Multiscale, resurgent epidemics in a hierarchical metapopulation model

Duncan J. Watts*†‡§, Roby Muhamad*, Daniel C. Medina¶, and Peter S. Dodds†

*Department of Sociology, and †Institute for Social and Economic Research and Policy, Columbia University, New York, NY 10027; ‡Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501; and ¶College of Physicians and Surgeons, Columbia University, New York, NY 10032

Edited by David O. Siegmund, Stanford University, Stanford, CA, and approved June 14, 2005 (received for review February 12, 2005)

Although population structure has long been recognized as relevant to the spread of infectious disease, traditional mathematical models have understated the role of nonhomogenous mixing in populations with geographical and social structure. Recently, a wide variety of spatial and network models have been proposed that incorporate various aspects of interaction structure among individuals. However, these more complex models necessarily suffer from limited tractability, rendering general conclusions difficult to draw. In seeking a compromise between parsimony and realism, we introduce a class of metapopulation models in which we assume homogeneous mixing holds within local contexts, and that these contexts are embedded in a nested hierarchy of successively larger domains. We model the movement of individuals between contexts via simple transport parameters and allow diseases to spread stochastically. Our model exhibits some important stylized features of real epidemics, including extreme size variation and temporal heterogeneity, that are difficult to characterize with traditional measures. In particular, our results suggest that when epidemics do occur the basic reproduction number R_0 may bear little relation to their final size. Informed by our model's behavior, we suggest measures for characterizing epidemic thresholds and discuss implications for the control of epidemics.

math model | population structure

The role and importance of interaction structure is a central yet unresolved issue in mathematical epidemiology (1). At the broadest level, the issue is straightforward: clearly not all people interact equally with all others; hence diseases of humans cannot spread in real populations precisely as they would if all individuals were to mix uniformly at random. Moving beyond this simple insight, however, poses considerable empirical and theoretical obstacles: empirical, because the amount and variety of structure present in real populations of different sizes defies existing measurement technologies; and theoretical, because without such knowledge it is difficult to model and thus assess the impact of interaction structure on the spread of human-to-human diseases. In this article, we focus on two key aspects of large populations that we believe have not received adequate attention in the existing literature: (*i*) that large populations exhibit structure at many scales; and (*ii*) that the movement of individuals between these scales is essential to the spread of a large epidemic. These features can be represented formally with a straightforward variation of a commonly studied class of disease-spreading models, metapopulation models (e.g., ref. 2), yet they nevertheless carry important implications for understanding and possibly controlling diseases, such as severe acute respiratory syndrome (SARS) and influenza, that have the potential to spread on many scales.

Metapopulation models can in general be characterized as a theoretical compromise between the simplest and most analytically tractable disease-spreading models, often called compartment models, and models in the recent network epidemiology tradition that attempt to capture population structure in a realistic way, but which necessarily exhibit far greater complexity. Compartment models assume a continuous population that is divided into a number of compartments (or states), typically susceptible, infected, and recovered. Disease transmission occurs because of contact between susceptible and infected individuals, and the mixing within and between compartments is assumed to be random, where transition rules (for example, the rate at which an infected person recovers) specify how individuals move from one compartment to another (3).

Population structure can be introduced into these simple models by specifying additional compartments, corresponding not only to the different stages of within-host behavior, but also to various differentiating features of the population, such as age (4), susceptibility (5), risk behavior (6), and social status (2, 7), along with a correspondingly complex set of mixing rates. Individual-level fluctuations can also be included by specifying fully stochastic versions of these models (8) without overly compromising their mathematical tractability. Nevertheless, compartment models rely heavily on the assumption that population structure can be represented solely in terms of individual attributes (e.g., disease state, age, behavior), an assumption that clearly cannot be satisfied in cases of diseases spreading over spatially extended regions, where the physical distribution of the population matters, or when disease transmission depends on specific types of interactions (such as for sexually transmitted diseases), whose structure may cut across physical locations and social categories in unknown and complicated ways.

Spatial models (4, 9–11) address part of this problem by modeling transmission as a function of geographical distance and have been effective in capturing the dynamics of diseases in wild (12) and domesticated (13) animals, as well as in suggesting control strategies. However, spatial models are less relevant to epidemics of modern human societies, in part because of the importance of modern modes of transportation that shortcut long geographical distances (14–16), and in part because many diseases are transmitted by close-contact networks that characterize families, organizational affiliations (e.g., school or work) (7), or sexual relations (17). In recent years, therefore, models that attempt to characterize the actual pattern of interactions associated with a particular population and disease transmission mechanism have become increasingly popular (17–21). However, although network models are appealing from a theoretical perspective, the more elements of interaction structure that any such model incorporates, the more free parameters and assumptions are required, and the harder it becomes to perform robust and reliable analyses (1). Exacerbating this problem of model complexity is the difficulty of determining parameters or justifying assumptions empirically.

Metapopulation models (2) therefore offer a potentially useful compromise between compartment models and networks. Like compartment models, metapopulation models assume random mixing within subpopulations (or patches) that are typically

This paper was submitted directly (Track II) to the PNAS office.

Abbreviation: SARS, severe acute respiratory syndrome.

§To whom correspondence should be addressed. E-mail: djw24@columbia.edu.

defined in terms of geographical regions such as cities (14–16), districts within a city (22), villages (2), or even homes, schools, and shopping malls (7, 23). However, they also incorporate the central insight from spatial and network models that interactions between subgroups depend not just on individual similarities and differences, but on actual transportation networks (15, 16) and routine behavioral patterns such as going to work, attending school, etc. (7, 23). The model that we study here follows in the tradition of metapopulation models, most closely that of refs. 15, 16, and 24–26. However, where these models are restricted to mixing on only two scales, which we might call "local" and "global," our model explicitly incorporates mixing at multiple scales, a feature that has been identified as important by a number of authors (9, 27) but not yet examined in formal models. As we show below, the introduction of mixing on multiple scales does indeed have important consequences for the properties of epidemics.

Hierarchical Metapopulation Model

To specify our model precisely, we make the following assumptions.

(*i*) We assume that individuals occupy one of three states (susceptible, infected, and recovered) where the fraction of the population in each state at time t is $S(t)$, $I(t)$, and $R(t)$, respectively. At each discrete time step $t = 0, 1, 2, \ldots$, each individual is brought into contact with one other individual (in a manner specified below), and if any such pair comprises an infective and a susceptible, then the susceptible becomes infected with probability β, the infectiousness of the disease. Infectives subsequently recover with probability γ in each time step, after which they remain recovered permanently. Thus we have a stochastic susceptible-infected-recovered (SIR) model, from which we can immediately compute the basic reproduction number $R_0 = \beta/\gamma$, defined as the expected number of secondary infections generated by a single infective in an otherwise wholly susceptible population (28).

(*ii*) We assume that in sufficiently localized contexts, such as hospitals, schools, and workplaces, uniform mixing is approximately satisfied. Specifically, we partition the total population of N individuals into subpopulations of equal size n, within which the standard compartment model can be applied. Furthermore, we assume that all disease transmission occurs between infectives and susceptibles who occupy the same local context. It follows trivially that $R_0 > 1$ is a necessary condition for epidemics to start, although we note that because our model is stochastic, the condition $R_0 > 1$ does not guarantee that an epidemic will occur, only that it will do so with nonzero probability. Once initiated, an epidemic will proceed within its initial subpopulation in a manner analogous to a standard stochastic compartment model. How it may spread beyond any particular subpopulation depends on the large-scale structure of the population as a whole, as specified next.

(*iii*) We assume that at any given point in time each individual can be assigned uniquely to a single local context within which they experience sufficiently sustained, close interaction for disease transmission. However, we also want to capture the notion that these same individuals simultaneously belong to a nested hierarchy of larger communities, such as neighborhoods, cities, regions, states, nations, and so on, each of which is successively larger and more diffuse. These communities, therefore, cannot be represented as homogenously mixing populations; only when two individuals are present in the same lowest-level group are they able to engage in person-to-person disease transmission. However, their likelihood of being in the same lowest-level group depends on which level group (in the nested hierarchy) they have recently shared. Individuals from the same neighborhood, for example, or even a large city, are much more likely to be in the same store or school than individuals from different countries.

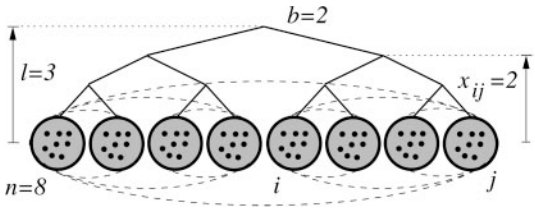

Fig. 1. Hierarchical metapopulation model of disease spreading. In our model, individuals (dots) belong to groups (solid circles) that in turn belong to groups of groups, and so on, giving rise to a hierarchy of scales. Alternatively, one can imagine the hierarchy as a nested set of subpopulations of increasing size, as indicated schematically by the dashed boundaries. In this example, there are $n = 8$ individuals in each group and the hierarchy has $l = 3$ levels and a branching ratio of $b = 2$. Individuals in the same group are considered to be a distance $x = 0$ apart. As an example, the distance between individuals in groups i and j is $x_{ij} = 2$.

We model this intuitive idea as follows. At time $t = 0$, every individual is assigned uniquely to a single local context, and one individual is infected. Subsequently, in addition to the infection dynamics specified above, individuals are permitted to travel between contexts in the following manner. With probability p, at each time step, each individual leaves his local context i and enters a new context j with probability $q_{ij} \propto e^{-x_{ij}/\xi}$, where x_{ij} is the ultra-metric distance between contexts i and j and ξ is a tunable parameter. Local contexts are related to each other via a nested hierarchy of scales, as represented schematically in Fig. 1, where b is the branching ratio of the hierarchy, and l is its total depth. We emphasize that the hierarchy of scales specified here is quite different from that in recent network models (29–32) in that the present model permits interactions between individuals (and hence infections) to occur only within small, locally mixing contexts. Individuals may travel long distances, and this feature is also important, but unlike in the network models above, they cannot simultaneously infect others who are not themselves nearby.

Although the extension from two to multiple scales is modest from a technical perspective, it nevertheless enables us to shed light on some important properties of real epidemics that are difficult to capture with models that admit mixing at only one or two scales. For example, as much as they vary in certain respects, almost all models of disease spreading share the property that any outbreak of disease can suffer only one of two fates: either the relevant epidemic threshold condition is not satisfied and the disease burns itself out before infecting more than a local population; or the condition is satisfied, in which case, with a nonzero probability, it spreads globally to a scale proportional to the population size N (1). The resulting distribution of event sizes is therefore always bimodal, where one mode corresponds to local outbreaks, and the other to "successful" epidemics, a result that holds even for relatively complicated network models in which individuals may interact along distinct "social dimensions" (30, 32).

Real epidemics, by contrast, occur on many scales. Our particular interest here is in novel epidemics that have the potential to spread on a global scale, such as SARS and influenza. Unfortunately, although the claim that historical size variations of, say, influenza epidemics, are extremely large is plausible [the 1918 epidemic is estimated to have infected 20 million to 40 million people (33)] sufficiently comprehensive data are not available to test it. However, detailed data regarding epidemics of various childhood diseases have been collected for more than a century in Iceland (1888–1990). Although these data are not ideal for our purposes [they are confined to a relatively small population (≈300,000) and childhood diseases differ in some important respects (discussed below) from the

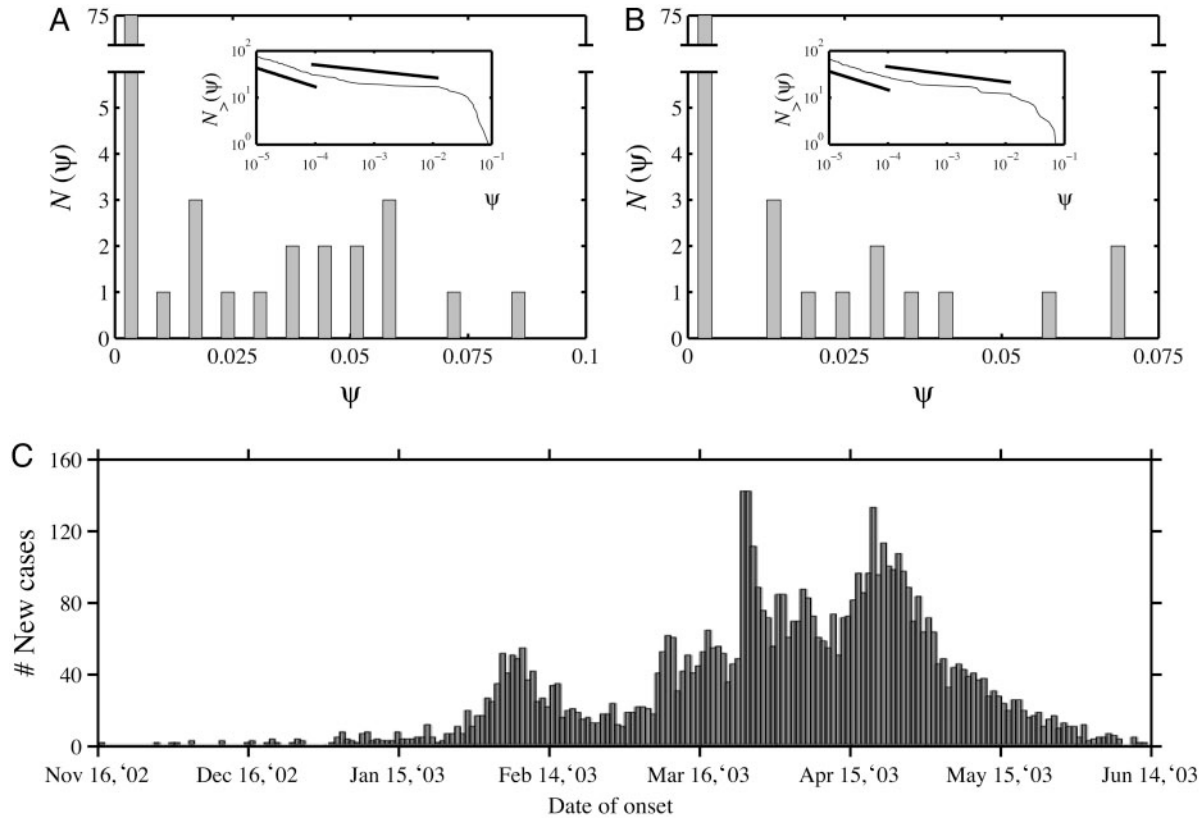

Fig. 2. Multiscale, resurgent behavior of real epidemics. (*A* and *B*) Frequency of epidemic sizes for measles (*A*) and pertussis (*B*) for Iceland, 1888–1990 (9). Each epidemic is identified as a contiguous sequence of months with nonzero caseloads bounded by months with zero caseloads, and the corresponding epidemic size ψ is normalized by the population size of Iceland at the midpoint of the epidemic. Both examples show a single mode for small outbreak sizes along with a relatively flat distribution of larger epidemic sizes. (*Insets*) The complementary cumulative frequency distributions (i.e., the number of epidemics of size at least ψ) for the same diseases on double logarithmic axes. A linear fit with slope $-\theta$ corresponds to a power-law distribution $P[\Psi > \psi] \propto \psi^{-\theta}$ over some interval $\psi_{\min} \leq \psi \leq \psi_{\max}$, usually called the ''scaling region.'' Here, the largest, and most likely, candidate for a scaling region is $10^{-4} \leq \psi \leq 10^{-2}$, over which the exponent θ is 0.13 for measles and 0.16 for pertussis; both values are well outside the usual range of $1 \leq \theta < 2$. We also identify a second, smaller scaling region, $10^{-5} \leq \psi \leq 10^{-4}$, in which θ is 0.40 for measles and 0.39 for pertussis, where the slightly larger exponents can be attributed to the spike near zero. Our conclusion is that neither distribution is well fit by a power law (see text). (*C*) Estimated worldwide number of cases of SARS as a function of time, showing resurgent behavior. Source: Mark Lipsitch, Harvard University, Boston.

simple model we present here] they nevertheless serve to illustrate the qualitative point that epidemics, when they succeed, occur at many scales, as shown in Fig. 2 *A* and *B* for measles and pertussis (whooping cough).

An important aspect of childhood diseases like measles and pertussis is that their propagation depends in part on the long-term immunity granted to survivors of previous epidemics. Most models of childhood diseases therefore incorporate the effects of what is called ''herd immunity'' by allowing the susceptible population to vary over time, [sometimes with realistic age profiles (34)], as a function of births, deaths, and recoveries. Successive epidemics arise either as a result of periodic exogenous introductions (as is usually done in the case of small populations like Iceland) or because of endemic persistence of the disease (in larger populations) (3). Although our primary purpose is to understand the size distributions of novel epidemics like SARS rather than recurring childhood epidemics like measles, we have checked [using parameters approximating the population, birth rate, and rate of introduction of measles into Iceland in the 20th century (9)] that the effects of herd immunity are unlikely to account for the size distributions in Fig. 2 *A* and *B*. Our results (data not shown) indicate that although fluctuations in the relevant susceptible population over time do increase the heterogeneity of epidemic sizes around the upper mode, the size distribution remains essentially bimodal.

In contrast with the usual prediction of bimodal size distributions, recent work (35) has suggested that the sizes of epidemics in isolated communities, including Iceland, are ''power law'' distributed ($P[\Psi > \psi] \propto \psi^{-\theta}$), and that such epidemics might therefore be the result of self-organized criticality (36). We find, however, that the size distributions of epidemics in Iceland are not well described by power laws. As shown in Fig. 2 *A* and *B Insets*, fitting a power law to the complementary cumulative distributions reveals exponents θ close to zero, well outside the typical range ($1 \leq \theta < 2$) either observed for empirical power-law distributions or generated by familiar mechanisms (36, 37). Furthermore, the forest-fire model used in ref. 35 remains poorly understood (38) especially when generalized beyond a 2D lattice. Thus although it does appear that epidemics occur on many scales, this observation neither implies that epidemic sizes conform to a power-law distribution nor that self-organized criticality is present.

Another feature of at least some real-world epidemics that has received relatively little attention in the modeling literature is their striking temporal heterogeneity, as exemplified by the daily caseload data of the 2003 SARS epidemic, shown in Fig. 2*C*. Although periodic and chaotic time series of epidemics have been extensively investigated in the context of epidemic models that incorporate exogenous (i.e., nondisease related) birth and death processes (3, 39), such ''recurrence'' takes place on time

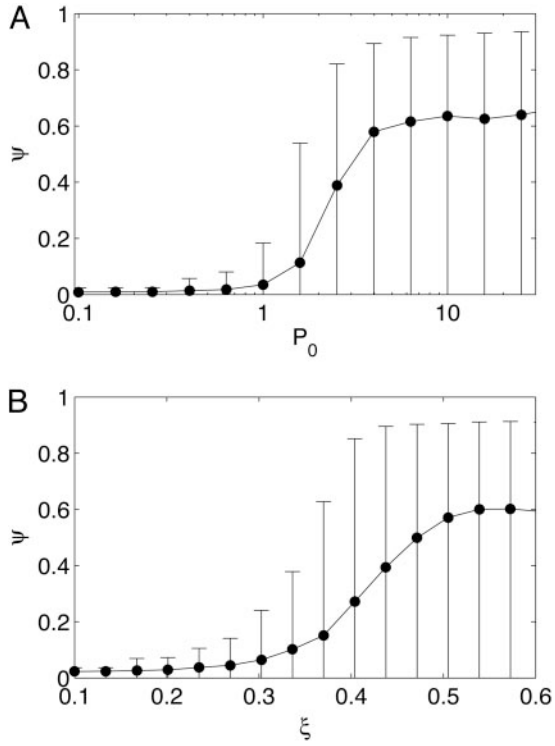

Fig. 3. Average epidemic size ψ as a function of model parameters. (*A*) ψ and associated 95% confidence interval delimited by the 2.5th and 97.5th percentiles (vertical bars) as a function of P_0 as approximated by $P_0 \approx p\psi' n/\gamma$. This approximation, which becomes exact as $p \to 0$, can be understood as follows. The quantity ψ' is the expected normalized epidemic size for a stochastic, homogenous mixing model of size n and reproduction number R_0. Because infectives typically recover after $1/\gamma$ time steps, there are roughly $\psi' n/\gamma$ opportunities for infected individuals to leave the initially infected group. Multiplying by p then gives the expected number who do. As discussed in the text, when $P_0 > 1$, an epidemic is expected to leave the initial context and spread nonlocally (providing $R_0 > 1$). The model parameters for the data shown here are basic reproduction number $R_0 = 3$, transport parameter $\xi = 0.6$, branching ratio $b = 3$, hierarchy depth $l = 4$, and group size $n = 100$. The results are averaged over 1,000 simulations. (*B*) The average epidemic size ψ (again with 95% confidence intervals) as a function of ξ (all parameters are the same but now P_0 is fixed at 5). For sufficiently low ξ, nonlocal epidemics never arise, regardless of whether or not P_0 and R_0 exceed unity. When $\xi = 1/\ln b$, individuals are equally likely to travel to all distances; the disease typically spreads globally; and the epidemic size distribution is bimodal. For both plots, intermediate values of P_0 and ξ yield epidemics of a wide range of sizes (see Fig. 4).

scales longer than that of a single epidemic. In all of the models surveyed above, whether compartment, spatial, network, or metapopulation models, single epidemics, when they occur, invariably exhibit a single peak in the number of infectives, followed by a "burnout" period. Fig. 2*C*, by contrast, displays severe downturns followed by dramatic surges, as the epidemic "discovers" new pools of susceptibles, but without any change in the underlying population itself, a phenomenon that has been labeled resurgence (40, 41).

Results and Discussion

Although our model is simplistic, it nevertheless leads to insights regarding the potential for novel epidemics to display multiscale and resurgent dynamics. Fig. 3*A* shows the average epidemic size ψ (defined as the cumulative fraction of a population infected during an epidemic), along with 95% confidence intervals (vertical bars) as a function of P_0, for fixed ξ. The quantity P_0 is the expected number of infected individuals leaving a single context over the mean infectious period (see Fig. 3 for details). For the example of Fig. 3*A*, $\beta = 0.3$ and $\gamma = 0.1$ have been chosen such that $R_0 = 3$; thus it is necessarily the case (because $R_0 > 1$) that with nonzero probability epidemics of some size occur, regardless of the value of P_0. However, it is clear from Fig. 3*A* that the final size of an epidemic exhibits a dramatic transition from local to nonlocal in the vicinity of $P_0 = 1$. The reason is that when $P_0 < 1$, infected individuals are likely to be confined to whichever context they are assigned at $t = 0$; hence the size of any resulting epidemics is bounded by the local scale n. Because $n \ll N$ and, moreover, n is determined by factors, such as the transmission mechanism of the disease and population density, that do not scale with the entire size of the global population, then for a sufficiently large N, epidemics that occur when $P_0 < 1$ will not reach more than a negligible fraction of the total population. When $P_0 > 1$, however, epidemics of different scales can occur, including truly global epidemics in the sense that a nonzero fraction of the population becomes infected, regardless of the population size.

When nonlocal epidemics are possible (i.e., when $R_0 > 1$ and $P_0 > 1$), their size depends sensitively on ξ, our second transport parameter, which characterizes the range over which individuals can travel. Fig. 3*B* shows ψ as a function of ξ for fixed P_0 and R_0. In both the limits $\xi \to 0$ and $\xi \to 1/\ln b$, our model reduces to an approximation of a stochastic, homogenously mixed model, where the effective susceptible population corresponds to n and N, respectively, and where the stochastic fluctuations correspondingly occur at the level of individuals or groups. However, neither of these limits is a plausible representation of modern-day transport, which is neither entirely local $\xi \to 0$, nor uniform at all scales $\xi \to 1/\ln b$. We thus focus our attention on the intermediate range of ξ where changes to its value impact the expected epidemic size. In this range, although the shape of the curve looks similar to that for P_0 (Fig. 3*A*), note that P_0 is plotted on semilog scale, whereas ξ is plotted on a linear scale, meaning that small changes in ξ have a much greater impact on the size of resulting epidemics than equivalent changes in P_0. This result has possibly useful policy implications, as it suggests that restricting how far, rather than how often, individuals travel during an epidemic (say, by issuing travel advisories) may be the most effective way of minimizing the epidemic's eventual impact.

Nonlocal epidemics in our hierarchical metapopulation model also display the properties of multimodality and resurgence discussed above. In Fig. 4, all model parameters are kept fixed (with one exception, discussed below) and all epidemics are allowed to run their course without any exogenous interventions. Fig. 4*A* displays the bimodal epidemic size distribution for a homogeneous mixing model (i.e., a single group with $N = 102{,}400$), whereas *B* and *C* in Fig. 4 show two example epidemic size distributions generated by our model when structure is introduced into the same size population ($b = 4$, $l = 5$, and $n = 100$, along with $P_0 = 5$ and $\xi = 0.35$). Epidemics of all sizes are observed with a single mode near $\psi = 0$ and an otherwise mostly flat distribution. In Fig. 4 *E–G* show sample trajectories of two simulated epidemics generated by our model for the same values of P_0 and ξ; Fig. 4*D* shows the same information for the corresponding homogeneously mixed model. Whereas the homogeneous case (Fig. 4*D*) shows a simple rise and fall in the number of new cases with no evidence (or possibility) of resurgence, *E–G* in Fig. 4 display two striking features: (*i*) despite all three epidemics having succeeded in breaking out from their initial group, the trajectories display very different shapes and durations; and (*ii*) they all display the resurgence property that is apparent in Fig. 2*C* (and also noted in refs. 40 and 41); that is, the epidemic dies down, only to flare back up again.

In addition to reproducing some stylized features of real-world epidemics, such as multimodal size distributions and temporal resurgence, our results have implications for the control of

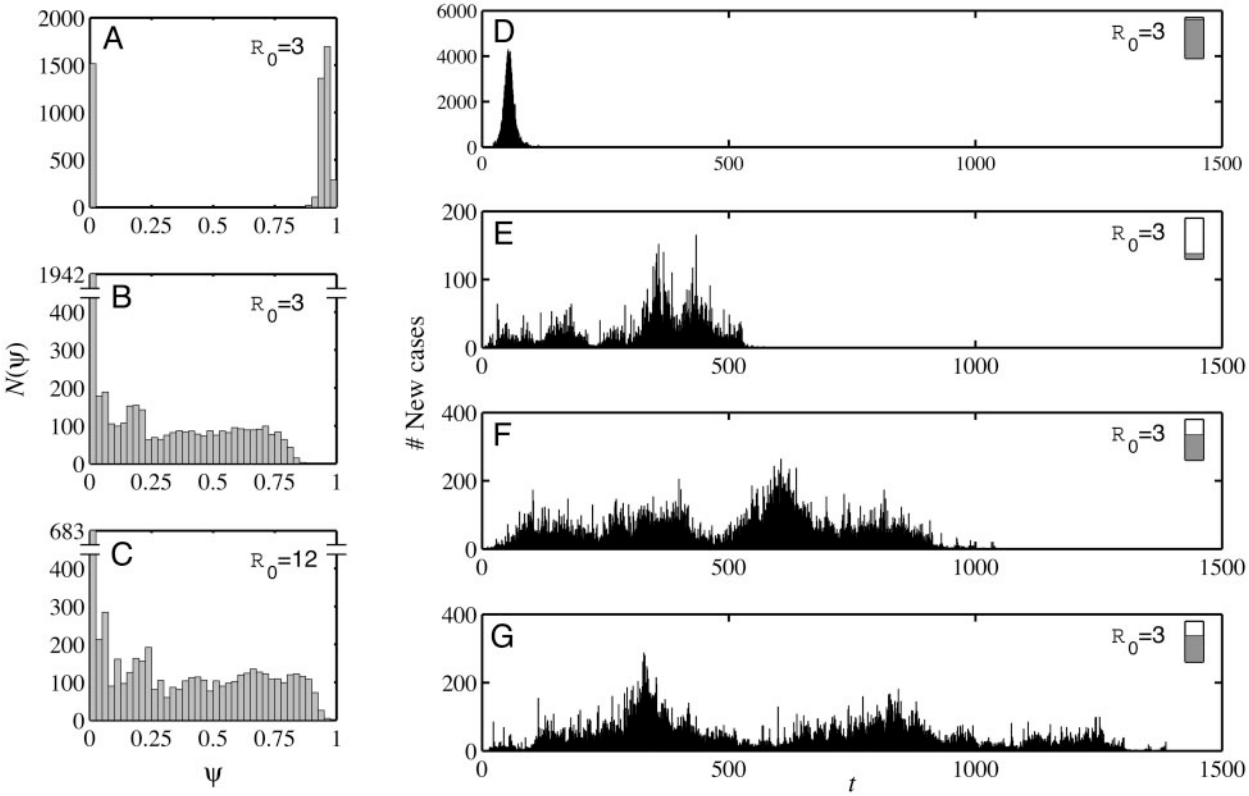

Fig. 4. Evidence of multiscale, resurgent behavior in simulated data. (*A–C*) Histograms of epidemic size ψ for three configurations of the model. (*A*) The bimodal size distribution generated from 5,000 simulations of a homogeneously mixing population of size $N = 102{,}400$ and reproduction number $R_0 = 3$. (*B* and *C*) The same size population as in *A* is hierarchically structured with branching ratio $b = 4$, depth $l = 5$, and group size $n = 100$, and for both of these examples, $P_0 = 0.35$ and $\xi = 0.35$. The models for *B* and *C* differ in the reproductive number: $R_0 = 3$ and 12, respectively. Nevertheless, both *B* and *C* show similar size distributions with modes near $\psi = 0$ and a relatively flat distribution for $\psi > 0$, qualitatively similar in form to that of the Icelandic data. (*D–G*) Example time series of total new cases, where in all cases, simulation parameters are $R_0 = 3$ and, again, $N = 102{,}400$. For *D*, the population is homogeneously mixing, and for *E–G*, the population is structured according to our model with the same parameters as for *B* and *C*. For the random mixing case of *D*, a typical epidemic trajectory rises rapidly once and then declines to zero, infecting most of the population in the process (see gray sidebar). In the examples with population structure, by contrast, epidemic trajectories exhibit dramatic resurgence, endure for markedly different time intervals, and infect very different fractions of the population (see gray sidebars), depending solely on stochastic fluctuations (i.e., all parameters are held constant).

epidemics. For example, an important variation between Fig. 4 *B* and *C* is that $R_0 = 3$ and $R_0 = 12$, respectively. Otherwise, *B* and *C* in Fig. 4 are remarkably similar: although Fig. 4*C* ($R_0 = 12$) displays greater mass in the tail, both exhibit a single peak near zero and are otherwise roughly flat distributions, extended across all scales. The shape and similarity of Fig. 4 *B* and *C* therefore suggest (*i*) that the same value of R_0 can be consistent with a very large range of epidemic sizes; and (*ii*) that very different values of R_0 can result in very similar epidemic size distributions. We note that this conclusion does not hold for all model parameters: As ξ increases, the size distribution becomes increasingly bimodal; and as ξ decreases, the distribution tends toward an exponential with a single mode. Thus as anticipated above, when transport is either entirely local or entirely global, our model reduces effectively to one with homogenous mixing, and R_0 can be interpreted in the usual manner. However, when transport is neither entirely local nor global, *B* and *C* in Fig. 4, together with our condition for a nonlocal epidemic $P_0 > 1$, imply that although $R_0 > 1$ remains a necessary condition for an epidemic to occur, the value of R_0 may otherwise provide little insight into the likely outcome of a disease outbreak.

Because much of the mathematical epidemiological literature has focused on R_0 as the principal parameter of interest (42), it is appropriate to elaborate somewhat on this last claim. For a deterministic metapopulation model, it would be straightforward to define R_0 in a way that accounts for the population structure (22, 43), by estimating the largest eigenvalue of the interpatch mixing matrix. In this manner, one could presumably obtain a modified reproduction number R_0 such that the condition $R_0 > 1$ would be equivalent to our dual condition $R_0 > 1$ and $P_0 > 1$ (see, for example, ref. 18). Simply redefining R_0, however, would not improve its relation to epidemic size in a stochastic, multiscale model. The reason is that the large variations in epidemic size apparent in Fig. 4 *B* and *C*, along with the resurgence behavior in Fig. 4 *E–G*, derive not from average statistics like R_0, but from rare, stochastic events, in which the epidemic "escapes" from currently infected contexts into newly susceptible populations. Thus the final size of an epidemic is largely determined by the small number of infected individuals who, by traveling, introduce the disease to previously unaffected groups. Although stochasticity is well understood to be important at the outset of a potential epidemic (1, 8), we see here that it continues to be relevant throughout the entire progress of an epidemic, even when large numbers of individuals have been infected.

Aside from highlighting the importance of stochastic, rare events, the multiscale and resurgent properties of epidemics in our model suggest that population structure itself can act as a kind of control. It has recently been shown, for example, that the same disease (SARS) can display very different trajectories in different regions, even when the corresponding environmental conditions are thought to be similar (41). Furthermore it is well known that different epidemics with similar estimated R_0 values

can experience dramatically different fates. For example, both the 2003 SARS epidemic and the 1918 influenza pandemic are estimated to have values of R_0 close to 3 (22, 33, 44), yet the latter is thought to have infected tens of millions of individuals, compared with only thousands for the former. Typically these differences are attributed to variations in the intensity, timing, and consistency of control measures. Although not disputing the relevance of active control measures to epidemic size and duration, our results indicate that population structure alone can generate purely stochastic fluctuations in epidemic size that are as large as any historically observed variations. The magnitude of this effect suggests that transport-oriented control measures such as issuing travel advisories, by effectively reducing ξ, could have a dramatic impact on the resulting epidemic size, possibly as large as more traditional, and far more interventionist, strategies like vaccination and quarantine.

Conclusion

Whenever a major outbreak of a novel infectious disease occurs, as happened most recently in 2003 with SARS, public health officials are invariably confronted with the question: how big will it be? As reasonable a question as this would seem, mathematical epidemiology currently provides no answer. To estimate the final size and duration of an epidemic, even sophisticated models of disease spreading require, as a parameter, the size (N) of the relevant susceptible population. After an epidemic has been observed it is, of course, always possible to estimate N, which is the usual modeling practice. Once we know, for example, that Hong Kong has suffered an epidemic of SARS, we may fit the parameters of a standard model such that approximately the correct (i.e., observed) number of individuals becomes infected in our simulations. Such an approach, however, is unable to shed light on the likelihood of very different scenarios having unfolded; thus its findings are of limited relevance to other outbreaks, past or future. For example, when a new respiratory disease with an R_0 and latency period comparable to SARS is identified in southern China, what outcome should we expect, and why? The population susceptible to a SARS-like virus is arguably no less than the entire population of the planet. Does it then make sense to estimate a worst-case scenario of hundreds of millions of potential victims? Alternatively, should we expect a similar outcome to the one experienced in 2003, with several hundred deaths spread across several continents, but concentrated in a few cities? Or was even that outcome atypically severe? In short, what is the likely size distribution for a given epidemic?

Although the model we have proposed in this article is too simplistic to answer these questions for realistic situations, it is at least formulated to address them. By representing population structure as a nested hierarchy of subpopulations, it has the advantage over existing metapopulation models in that it can accommodate populations of very large, and possibly even global, scales, without assuming that uniform mixing is satisfied at any scale above small, localized contexts. Our results can therefore be regarded as formalizing an early suggestion of Bailey's (45) that a global epidemic should properly be considered as many smaller epidemics occurring in different subpopulations, where (*i*) most transmission occurs at this local level, and (*ii*) broader spreading of a disease is driven by occasional long-range individual transport. Notwithstanding our model's simplicity, our results suggest an important empirical consequence: that the final size and duration of an epidemic are highly sensitive to the structure of the population through which it is spreading, even when the basic reproduction number R_0 is held constant. Conversely, similar distributions of epidemic size can correspond to very different values of R_0. Thus, in addition to the usual suite of intervention measures, control of epidemics could be exerted through effective manipulation of the natural barriers to disease spread that are inherent in the multiscale structure of large populations. We hope that future work will investigate increasingly realistic multiscale models of disease spreading, and explore their consequences for more effective control of novel epidemics.

1. Ferguson, N. M., Keeling, M. J., Edmunds, W. J., Gani, R., Grenfell, B. T., Anderson, R. M. & Leach, S. (2003) *Nature* **425,** 681–685.
2. Sattenspiel, L. & Dietz, K. (1995) *Math. Biosci.* **128,** 71–91.
3. Anderson, R. M. & May, R. M. (1991) *Infectious Diseases of Humans* (Oxford Univ. Press, Oxford).
4. Hethcote, H. W. (2000) *SIAM Rev.* **42,** 599–653.
5. Hethcote, H. W. & Yorke, J. A. (1984) *Gonorrhea Transmission Dynamics and Control* (Springer, New York).
6. Blower, S. M. & Farmer, P. (2003) *AIDScience*. Available at http://aidscience.org/Articles/AIDScience033.asp. Accessed July 13, 2005.
7. Halloran, M. E., Longini, I. M., Nizam, A. & Yang, Y. (2002) *Science* **298,** 1428–1432.
8. Keeling, M. J. & Grenfell, B. T. (2000) *J. Theor. Biol.* **203,** 51–61.
9. Cliff, A. D., Haggett, P., Ord, J. K. & Versey, G. R. (1981) *Spatial Diffusion: An Historical Geography of Epidemics in an Island Community*, Cambridge Geographical Studies (Cambridge Univ. Press, Cambridge, U.K.).
10. Durrett, R. & Levin, S. (1994) *Theor. Popul. Biol.* **46,** 363–394.
11. Holmes, E. E. (1997) in *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*, eds. Tilman, D. & Kareiva, P. (Princeton Univ. Press, Princeton), pp. 111–136.
12. Murray, J. D., Stanley, E. A. & Brown, D. L. (1986) *Proc. R. Soc. London Ser. B* **229,** 111–150.
13. Keeling, M. J., Woolhouse, M. E. J., May, R. M., Davies, G. & Grenfell, B. T. (2003) *Nature* **421,** 136–142.
14. Rvachev, L. A. & Longini, I. M. (1985) *Math. Biosci.* **75,** 3–22.
15. Longini, I. M. (1988) *Math. Biosci.* **90,** 367–383.
16. Hufnagel, L., Brockmann, D. & Geisel, T. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 15124–15129.
17. Klovdahl, A. S., Potterat, J. J., Woodhouse, D. E., Muth, J. B., Muth, S. Q. & Darrow, W. W. (1994) *Soc. Sci. Med.* **38,** 79–88.
18. Ball, F., Mollison, D. & Scalia-Tomba, G. (1997) *Ann. Appl. Probabil.* **7,** 46–89.
19. Kretzschmar, M. & Morris, M. (1996) *Math. Biosci.* **133,** 165–195.
20. Ancel, L. W., Newman, M. E. J., Martin, M. & Schrag, S. (2003) *Emerg. Infect. Dis.* **9,** 204–210.
21. Newman, M. E. J. & Watts, D. J. (1999) *Phys. Rev. E* **60,** 7332–7342.
22. Riley, S., Fraser, C., Donnelly, C. A., Ghani, A. C., Abu-Raddad, L. J., Hedley, A. J., Leung, G. M., Ho, L.-M., Lam, T.-H., Thach, T. Q., *et al*. (2003) *Science* **300,** 1961–1966.
23. Eubank, S., Guclu, H., Anil Kumar, V. S., Marathe, M. V., Srinivasan, A., Toroczkai, Z. & Wang, N. (2004) *Nature* **429,** 180–184.
24. Watson, R. K. (1972) *J. Appl. Probl.* **9,** 659–666.
25. Ball, F. & Clancy, D. (1993) *Adv. Appl. Prob.* **25,** 721–726.
26. Ball, F. (1991) *Math. Biosci.* **107,** 299–324.
27. Ferguson, N. M., May, R. M. & Anderson, R. M. (1997) in *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*, eds. Tilman, D. & Kareiva, P. (Princeton Univ. Press, Princeton), pp. 137–157.
28. Dietz, K. (1993) *Stat. Methods Med. Res.* **2,** 23—41.
29. Kleinberg, J. M. (2000) *Nature* **406,** 845.
30. Watts, D. J., Dodds, P. S. & Newman, E. J. (2002) *Science* **296,** 1302–1305.
31. Keeling, M. J. (1999) *Proc. R. Soc. London Ser. B* **266,** 859–867.
32. Zheng, D.-F., Hui, P. M., Trimper, S. & Zheng, B. (2005) *Physica A* **352,** 659–668.
33. Mills, C. E., Robins, J. M. & Lipsitch, M. (2004) *Nature* **432,** 904–906.
34. Bolker, B. M. (1993) *IMA J. Math. Appl. Med. Biol.* **10,** 83–95.
35. Rhodes, C. J. & Anderson, R. M. (1996) *Nature* **381,** 600–602.
36. Jensen, H. J. (1998) *Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems*, Cambridge Lecture Notes in Physics (Cambridge Univ. Press, Cambridge, U.K.).
37. Mitzenmacher, M. (2003) *Internet Math.* **1,** 226–251.
38. Grassberger, P. (2002) *New J. Phys.* **4,** 17.1—17.15.
39. Earn, D. J. D., Rohani, P., Bolker, B. M. & Grenfell, B. T. (2000) *Science* **287,** 667–670.
40. Dye, C. & Gay, N. (2003) *Science* **300,** 1884–1885.
41. Wallinga, J. & Teunis, P. (2004) *Am. J. Epidemiol.* **160,** 509–516.
42. Keeling, M. J. (2001) *+Plus Magazine*. Issue 14 (March). Available at http://plus.maths.org/issue14/features/diseases. Accessed July 13, 2005.
43. Diekmann, O. & Heesterbeek, J. A. P. (2000) *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis, and Interpretation* (Wiley, New York).
44. Lipsitch, M., Cohen, T., Cooper, B., Robins, J. M., Ma, S., James, L., Gopalakrishna, G., Chew, S. K., Tan, C. C., Samore, M. H., *et al*. (2003) *Science* **300,** 1966–1970.
45. Bailey, N. T. J. (1975) *The Mathematical Theory of Infectious Diseases and Its Applications* (Hafner, New York).