistically significant dominant period through time as detected by wavelet spectral analysis is also plotted in each panel.

dynamics. From an ecological perspective, this observation is particularly interesting because of the historical, recurrent and occasionally charged debate over the relative importance of exogenous (stochastic) and endogenous (density-dependent) forces in shaping population dynamics [44,45]. Analyses of childhood disease data have shed light on when stochasticity is dynamically important, identifying the epidemiological traits (e.g. infectious period and transmission rate) that determine the outcome of seasonality and demographic noise [46,47]. Emerging theory on this front has very elegantly revealed the ingredients necessary for noise amplification in such systems [48–50] and the accompanying response to seasonality [43].

Bifurcations, chaos and natural experiments

One of the guises under which the noise versus nonlinearity controversy reappeared was the 1980s and 1990s discussion surrounding the possibility of chaos in ecological systems [51]. The idea was that the nonlinearity inherent in pathogen transmission conjoined with seasonal forcing makes childhood diseases prime candidates for chaotic dynamics. The high-profile work of Schaffer and colleagues [52,53], Sugihara and May [54] and Ellner and Turchin [55] exploited long time-series data for case notifications for childhood diseases and novel theoretical approaches to identify the fingerprints of chaos. Ultimately, unequivocal evidence of chaos in these systems remains elusive, with perhaps the most likely example being that of measles in Niger discussed above [40]. In retrospect, the lasting impact of the hunt for chaos in ecology has been methodological. The question of whether any particular ecological system is chaotic has been eclipsed by a more basic ques-

tion: what features must a mechanistic model have to explain ecological dynamics?

More recently, Earn *et al.* [42] argued that because measles is highly transmissible and elicits long-lasting immunity, epidemics are determined by the replenishment rate of the susceptible pool: so-called supply-side epidemiology. They pointed out that changes in the influx of susceptible subjects resulting from, for example, secular trends in birth rates or vaccination coverage might result in shifts in dynamical patterns and suggested that the recurrent annual epidemics of measles observed in developing nations [56] might be explained as a consequence of high fecundity, whereas the aperiodic dynamics observed in the vaccine era in developed countries, previously considered to be an example of chaos, can be more parsimoniously attributed to the interaction between stochasticity and multiple attractors.

Direct experimental confirmation of the changes predicted by bifurcation analyses and stochastic simulations have here, as elsewhere in ecology, been practically impossible. However, informative studies have exploited natural experiments of four types: (i) changes in host demography have afforded some of the most elegant and direct confirmations of predicted bifurcations [27,42,57]; (ii) the commencement of mass vaccination campaigns [58]; (iii) differential immunization strategies across countries; and (iv) comparison of outbreak data among communities (towns, countries) of different size, which has facilitated assessment of the relative importance of demographic stochasticity and extinction dynamics [37,59–61].

Metapopulations, spatial synchrony, travelling waves and transmission networks

The systematically collected and spatially resolved UK incidence data for measles and whooping cough represent a special, perhaps unique, resource. Recognizing the significance of these data for long-standing questions in population ecology, Grenfell and colleagues spearheaded a campaign to digitize such information [62]. Subsequent analysis revealed that the measles metapopulation in England and Wales in the pre-vaccine era was characterized by highly synchronous biennial outbreaks [37]. In the vaccination era, however, a significant reduction in spatial synchrony was observed [58,62]. Phase differences among outbreaks in different populations have been mooted as a possible explanation for the paradoxical observation that the critical community size has not increased substantially as a result of vaccination. If correct, this is a prominent manifestation of the ecological concept of the rescue effect [63] and, importantly, suggests a strategy of spatially targeted immunization programs [25].

Spatially explicit epidemiological models make spatiotemporal predictions and a great deal of attention has been focused on synchrony and traveling waves in disease systems [64]. Since Grenfell *et al.* [62] described such waves in measles incidence in England and Wales, they have been identified in a number of other systems, including spatially pulsed dengue outbreaks in Thailand that emanated from Bangkok [7]. A very active area of research has focused on the mechanisms of host and/or vector movement affecting such spatiotemporal patterns. It has been shown that

gravity models generate patterns that are consistent with the waves of measles outbreaks in England and Wales [65]. Such models (borrowed from transportation theory) assume that the extent of epidemiological interaction (or coupling) between two centers is determined by the geographical distance between them and their respective population sizes. By contrast, the pronounced spatial waves of seasonal H3N2 influenza epidemics in the US have been explained via coupling predicted by commuter movement between states [28].

Methodological development

Mechanistic models of epidemiological processes are nonlinear dynamical systems and as such are amenable to the tools of that field, most importantly numerical solution, stability and bifurcation analyses for deterministic models [24,66] and simulation, computation of stationary distributions and stochastic resonance for probabilistic models [46,48,50]. To date, less attention has been focused on formal statistical inference in disease systems (estimating key parameters and evaluating competing hypotheses) than on analysis of models. The most widespread approach to formal statistical inference has used the time-series SIR (TSIR) approach [39], in which the dynamics of transmission are approximated by a simple discrete-time stochastic model that can be fitted to time series data via nonlinear regression. Although this approach has been applied to a variety of diseases [43,67], and is a rough-and-ready tool, the approximations it makes begin to break down with increasing distance from the measles regime. Novel approaches based on the state–space framework have been applied to diseases such as influenza [68], cholera [29] and plant diseases [69] and show promise for dealing with strain dynamics, age structure and environmental drivers. In a state–space framework, the underlying ecoepidemiological processes that are not directly observable but are responsible for observable patterns are viewed as distinct from the observation process itself. Statistical inference on state–space models is computationally demanding but recent algorithmic breakthroughs have greatly improved the outlook for rigorous inference. Worthy of special note are tailored MCMC approaches [68,69], indirect inference approaches based on nonlinear forecasting [70] and iterated filtering [29,71,72], the latter two of which have the plug-and-play property, requiring only model simulation and obviating the need for analytical tractability of the model. These methods have yielded considerable success in infectious disease settings and will probably lead to important insights in other ecological systems.

Theoretical challenges and data needs

Here, we look to the future, outlining some of the research frontiers in disease ecology and advocating for new and different types of long-term data.

Strain evolution, phylodynamics and the community perspective

The broader ecological stage on which infectious disease dynamics play out, their community context, is increasingly recognized as critical [4,26,73]. Whereas the single host–

single pathogen paradigm deepened our understanding of the epidemiology of measles and chickenpox, for example, there are many systems for which its explanatory power is limited. Obvious examples include strain-polymorphic pathogens, such as those responsible for malaria, influenza, dengue and polio. Numerous open questions remain that can only be adequately answered by additional data. For instance, in disease systems with antigenic variability, much uncertainty surrounds the determinants of strain diversity [74], the limits to strain coexistence [75], the mechanisms responsible for the patterns of strain replacement observed, and the strength, duration and impact of immunity [76,77].

In the context of pathogens with limited diversity (e.g. cholera and dengue), existing theory is complex but relatively straightforward [75,78]. When genetic novelty continually arises, as in the case of influenza A, the theoretical challenges are greater. The so-called phylodynamics perspective attempts to infer aspects of the ecology and evolution of hosts and pathogens from the shapes of pathogen phylogenies [74,79,80]. Identification of ways to better integrate genetic and epidemiological data beyond visual or descriptive comparisons of phylogenies remains a challenge.

Beyond multi-strain systems, we now recognize polymicrobial diseases, in which transmission and pathogenicity involve interactions among distinct pathogens. Examples include opportunistic bacterial and viral infections (with numerous high-profile demonstrations in HIV/AIDS patients), periodontal diseases and some respiratory infections, including *Haemophilus influenzae* and *Streptococcus pneumoniae* [81]. Despite the recognized importance of multi-pathogen diseases in general, appropriate long-term data are still scarce. We believe that breakthroughs in the understanding of these processes will require the collation of data of different types, especially serological cross-sectional information shedding light on the kinetics of population immunological profiles and interactions among infectious agents.

Within-host dynamics

Most epidemiological models that track the prevalence of an infectious disease within a population categorize individuals according to infection and immunity status (e.g. susceptible, infectious or recovered and immune). Conceptually and mathematically, this resembles the Levins metapopulation model, in which habitat patches are either empty or fully colonized, irrespective of the population density [82]. This approach has limited value in a number of applications [83], including attempts to understand the epidemiological outcome of mixed infections [84], the evolutionary consequences of pathogen life-history traits [59,85], and the evolution of drug resistance [86]. In such cases, attempts to understand the underlying processes using mathematical models have been frustrated by the absence of long-term data at the individual infection scale [87]. Short-term or snapshot data for the initial stages of an infection are often available, but greater longitudinal information is likely to be the key to further progress for persistent infections such as HIV [88] and pathogens that can reinfect, such as influenza [89].

Another field that increasingly calls for a finer-scale understanding is immunity dynamics. Although infections by chickenpox, smallpox and morbilliviruses (including measles, rinderpest, and canine and phocine distemper viruses) induce life-long immunity, this does not seem to be the norm. Population-level data have been used to infer the dynamics of immunity relating to cholera [29,67], *H. influenzae* type B [90] and pertussis [30]. Ultimately, however, this question will need to be resolved using better specific within-host models of infection with appropriate empirical information, to some extent obtainable from animal models.

Environmental drivers

From a public health and wildlife management perspective, research on the ecology and evolution of infectious diseases would ideally translate into the development of early warning systems. This effort has for the most part focused on the use of climatological variables to inform epidemic predictions. This is largely because mechanisms linking environmental conditions, such as rainfall and temperature, to disease transmission are known [91]. For example, temperature determines the developmental rate of the malaria parasite *Plasmodium falciparum* [92] and the persistence of avian influenza viruses in aquatic environments [93], with qualitative impacts for transmission dynamics [94]. Perhaps the best-studied aspect has been the impact of climatic variables on disease vectors; ecological niche models have been used to predict the presence of vectors by reference to abiotic determinants of habitat suitability [95]. Although plausible, climatological determinants of epidemic risk and reliable early warning systems based on them require further empirical support. This has led a number of scientists to use long-term data to examine the statistical association between climate variation and the incidence and dynamics of infectious diseases [91], especially cholera [6] and Lyme disease [96], with a view to predicting the consequences of climate change [97]. These efforts are likely to be transformed by the increasingly abundant, highly spatially resolved satellite data on environmental drivers, whereas the acquisition of similarly resolved epidemiological data remains an active frontier.

Surveillance networks and policy

We have highlighted several epidemiological success stories made possible by long-term data sets accumulated through surveillance systems. Naturally, improved understanding of such systems leads to an expectation of reliable quantitative predictions. Indeed, epidemiological models are increasingly expected to quantify unobserved variables in an outbreak in progress (so-called nowcasting) and to make forecasts. For example, in the 2001 outbreak of foot-and-mouth disease in the UK, policy-makers and politicians relied heavily on mathematical modeling in their selection of epidemic control measures, with great success [98,99]. The recent H1N1 pandemic, however, provides a sobering counter-example. Following the first wave of transmission in the Northern Hemisphere in the summer of 2009, epidemiological models were scrutinized for predictions about the severity and impact of the autumn influenza season.

The models dramatically overpredicted the size of winter outbreaks and numbers of likely fatalities. This prominent setback can be largely attributed to inadequate information, both in terms of the epidemiology of the virus (particularly the case fatality rate) and the true extent of the first epidemic. The recent study by Miller *et al.* [100] goes a long way to explaining why: they demonstrate that the first wave of the epidemic in the UK probably involved ten times more children than was initially estimated. The resulting overestimation of the number of susceptible individuals seems to have led directly to overestimation of the severity of the second wave. This observation points to the need for systematic cross-sectional serological surveys as a prerequisite for better real-time modeling of the dynamics of emerging threats.

Increasing reliability of models for forecasting and now-casting will depend on better data on the contact patterns and transmission networks within and between populations. Promising recent developments in this regard include detailed studies of contact networks in Portland, Oregon [101] and self-reported mixing-pattern data for European populations [102]. We are also better placed to understand the mechanistic basis of individual movements thanks to mobile phone geolocation data [103] and the geographical dynamics of monetary currency (the Where is George? project [104]).

A universal challenge in the interpretation of incidence data is the reporting bias that arises, for example, when subclinical infections play an important epidemiological role but are less likely to be reported than severe disease [29]. More troubling are the potential dynamic interactions between reporting fidelity and epidemiological processes. For example, the 2009 H1N1 pandemic showed that sensationalism and fear can lead to increased clinic visits and thus higher reported incidence. Making the best use of long-term incidence data will require a better understanding of the interactions between disease dynamics, transmission, behavioral changes and the processes by which incidence data are recorded.

Data-sharing policies

Policies and practices for systematic sharing of and access to data have yet to be formulated and adopted by the epidemiological research community. This leads to tension between those who have invested in data collection and digitization and those who have invested in the development of analytic tools. Whatever community-wide (or, more probably, funding-agency-mandated) policies are eventually agreed on must adequately reward the initial investment in data mining and collection efforts.

Conclusions

Infectious disease ecology is a vibrant field of research. Long-term epidemiological data continue to feature prominently in the development and utility of the field. Epidemiology has furnished some of the most definitive tests of ecological principles and has proved an unrivalled test bed for ecological theory and method. In turn, epidemiology is beginning to benefit from an ecological perspective on complex multi-host and multi-pathogen systems. Continued progress will depend on our ability to gather new and

different long-term data and effectively query them using more realistic models.

Acknowledgements

We thank Marc Choisy and two anonymous reviewers for comments on the manuscript. We are grateful to Matt Ferrari and Natalia Mantilla-Beniers for graciously sharing Figures 1 and 2, respectively. The authors are supported by the Research and Policy in Infectious Disease Dynamics program of the Science and Technology Directorate, Department of Homeland Security, and the Fogarty International Center, National Institutes of Health. PR was also supported by the Vaccine Modeling Initiative of the Bill & Melinda Gates Foundation.

References

- Graunt, J. (1662) *Natural and Political Observations Made upon the Bills of Mortality*, Thomas Roycroft
- Newsholme, A. (1896) A national system of notification and registration of sickness. *J.R. Stat. Soc.* 59, 1–37
- Creighton, C. (1894) *A History of Epidemics in Britain*, Cambridge University Press
- Rohani, P. *et al.* (2003) Ecological interference between fatal diseases. *Nature* 422, 885–888
- Cliff, A. *et al.* (1998) *Deciphering Global Epidemics: Analytical Approaches to the Disease Records of World Cities, 1888–1912 (Cambridge Studies in Historical Geography)*, Cambridge University Press
- Pascual, M. *et al.* (2000) Cholera dynamics and El Niño-southern oscillation. *Science* 289, 1766–1769
- Cummings, D.A.T. *et al.* (2004) Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. *Nature* 427, 344–347
- Smith, D.L. *et al.* (2002) Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes. *Proc. Natl. Acad. Sci. U. S. A.* 99, 3668–3672
- Davis, S. *et al.* (2004) Predictive thresholds for plague in Kazakhstan. *Science* 304, 736–738
- Bjørnstad, O. *et al.* (1999) Cycles and trends in cod populations. *Proc. Natl. Acad. Sci. U. S. A.* 96, 5066–5071
- Hansen, T. *et al.* (1999) Multiannual vole cycles and population regulation during long winters: an analysis of seasonal density dependence. *Am. Nat.* 154, 129–139
- Esper, J. *et al.* (2007) 1200 years of regular outbreaks in Alpine insects. *Proc. R. Soc. B* 274, 671–679
- Ransome, A. (1880) On epidemic cycles. *Proc. Manch. Lit. Philos. Soc.* 19, 75–96
- Soper, H. (1929) The interpretation of periodicity in disease prevalence. *J. R. Stat. Soc.* 92, 34–73
- Hamer, W. (1906) Epidemic diseases in England – the evidence of variability and of persistency of type. *Lancet* 1, 733–739
- Ross, R. (1911) *The Prevention of Malaria*, John Murray
- Kermack, W. and McKendrick, A. (1927) A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A* 115, 700–721
- Bartlett, M. (1960) The critical community size for measles in the United States. *J. R. Stat. Soc. A* 123, 37–44
- Bartlett, M. (1953) Stochastic processes or the statistics of change. *J. R. Stat. Soc. C* 2, 44–64
- Dietz, K. (1976) The incidence of infectious diseases under the influence of seasonal fluctuations. *Lect. Notes Biomath.* 11, 1–15
- Bailey, N. (1975) *The Mathematical Theory of Infectious Diseases*, Charles Griffin & Company
- Anderson, R. and May, R. (1979) Population biology of infectious diseases: part I. *Nature* 280, 361–367
- May, R. and Anderson, R. (1979) Population biology of infectious diseases: part II. *Nature* 280, 455–461
- Anderson, R. and May, R. (1991) *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press
- Earn, D. *et al.* (1998) Persistence, chaos and synchrony in ecology and epidemiology. *Proc. R. Soc. B* 265, 7
- Ostfeld, R. *et al.* (2008) *Infectious Disease Ecology: Effects of Ecosystems on Disease and of Disease on Ecosystems*, Princeton University Press
- Cummings, D.A.T. *et al.* (2009) The impact of the demographic transition on dengue in Thailand: insights from a statistical analysis and mathematical modeling. *PLoS Med.* 6, e1000139

- 28 Viboud, C. *et al.* (2006) Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* 312, 447–451
- 29 King, A. *et al.* (2008) Inapparent infections and cholera dynamics. *Nature* 454, 877–880
- 30 Wearing, H. and Rohani, P. (2009) Estimating the duration of pertussis immunity using epidemiological signatures. *PLoS Pathog.* 5, e1000647
- 31 Tidd, C. *et al.* (1993) The case for chaos in childhood epidemics. I. Predicting historical epidemics from mathematical models. *Proc. R. Soc. Lond. B* 254, 257–273
- 32 Cazelles, B. *et al.* (2008) Wavelet analysis of ecological time series. *Oecologia* 156, 287–304
- 33 Taper, M.L. and Lele, S.R. (2004) *The Nature of Scientific Evidence: Statistical, Philosophical, and Empirical Considerations*, University of Chicago Press
- 34 Fussmann, G. *et al.* (2000) Crossing the Hopf bifurcation in a live predator–prey system. *Science* 290, 1358–1360
- 35 McCauley, E. *et al.* (2008) Small-amplitude cycles emerge from stage-structured interactions in Daphnia–algal systems. *Nature* 455, 1240–1243
- 36 Costantino, R.F. *et al.* (2005) Nonlinear stochastic population dynamics: the flour beetle *Tribolium* as an effective tool of discovery. In *Population Dynamics and Laboratory Ecology (Advances in Ecological Research, Vol. 37)* (Desharnais, R.A., ed.), pp. 101–141, Academic Press
- 37 Grenfell, B.T. *et al.* (2002) Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model. *Ecol. Monogr.* 72, 169–184
- 38 London, W. and Yorke, J.A. (1973) Recurrent outbreaks of measles, chickenpox and mumps. I. Seasonal variations in contact rates. *Am. J. Epidemiol.* 98, 453–468
- 39 Finkenstädt, B. and Grenfell, B. (2000) Time series modelling of childhood diseases: a dynamical systems approach. *Appl. Stat.* 49, 187–205
- 40 Ferrari, M. *et al.* (2008) The dynamics of measles in sub-Saharan Africa. *Nature* 451, 679–684
- 41 Bolker, B. and Grenfell, B. (1993) Chaos and biological complexity in measles dynamics. *Proc. R. Soc. B* 251, 75–81
- 42 Earn, D. *et al.* (2000) A simple model for complex dynamical transitions in epidemics. *Science* 287, 667–670
- 43 Metcalf, C. *et al.* (2009) Seasonality and comparative dynamics of six childhood infections in pre-vaccination Copenhagen. *Proc. R. Soc. Lond. B* 276, 4111–4118
- 44 Nicholson, A. (1957) The self-adjustment of populations to change. *Cold Spring Harbor Symp. Quant. Biol.* 22, 153–172
- 45 Andrewartha, H. and Birch, L. (1954) *The Distribution and Abundance of Animals*, Chicago University Press
- 46 Rohani, P. *et al.* (2002) The interplay between determinism and stochasticity in childhood diseases. *Am. Nat.* 159, 469–481
- 47 Bauch, C.T. and Earn, D.J.D. (2003) Transients and attractors in epidemics. *Proc. Biol. Sci.* 270, 1573–1578
- 48 Alonso, D. *et al.* (2007) Stochastic amplification in epidemics. *J. R. Soc. Interface* 4, 575–582
- 49 Kuske, R. *et al.* (2007) Sustained oscillations via coherence resonance in SIR. *J. Theor. Biol.* 245, 459–469
- 50 Black, A. *et al.* (2009) Stochastic fluctuations in the susceptible–infective–recovered model with distributed infectious periods. *Phys. Rev. E* 80, 021922
- 51 Hastings, A. *et al.* (1993) Chaos in ecology: is Mother Nature a strange attractor? *Annu. Rev. Ecol. Syst.* 24, 1–33
- 52 Schaffer, W. and Kot, M. (1986) Chaos in ecological systems: the coals that Newcastle forgot. *Trends Ecol. Evol.* 1, 58–63
- 53 Olsen, L. and Schaffer, W. (1990) Chaos versus noisy periodicity: alternative hypotheses for childhood epidemics. *Science* 249, 499
- 54 Sugihara, G. and May, R. (1990) Nonlinear forecasting as a way of distinguishing chaos from measurement error in time-series. *Nature* 344, 734–741
- 55 Ellner, S. and Turchin, P. (1995) Chaos in a noisy world: new methods and evidence from time-series analysis. *Am. Nat.* 145, 343–375
- 56 McLean, A. and Anderson, R. (1988) Measles in developing countries part I. Epidemiological parameters and patterns. *Epidemiol. Infect.* 100, 111–133
- 57 Pitzer, V.E. *et al.* (2009) Demographic variability, vaccination, and the spatiotemporal dynamics of rotavirus epidemics. *Science* 325, 290–294
- 58 Rohani, P. *et al.* (1999) Opposite patterns of synchrony in sympatric disease metapopulations. *Science* 286, 968–971
- 59 King, A.A. *et al.* (2009) Evolution of acute infections and the invasion–persistence trade-off. *Am. Nat.* 173, 446–455
- 60 Black, F. (1966) Measles endemicity in insular populations: critical community size and its evolutionary implication. *J. Theor. Biol.* 11, 207
- 61 Broutin, H. *et al.* (2005) Epidemiological impact of vaccination on the dynamics of two childhood diseases in rural Senegal. *Microbes Infect.* 7, 593–599
- 62 Grenfell, B. *et al.* (2001) Travelling waves and spatial hierarchies in measles epidemics. *Nature* 414, 716–723
- 63 Brown, J. and Kodric-Brown, A. (1977) Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology* 58, 445–449
- 64 Zadokis, J. and Van den Bosch, F. (1994) On the spread of plant disease – a theory on foci. *Annu. Rev. Phytopathol.* 32, 503–521
- 65 Xia, Y.C. *et al.* (2004) Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics. *Am. Nat.* 164, 267–281
- 66 Diekmann, O. and Heesterbeek, J. (2000) *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*, Wiley
- 67 Koelle, K. and Pascual, M. (2004) Disentangling extrinsic from intrinsic factors in disease dynamics: a nonlinear time series approach with an application to cholera. *Am. Nat.* 163, 901–913
- 68 Cauchemez, S. *et al.* (2004) A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data. *Stat. Med.* 23, 3469–3487
- 69 Gibson, G.J. *et al.* (2004) Bayesian analysis of botanical epidemics using stochastic compartmental models. *Proc. Natl. Acad. Sci. U. S. A.* 101, 12120–12124
- 70 Kendall, B. *et al.* (2005) Population cycles in the pine looper moth: dynamical tests of mechanistic hypotheses. *Ecol. Monogr.* 75, 259–276
- 71 Ionides, E. *et al.* (2006) Inference for nonlinear dynamical systems. *Proc. Natl. Acad. Sci. U. S. A.* 103, 18438
- 72 He, D. *et al.* (2010) Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *J. R. Soc. Interface* 7, 271–283
- 73 Fenton, A. and Pedersen, A. (2005) Community epidemiology framework for classifying disease threats. *Emerg. Infect. Dis.* 11, 1815–1821
- 74 Ferguson, N. *et al.* (2003) Ecological and immunological determinants of influenza evolution. *Nature* 422, 428–433
- 75 Gog, J. and Grenfell, B. (2002) Dynamics and selection of many-strain pathogens. *Proc. Natl. Acad. Sci. U. S. A.* 99, 17209–17214
- 76 Wearing, H. and Rohani, P. (2006) Ecological and immunological determinants of dengue epidemics. *Proc. Natl. Acad. Sci. U. S. A.* 103, 11802–11807
- 77 Adams, B. *et al.* (2006) Cross-protective immunity can account for the alternating epidemic pattern of dengue virus serotypes circulating in Bangkok. *Proc. Natl. Acad. Sci. U. S. A.* 103, 14234–14239
- 78 Gupta, S. *et al.* (1998) Chaos, persistence and evolution of strain structure in antigenically diverse infectious agents. *Science* 280, 912–915
- 79 Biek, R. *et al.* (2006) A virus reveals population structure and recent demographic history of its carnivore host. *Science* 311, 538–541
- 80 Grenfell, B. *et al.* (2004) Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* 303, 327–332
- 81 Brogden, K. and Guthmiller, J. (2002) *Polymicrobial Diseases*, ASM Press
- 82 Levins, R. (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bull. Entomol. Soc. Am.* 15, 237–240
- 83 Mideo, N. *et al.* (2008) Linking within- and between-host dynamics in the evolutionary epidemiology of infectious diseases. *Trends Ecol. Evol.* 23, 511–517
- 84 Antia, R. *et al.* (2008) The dynamics of acute malaria infections. I. Effect of the parasite’s red blood cell preference. *Proc. R. Soc. B* 275, 1449
- 85 Read, J. and Keeling, M. (2006) Disease evolution across a range of spatio-temporal scales. *Theor. Pop. Biol.* 70, 201–213
- 86 Breban, R. *et al.* (2009) Quantifying the treatment efficacy of reverse transcriptase inhibitors: new analyses of clinical data based on within-host modelling. *BMC Pub. Health* 9, S11

- 87 Handel, A. *et al.* (2007) Neuraminidase inhibitor resistance in influenza: assessing the danger of its generation and spread. *PLoS Comput. Biol.* 3, e240
- 88 Dixit, N. and Perelson, A. (2005) HIV dynamics with multiple infections of target cells. *Proc. Natl. Acad. Sci. U. S. A.* 102, 8198–8203
- 89 Handel, A. *et al.* (2010) Towards a quantitative understanding of the within-host dynamics of influenza A infections. *J. Roy. Soc. Interface* 7, 35–47
- 90 Auranen, K. *et al.* (2004) A hierarchical Bayesian model to predict the duration of immunity to Haemophilus influenzae type B. *Biometrics* 55, 1306–1313
- 91 Rogers, L. (1923) The world incidence of leprosy in relation to meteorological conditions and its bearings on the probable mode of transmission. *Trans. R. Soc. Trop. Med. Hyg.* 16, 440–464
- 92 MacDonald, G. (1957) *The Epidemiology and Control of Malaria*, Oxford University Press
- 93 Brown, J. *et al.* (2009) Avian influenza virus infectivity in water: viral responses to varying pH, salinity, and temperature conditions. *Vet. Microbiol.* 136, 20–26
- 94 Rohani, P. *et al.* (2009) Environmental transmission of low pathogenicity avian influenza viruses and its implications for pathogen invasion. *Proc. Natl. Acad. Sci. U. S. A.* 106, 10365–10369
- 95 Craig, M. *et al.* (1999) A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitol. Today* 15, 105–110
- 96 Brownstein, J.S. *et al.* (2005) Effect of climate change on Lyme disease risk in North America. *EcoHealth* 2, 38–46
- 97 Rogers, D.J. and Randolph, S.E. (2000) The global spread of malaria in a future, warmer world. *Science* 289, 1763–1766
- 98 Keeling, M. *et al.* (2001) Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science* 294, 813
- 99 Ferguson, N. *et al.* (2001) Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. *Nature* 413, 542–547
- 100 Miller, E. *et al.* (2010) Incidence of 2009 pandemic influenza A H1N1 infection in England: a cross-sectional serological study. *Lancet* 375, 1100–1108
- 101 Eubank, S. *et al.* (2004) Modelling disease outbreaks in realistic urban social networks. *Nature* 429, 180–184
- 102 Mossong, J. *et al.* (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* 5, e74
- 103 González, M. *et al.* (2008) Understanding individual human mobility patterns. *Nature* 453, 779–782
- 104 Brockmann, D. *et al.* (2006) The scaling laws of human travel. *Nature* 439, 462–465