Disease transmission in heterogeneous populations

by

James Oliver Lloyd-Smith
B.Sc.(Hons.) (Queen's University) 1996
M.Sc. (University of British Columbia) 1999

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy
in
Biophysics
in the
GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Wayne M. Getz, Chair
Professor Adam P. Arkin
Professor Cheryl J. Briggs

Spring 2005

Disease transmission in heterogeneous populations

by

James Oliver Lloyd-Smith

Abstract

Disease transmission in heterogeneous populations

by

James Oliver Lloyd-Smith

Doctor of Philosophy in Biophysics

University of California, Berkeley

Professor Wayne M. Getz, Chair

Disease transmission between host individuals is the defining characteristic of infectious disease dynamics, and the transmission process is the core element of epidemic models. This dissertation describes advances in the modeling of disease transmission in populations with structural and dynamical heterogeneities, including behavioral changes due to illness, distinctions in contact patterns and control measures between hospitals and general communities, and individual-level variation in infectiousness arising from host, pathogen and environmental factors. Theoretical and empirical approaches are presented, motivated sometimes by fundamental questions about disease spread, and sometimes by applied problems related to specific outbreaks.

In Chapter 2, I present a mechanistic derivation of the frequency-dependent transmission model from a pair-based contact process, then extend this classical model to incorporate effects of illness on pairing behavior. Frequency-dependent transmission is the standard model for sexually-transmitted diseases (STDs), but a timescale approximation

required for the derivation means that pair-based STD transmission is portrayed accurately only for promiscuous populations and chronic, less-transmissible infections. Simulations define the limits of the classical model for two broad classes of STD. I then extend the derivation to include situations where infected individuals exhibit altered pairing behavior, presenting analytic expressions for the generalized frequency-dependent transmission rate, basic reproductive number (R_0), and steady-state prevalence of an epidemic, for four cases of increasing behavioral complexity. Potentially significant effects of infection-induced changes in contact behavior are illustrated by simulating epidemics of bacterial and viral STDs.

In Chapter 3, I investigate the stochastic invasion dynamics of an emerging disease in a community and its associated hospital, exploring for the first time the potential amplifying role of hospitals in an outbreak characterized by nosocomial spread. Severe acute respiratory syndrome (SARS) was transmitted extensively within hospitals, and healthcare workers comprised a large proportion of SARS cases worldwide. I evaluate contact precautions and case management (quarantine and isolation) as control measures for SARS, revealing that hospital infection control is the most potent measure and should be practiced by all individuals in affected hospitals, rather than only those interacting with known SARS cases. Delays of a few days in contact tracing and case identification severely degrade the utility of quarantine and isolation, and still more detrimental are delays between onset of an outbreak and implementation of control measures. If hospital-based transmission is not halted, measures which reduce community-healthcare worker contact are vital to preventing a widespread epidemic. These results have implications for future outbreaks of SARS or other emerging pathogens.

In Chapter 4, I address the impact of individual-level variation in infectiousness, and resulting superspreading events (SSEs), on disease emergence. I introduce the "individual reproductive number", a natural extension of the basic reproductive number R_0 from a population average to a distribution incorporating individual variation. The degree of individual

variation is quantified from outbreak data for ten diseases of casual contact (including SARS, smallpox, plague and H5N1 avian influenza), showing conclusively that conventional models are inadequate to represent real transmission patterns. I provide the first in-depth discussion and analysis of SSEs, including an extensive review of their causes and a general, probabilistic definition that allows prediction of the proportion of cases causing SSEs for a given disease outbreak. I introduce a rigorous stochastic theory for disease invasions based on the individual reproductive number, and analyze it to demonstrate profound effects of individual variation on disease emergence. The model replicates real-world patterns, such as the explosive outbreaks characteristic of SARS in 2003—a test failed by conventional models. Finally, I present theory and data on outbreak control that distinguish between population-wide and individual-specific control measures. Data from four outbreaks are more consistent with individual-specific control, which is proven to be more effective at disease containment for a given reduction in R_0. I show that targeting highly-infectious individuals substantially improves the efficiency of control, opening important research challenges in predicting individual infectiousness.

Table of Contents

ACKNOWLEDGEMENTS

I've been fortunate at Berkeley to work in a supportive and intellectually vibrant environment. Wayne Getz has taught me much about the practice and culture of science, and has set a powerful example by his integrity and humanism. Working with Wayne has been challenging, exciting and satisfying, and he has been incredibly generous with opportunities and his wholehearted support. Travis Porco has shared freely and graciously of his compendious knowledge of disease dynamics, and has shown me the beauty and pleasure in thinking a problem right through. Cherie Briggs gave me most of my training in ecology, as well as insightful guidance on matters of research, teaching and career. I thank Adam Arkin for our series of stimulating meetings, which I always left with my head brimming with new ideas. Joe Eisenberg provided excellent early training in disease modeling, and continued valuable advice over the years. I thank Udi Isacoff and the Berkeley Biophysics program for first drawing me to this great university, then granting me the freedom to discover and pursue a surprising new discipline. Shizuka Gannon, Diane Sigman and the Wellman office staff kept the bureaucrats at bay and otherwise softened the edges of institutional life. Away from Berkeley, Brian Williams has helped me to expand my research horizons to include global public health, and made possible a fascinating research visit at the World Health Organization. I've benefited greatly from collaborations with Sebastian Schreiber and Ekkehard Kopp, and separately with Alison Galvani.

My day-to-day life at Berkeley has been shaped and greatly enhanced by my brilliant and eclectic peers. I've learned a tremendous amount from (and shared

countless pleasurable coffees, lunches or beers with) the students and postdocs in our lab group: Shirli Bar-David, Peter Baxter, Annaliese Beery, Allison Bidlack, Paul Cross, John Eppley, Alison Galvani, Kimber Gross, Eran Karmon, Karen Levy, Andy Lyons, Leo Polansky, Phil Starks, Jessica Redfern, Sadie Ryan, Maria Sanchez, Wendy Turner, Michael Westphal, Chris Wilmers, and George Wittemyer. I particularly thank Paul Cross, my partner on a series of enjoyable and fruitful projects and tireless wrangler of unruly field data (and buffalo). My classmates in Biophysics have provided a loose-knit but strongly-bonded community throughout my time here, and given me a glimpse of their diverse fields of study. I sadly acknowledge the loss of Eran Karmon, whose unexpected departure robbed me of a valued friend, respected colleague, intrepid ski buddy and all-around comic relief. We're all poorer for his absence, and laugh less.

Everything I've done is built on the foundation of unwavering support provided by my family. I am deeply grateful for this consistent encouragement, and for the taste for challenge and exploration that was instilled at an early age. I'm endebted to family and friends—old and new, local and distant—for their grounding influence and reliable injection of fun into a busy lifestyle. I thank Marc and Julie for opening their home to a stray foreigner and being steadfast friends ever since, and the McNamara and Henderson families for their warm welcome and inclusiveness. Finally I thank Katy for sharing her life with me: for joy and adventure when times are good, for comfort and strength when they're bad, and for a broader, richer experience of the world through it all.

I gratefully acknowledge financial support during my dissertation from a Berkeley Fellowship, a post-graduate fellowship from the Natural Sciences &

Engineering Research Council of Canada, and support from the National Science Foundation and National Institutes of Health. Funds for travel and collaborative work were provided by the South African Centre for Epidemiological Modelling and Analysis, the American Institute for Mathematical Sciences, the James S. McDonnell Foundation, and the U.C. Berkeley Graduate Division.

Chapter One

Introduction

J.O. Lloyd-Smith

Epidemic modelers face an essential dilemma. On one hand, the social and biological phenomena we seek to model are far too complex to be encapsulated accurately by any reasonable set of equations or simulation rules. On the other hand, infectious diseases exact a relentless toll on their human and animal hosts, often raising compelling questions that cannot ethically or logistically be addressed by experiments. Modelers are thus engaged in an on-going struggle to improve our own science, moving from empirical observation of complex disease dynamics to elaboration of basic epidemiological theory, and back again, often motivated by the principal disease threats of the day.

This dissertation addresses a central challenge of disease ecology: modeling the transmission of infectious diseases in heterogeneous populations (Levin et al. 1997; Koopman 2002). Transmission among host individuals is the defining characteristic of infectious diseases, and the depiction of transmission is known to profoundly influence the results of epidemic models (Getz and Pickering 1983; Hethcote and Van Ark 1987; McCallum et al. 2001). Disease transmission is a complicated process, however, involving spatial and social relations among hosts, attributes of individual hosts, host-pathogen interactions such as immune response and evasion, and environmental factors (Anderson and May 1991; Levin et al. 1997; Koopman 2002). These factors can be divided into structural heterogeneities, which separate host individuals into groups with limited contact among them, and dynamical heterogeneities, which cause hosts to exhibit different disease-related behavior. Epidemic modelers must balance these known complexities of real populations with the need for clear and tractable models—

all the while aware that the broader conclusions of any disease modeling study may depend sensitively on the basic treatment of transmission.

Challenges related to transmission therefore demand the attention of all researchers of infectious disease dynamics. Fundamental research on epidemiological theory must work to improve our understanding of commonly-used transmission models, and extend them to include known heterogeneities that traditionally are ignored (e.g. Dietz and Hadeler 1988; Diekmann et al. 1990; Ball et al. 1997; Keeling 1999). Empirical research must seek to identify important patterns of heterogeneity and describe their effects on disease spread (e.g. Longini et al. 1982; Grenfell et al. 2001; Bjornstad et al. 2002; Gani and Leach 2004). Applied studies of particular outbreaks must structure their analyses to incorporate the dimensions of complexity that are essential to the questions at hand (e.g. Hethcote and Yorke 1984; Keeling et al. 2001; Smith et al. 2002; Keeling et al. 2003; Porco et al. 2004). In this dissertation on disease transmission in heterogeneous populations, I have endeavored to contribute on each of these three frontiers: theoretical, empirical, and applied.

In Chapter 2, I address the basic theory of modeling transmission in its most common manifestation, as the incidence term in compartmental (or SEIR-type) models of epidemics. For models without social or spatial structure, all details of the transmission process are encapsulated in the incidence term, which is a mathematical function of the numbers (or densities) of susceptible and infectious individuals that represents the rate at which new infections arise (Anderson and May 1991). The formulation of the incidence rate is the topic of on-going research and controversy (De Jong et al. 1995; McCallum et al. 2001; Begon et al. 2002). For diseases transmitted by

casual contact (DCC), common usage has shifted in the past decade from so-called "mass action" transmission to "frequency-dependent" transmission (McCallum et al. 2001); for sexually-transmitted diseases (STDs), frequency-dependent transmission has long been the preferred approach (Getz and Pickering 1983; Hethcote and Yorke 1984). In particular, I explore the mechanistic basis of the incidence term for STDs in particular, by modeling a pair-based contact process underlying transmission. (For STDs, sexual partnerships are the known mechanism of contact transmission, while for DCC pair-wise interactions are an approximate representation of daily social contact.) This approach introduces structural heterogeneity by dividing the population into single and paired individuals. Dynamical heterogeneity is added later in the chapter, by allowing infected individuals to exhibit different pairing behavior from healthy individuals—a phenomenon reported in empirical studies, but not previously incorporated into epidemiological theory.

In Chapter 2, I present the first mechanistic derivation of frequency-dependent incidence from the biologically-relevant contact process of pair formation and dissolution. This strengthens the case for modeling STD spread using the frequency-dependent model, but also clarifies the conditions required for the model to accurately represent pair-based transmission. The derivation requires a "timescale approximation" that pairing processes are much faster than disease processes, which is fulfilled only for relatively promiscuous populations, and is more accurate for chronic, less-transmissible infections than for transient, highly-transmissible infections. Frequency-dependent incidence thus can be thought to portray pair-based STD transmission only for a limited subset of STD-host combinations, and will overestimate incidence rates and steady-state

prevalence in many settings. For two major classes of STD, I use numerical simulations of a full pair-formation/epidemic model to define partnership durations beyond which the approximation breaks down and frequency-dependent incidence should not be used. For populations where the timescale approximation is reasonable, such as core groups and non-pair-bonding animals, I extend the mechanistic derivation to treat cases where the infection status of individuals influences their pairing behavior. I present expressions for the incidence rate, basic reproductive number (R_0), and steady-state prevalence (i_∞) of an epidemic, which can be applied directly in SEIR-type models where appropriate. The expression for R_0 is identical for all cases, giving refined insights into determinants of invasibility of STDs. Simulation results demonstrate the potentially dramatic impact of infection-induced behavioral change on epidemic dynamics.

Chapter 3 is an applied study, addressing the spread of severe acute respiratory syndrome (SARS)—or any infectious disease with a tendency for nosocomial spread—in a community and its affiliated hospital. Hospital-based transmission is a long-recognized phenomenon, but this is the first study to explicitly model the importation of patients from the community to the hospital, the resulting exposure of healthcare workers (HCWs), and the possibility for spread back to the community during HCWs' off-duty time. These patterns of spread were strikingly evident during the global emergence of SARS in 2003 (Anderson et al. 2004; Kamps and Hoffmann 2004), hence this work is an example of a real-world epidemic raising awareness of an applied problem, motivating its explicit study for the first time. In this model, the host population exhibits both structural heterogeneity (in the division between community,

HCW, and patient pools) and dynamical heterogeneity (in the different transmission pathways and control measures applicable to each pool). The study was conducted primarily by simulation, to allow incorporation of relevant biological detail and complex control policies in a stochastic framework, complemented by a simplified analytic model that allowed fundamental relationships to be elucidated.

The results of chapter 3 delineate priorities in outbreak containment in a hospital-community system, for a disease such as SARS for which no treatment or vaccine exists. The relative merits of case management (i.e. isolation and quarantine) and contact precautions (such as masks, gowns, and hand-washing) are weighed, including effects of both timing and efficacy of different control measures. A broad sensitivity analysis explores synergies and tradeoffs among measures imposed simultaneously, identifying the extent to which one measure can compensate for another which is not available in a given setting. A number of unintuitive and applicable conclusions arise from this analysis. For instance, quarantine is found to be largely unproductive for SARS, except via its indirect effect of hastening identification of symptomatic cases. Hospital-wide contact precautions as the single most potent containment measure—an encouraging finding since these are easily implemented and inexpensive. (It is noteworthy that subsequent work (in chapter 4) strongly reinforces this conclusion, but for a different reason: the greatest danger averted by hospital-wide infection control is transmission by misdiagnosed patients, rather than by unidentified HCW cases as in this model.) I investigate impacts of patient-level and population-level delays in control efforts, and emphasize the importance of immediate action by health authorities once an outbreak begins. In settings where hospital-based

transmission is continuing, analysis of pathways of transmission shows that measures reducing contact between HCWs and the community have dramatic benefits in preventing a widespread epidemic.

In chapter 4, I use empirical evidence for individual-level variation in infectiousness to motivate development of new theoretical results for disease invasions and provide new perspectives on the applied issue of outbreak control. Classical epidemiological theory uses population-average parameters to describe all individuals in a given group, neglecting the variation evident in epidemiologically-relevant traits from one individual to another (Anderson and May 1991; Diekmann and Heesterbeek 2000; Hethcote 2000). The most conspicuous example is the basic reproductive number, R_0, defined as the expected number of secondary cases caused by a "typical" individual in a wholly susceptible population, and certainly the central quantity in modern infectious disease modeling (Anderson and May 1991; Diekmann and Heesterbeek 2000). R_0 describes the average level of infectiousness for a given disease and population, and has very useful threshold properties, but does not account for differences in social contact rates, pathogen shedding, immunocompetence, hygiene, circumstance, or other factors that will influence the degree of transmission by a given individual. The effect of heterogeneity in contact rates has been considered for STDs and vector-borne diseases (May and Anderson 1987; Woolhouse et al. 1997), but for diseases of casual contact (DCC) individual-level heterogeneity has been largely ignored. Again, highly-publicized patterns in the 2003 SARS outbreaks—most notably the numerous superspreading events (SSEs) in which particular individuals generated anomalously high numbers of secondary cases (Anderson et al. 2004; Kamps and

Hoffmann 2004; Shen et al. 2004)—exposed the shortcomings of conventional theory, provoking the research described below. This work incorporates dynamical heterogeneity via differing infectiousness among individuals, but not structural heterogeneity because the population is assumed to be well-mixed.

Chapter 4 begins by introducing the "individual reproductive number", ν, the natural generalization of R_0 from a population average to a distribution incorporating individual variation in infectiousness. In contact tracing data this variation will be manifested in the "offspring distribution" defining the probability that an index case will cause a given number of secondary cases. Outbreak and surveillance data for SARS and nine other DCC exhibit significant levels of individual variation, shown by model selection techniques to be inconsistent with conventional theory assuming homogeneity. Quantifying this variation, I find that the most infectious 20% of cases is responsible for anywhere from 34 to 100% of transmission in these datasets (in contrast to a 20/80 rule proposed to be universal for STDs and vector-borne diseases (Woolhouse et al. 1997)). I propose a rigorous definition for SSEs, review their causes, and describe a method to predict their frequency based on assumptions about ν. Data on SSE frequency further support the existence of individual variation in infectiousness. To explore the impact of individual variation on disease emergence, I analyze stochastic branching process models incorporating a gamma-distributed individual reproductive number, which produces a negative binomial offspring distribution. High variation in infectiousness leads to either explosive outbreaks or rapid disease extinction: patterns consistent with outbreak data for SARS but unexplained by conventional theory. These insights are extended to a new theory for disease control that distinguishes individual-

specific from population-wide measures, and proves that stopping all transmission by some individuals controls epidemics better than partial reduction in everyone's transmission (for a given reduction in R_0). Returning to outbreak data, we see that changes in transmission patterns observed upon imposition of control are more consistent with individual-specific measures.

Through the interplay of developing theory, analyzing empirical data, and addressing applied problems, the work contained in this dissertation constitutes some measure of progress in our understanding of disease transmission in heterogeneous populations. Much more remains to be understood, including some direct links between the studies presented here. Community-hospital spread of SARS should be re-addressed in light of the marked individual variation in SARS transmission patterns and their demonstrated impact on invasion dynamics. Theory incorporating the individual reproductive number should be extended to populations with structural heterogeneity (i.e. multiple distinct groups), and generalized beyond the invasion scenarios described by branching process models. Individual-level heterogeneity in pairing behavior could be introduced to a pair-formation model, perhaps leading to situations where only part of the population satisfies the timescale approximation required for frequency-dependent incidence to represent pair-based transmission. In all cases, more and better empirical datasets are needed to test our theoretical ideas. The struggle against infectious diseases will not abate in the foreseeable future, so epidemic modelers must continue to advance our science in order to contribute meaningfully to the fight.

Chapter Two

Frequency-dependent incidence in sexually-transmitted disease models: portrayal of pair-based transmission and effects of illness on contact behaviour

J.O. Lloyd-Smith, W.M. Getz, and H.V. Westerhoff

1. Introduction

Sexually-transmitted diseases (STDs) are a significant and growing problem in public health worldwide (Piot et al. 2001; CDC 2002), and mathematical models have become an integral part of STD epidemiology (Hethcote and Yorke 1984; Castillo-Chavez 1989; Anderson and May 1991; Anderson and Garnett 2000; Garnett 2002). The core of every infectious disease model is its representation of the transmission process, and the fundamental question of how to formulate the transmission rate is the subject of active research (Antonovics et al. 1995; De Jong et al. 1995; McCallum et al. 2001; Fenton et al. 2002; McCallum et al. 2002). Increasingly complex models have yielded important insights into STD dynamics, by incorporating details of heterogeneity in sexual behavior (Hethcote and Yorke 1984; Anderson and May 1991), partnership dynamics (Dietz and Hadeler 1988; Waldstatter 1989; Diekmann et al. 1991; Kretzschmar and Dietz 1998; Kretzschmar 2000), and sexual network structure (Kretzschmar and Morris 1996; Bauch and Rand 2000; Ferguson and Garnett 2000; Kretzschmar 2000; Eames and Keeling 2002). These gains come at a price of increasing mathematical and computational demands, however, so simple analytic formulations continue to play an important role. STD models based on frequency-dependent transmission—the classical analytic formulation—are published often and prominently, typically nested in models addressing larger topics such as drug resistance or competition between strains (e.g. Thrall and Antonovics 1997; Blower et al. 1998; Blower et al. 2000; Bowden and Garnett 2000; Sullivan et al. 2001; Boots and Knell 2002). In their review of STD modeling, Anderson & Garnett (2000) point out that simpler models are complementary to complex simulations, and help to extract general

principles and reach robust conclusions which aid in policy development (though see Garnett et al. (1999) for a discussion of potential shortcomings). McCallum et al. (2002) observe that simple treatments of transmission will continue to be necessary, and that correspondence with more complex models is an important area for investigation.

In this spirit we undertake to explore the transmission process for sexually-transmitted diseases, and to relate complex dynamics to simple analytic expressions for the transmission rate. Using an approximation which applies to populations with rapid pairing dynamics (such as core groups or non-pair-bonding animals), we derive the classical frequency-dependent incidence mechanistically from a pair-formation model. We thus demonstrate a formal correspondence between the standard model of STD incidence and the pair-based contact process that is known to underlie STD dynamics. This mechanistic derivation clarifies the conditions required for the frequency-dependent model to accurately represent pair-based transmission, and provides a natural framework in which we assess the classical model's proper scope of application. We then extend the derivation to obtain generalized frequency-dependent transmission rates for situations where the pairing behavior of individuals is influenced by their infection status.

Frequency-dependent transmission (also called the standard incidence or density-independent transmission) is the standard approach to modeling STD transmission in compartmental disease models (Getz and Pickering 1983; Antonovics et al. 1995; De Jong et al. 1995; Hethcote 2000; McCallum et al. 2001). In this formulation, the rate at which susceptible individuals become infected is proportional to the prevalence (or "frequency") of the disease in the population. Let S and I represent

the densities of susceptible and infectious individuals, and $N=S+I$ the total density; the prevalence is thus $\frac{I}{N}$. Also, let c_{FD} represent the average per capita rate of acquiring new sex partners (assumed to be independent of population density), and p_{FD} represent the probability that transmission will occur over the course of an SI partnership (a partnership between a susceptible and an infectious individual). Then the frequency-dependent incidence rate, i.e. the rate at which new infections arise in the population, is (Anderson and May 1991; Hethcote 2000):

$$\varepsilon_{FD} = c_{FD} p_{FD} \left(\frac{I}{N} \right) S \tag{1}$$

The frequency-dependent formulation does not explicitly model sexual partnerships: contacts have no temporal extent and the probability of transmission during a partnership is ill-defined (Antonovics et al. 1995). Partnership dynamics have been recognized as a critical element of STD models since the seminal work of Dietz & Hadeler (1988), and are the basis for subsequent network models as well (Ferguson and Garnett 2000; Kretzschmar 2000). Pair-formation models account for the essential structure of STD epidemics, in that sexual contacts take place within partnerships (however short-lived) and individuals in monogamous pairs are removed from mixing with the rest of the population. Modeling the finite duration of partnerships can have significant effects for both transient and steady-state properties of STD epidemics (Waldstatter 1989; Kretzschmar and Dietz 1998), and resulting predictions for disease invasion or persistence can differ qualitatively from non-partnership models (Diekmann et al. 1991; Kretzschmar and Dietz 1998).

We seek to clarify the relationship of frequency-dependent incidence to the pairing processes underlying STD transmission. To do so, we construct a full model of

partnership, disease and demographic dynamics and consider its behavior in situations where pairing processes occur much faster than epidemic processes. In such situations we can separate the fast pair dynamics from the slower disease and demographic dynamics, and approximate that pair dynamics reach quasi-steady-state on epidemic timescales (i.e. that the pairing process constantly re-equilibrates as the epidemic progresses). We demonstrate numerically that solutions of the approximate system converge to those of the full system as the ratio of fast to slow timescales tends to infinity. This approach follows Heesterbeek & Metz (1993; see also Diekmann and Heesterbeek 2000), who broke new ground by deriving a saturating transmission function from a pair-formation model. The mass-action formulation of their pairing process, however, may be suitable for non-sexually transmitted infections but is inappropriate for STDs (particularly among humans) for which pair-entry rates are generally thought to be density-independent (Dietz and Hadeler 1988; Waldstatter 1989; Diekmann et al. 1991). In this study we derive analytic results from partnership models under the timescale approximation, then use comparisons with a full pair-formation/epidemic model to test the approximation. We find that the accuracy of the approximation—and hence the accuracy with which frequency-dependence portrays pair-based transmission—depends strongly on the natural history of the STD in question, as well as the rates of partnership formation and break-up.

Finally, for situations where the timescale approximation holds, we build on the foundation of pair-formation models to consider infection-induced changes in pairing behavior. Previous work on behavior change in STD epidemics has focused on population-level effects in humans, wherein education campaigns or community

awareness of disease prevalence lead to reduced risk behavior or changes in partner selection (e.g. Hadeler and Castillo-Chavez 1995; Hyman and Li 1997; Hsieh and Sheu 2001). We consider a different class of behavior change, applicable to humans and animals, in which individual contact behavior is influenced by individual infection status (i.e. sick individuals behave differently from healthy ones). There is strong evidence for this phenomenon in the biological and behavioral literature, but it has not yet been incorporated into epidemiological theory.

Infection status of individuals (both human and animal) has been shown to affect their contact behavior in sexually-transmitted and other diseases (Kennedy et al. 1987; Loehle 1995; Able 1996; Beckage 1997; Garnett et al. 1999; Kiesecker et al. 1999; Webster et al. 2003), and evolutionary biologists have long speculated that parasites play a role in mating behavior (Hamilton and Zuk 1982; Boots and Knell 2002). Proposed or observed mechanisms for lowered contact rates include debilitation and reduced vigor (e.g. Newshan et al. 1998; Schiltz and Sandfort 2000), social factors (Gold and Skinner 1996; Donovan 2000), scent cues (Penn and Potts 1998; Kavaliers et al. 1999), and secondary sexual characteristics (Hamilton and Zuk 1982; Loehle 1995; Able 1996). Pathogens can also act to modify host behavior to increase opportunity for transmission, as is commonly seen in macroparasitic infections (Beckage 1997); in sexual behavior, such effects have been observed in mice (Kavaliers et al. 1999) and milkwood leaf beetles (Abbot and Dill 2001), and postulated in humans (Starks et al. 2000). In sum, ample evidence exists that infection status of individuals may influence their partnering behavior. Below, we generalize the theory of frequency-dependent

transmission to incorporate this phenomenon for populations with rapid pairing dynamics.

2. Derivation of frequency-dependent transmission from a pair-formation model

We consider a model for STD spread in a population of individuals engaging in short-lived sexual partnerships, such as a core group within an HIV/AIDS epidemic. Long-term relationships are outside the scope of this study, but have been treated elsewhere (Diekmann et al. 1991; Kretzschmar et al. 1994). In the model presented here, the defining characteristics of a partnership are that (i) both individuals are removed from the mixing population for the duration of the pairing, and (ii) sexual contacts are occurring at some rate. We assume that sexual contacts do not occur outside of partnerships, and that concurrency is insignificant for these brief pairings. Hence disease transmission takes place only in partnerships between infected and susceptible individuals. We describe pairing dynamics using the standard formulation of recent STD models (e.g. Dietz and Hadeler 1988; Waldstatter 1989; Diekmann et al. 1991; Kretzschmar et al. 1994; Kretzschmar and Dietz 1998; Bauch and Rand 2000), assuming that individuals enter partnerships at a constant per capita rate and choose partners according to a defined mixing pattern.

The model is explained in Figure 1, and model equations are in the appendix. The basic assumption is that pairing processes occur on a timescale sufficiently faster than epidemic processes that the two systems can be separated, and pairing processes can be considered to be at quasi-steady-state relative to the epidemic. (Conversely, the epidemic variables are assumed to be constant at the pairing timescale.) This is the

approach taken by Heesterbeek & Metz (1993, also Diekmann and Heesterbeek 2000), who use a singular perturbation method to separate timescales which differ significantly. It is a quasi-steady-state because the partnership equilibria will be responding continuously as the epidemic progresses.

When we consider the fast pairing dynamics, we divide the population into five groups according to relationship and disease status. When we consider the slower epidemic dynamics, we divide the population into just two groups, susceptible and infected individuals, with densities S and I. Each of these groups includes both single and partnered individuals, but on the epidemic timescale we are concerned only with overall densities by disease state. The fast and slow dynamics are linked because new infections arise only in SI partnerships (at a constant rate β_{pair}), so the incidence rate derived from the rapid pairing model is:

$$\varepsilon_{\text{rapid pairing}} = \beta_{\text{pair}} P_{\text{SI}}^{*} \tag{2}$$

where P_{SI}^{*} is the steady-state density of SI partnerships.

Our goal is to express P_{SI}^{*} in terms of the basic quantities S and I, and thus to translate our complex model of partnership dynamics into the standard framework of compartmental epidemic models (SIS, SIR, SEIR, etc.) (Anderson and May 1991; Hethcote 2000). The timescale approximation makes this possible—in the appendix we derive exact solutions for P_{SI}^{*} in terms of S, I, and the pair formation and dissolution rate constants. In the simplest behavioral case, where single individuals of all types enter partnerships at rate k and partnerships of all types break up at rate l, we obtain:

$$P_{\text{SI}}^{*} = \left(\frac{k}{k+l} \right) \frac{SI}{N} . \tag{3}$$

This result can be understood intuitively. When the disease does not affect pairing behavior the total steady-state density of partnerships is $\left(\frac{k}{k+l}\right)\frac{N}{2}$, and the proportion of those which are SI pairs is $\frac{2SI}{N^2}$. (The former expression is derived in the appendix. The latter follows from the second term in the binomial expansion, $\left(\frac{S}{N}+\frac{I}{N}\right)^2 = \frac{S^2}{N^2}+\frac{2SI}{N^2}+\frac{I^2}{N^2}$, which gives the relative proportions of SS, SI, and II pairs assuming random mixing.) Equation (3) is simply the product of these two expressions.

Substituting our solution for P_{SI}^{*} into Equation (2), we obtain a total incidence rate of $\beta_{\mathrm{pair}}\left(\frac{k}{k+l}\right)\frac{SI}{N}$. This expression identifies with Equation (1), the classical frequency-dependent transmission function, if $c_{\mathrm{FD}}p_{\mathrm{FD}} = \beta_{\mathrm{pair}}\left(\frac{k}{k+l}\right)$. We seek to understand this equality by viewing c_{FD} and p_{FD} from the perspective of our pair-formation model. The mean duration of a partnership is $1/l$, and the mean period between partnerships is $1/k$. Assuming sequential monogamy, on average each individual will have one new sex partner every $\left(1/l + 1/k\right)$ units of time. The mean number of partners per unit time is thus $c_{FD} = \frac{1}{1/l + 1/k} = \frac{kl}{k+l}$. Also, the susceptible individual in an SI partnership has a constant risk of infection per unit time (or hazard rate) of β_{pair}. For a relationship of average duration ($1/l$), the probability that transmission occurs is then $p_{\mathrm{FD}}=1-\exp(-\beta_{\mathrm{pair}}/l)\approx\beta_{\mathrm{pair}}/l$ (the latter follows because we have assumed disease dynamics are much slower than pair dynamics, implying $\beta_{\mathrm{pair}} << l$). Expressing c_{FD} and p_{FD} in terms of our model parameters, therefore, we find

$c_{FD} p_{FD} = \left(\frac{kl}{k+l}\right) \times \left(\beta_{pair} / l\right) = \beta_{pair} \left(\frac{k}{k+l}\right)$. This is precisely the equality obtained above by setting $\varepsilon_{\text{rapid pairing}} = \varepsilon_{FD}$. In summary, we have derived both the functional form of the classical frequency-dependent incidence ($\varepsilon \propto \frac{SI}{N}$) and its conventional prefactor ($c_{FD} p_{FD}$) from a mechanistic model of partnership dynamics.

3. Validity of frequency-dependent models of STD transmission

The assumption that pairing processes reach steady-state on epidemic timescales allows us to link frequency-dependent incidence to pair-based transmission. To assess when our separation of fast and slow timescales is valid, we compare epidemics simulated with classical frequency-dependent transmission (i.e. using the timescale approximation) to those generated from a full pair-formation/epidemic system (i.e. without the approximation). The latter model closely matches previous studies of pair formation and STDs (e.g. Dietz and Hadeler 1988), and includes processes that are missed under the timescale approximation (namely the loss of SI partnerships due to recovery of the infected individual, the gain of SI partnerships due to recovery of one individual in an II pair, and partnerships ending due to the death of one partner). We thus can assess the accuracy with which frequency-dependent models represent pair-based transmission, for given sets of parameters, by the similarity of the simulated epidemics. To reach general conclusions, we compare epidemics for different types of STD, and in populations with pair dynamics occurring at a range of rates.

STDs have been divided into two groups by natural history: those with high transmission probability and short duration (chiefly bacterial pathogens) and those with low transmission probability and long duration (chiefly viruses) (Blanchard 2002;

Garnett 2002). We simulated epidemics for these two main classes of STD. To represent bacterial STDs, we drew parameter values from the literature on gonorrhea, chlamydia, and trichomoniasis (Kretzschmar et al. 1996; Garnett et al. 1999; Bowden and Garnett 2000): transmission rates are high, and infected individuals recover without immunity in roughly one month. To represent viral STDs, we chose parameters in the range appropriate for HIV and HSV-2 (Anderson et al. 1989; Castillo-Chavez 1989; Mertz et al. 1992; Blower et al. 1998; Blower et al. 2000): transmission is relatively slower, but there is no recovery. Parameter values for the two model diseases are specified in the caption to Figure 2, as are basic reproductive numbers R_0 (calculated as described in Section 4). In both cases individuals are assumed to remain in the population for 10 years. Pairing parameters were chosen such that healthy individuals divide their time about equally between single and partnered states, and we follow Kretzschmar et al. (1996) in assuming a mean of one sex act per day in casual partnerships. Asymptomatic cases and variable infectivity are not treated explicitly, but can be considered to be averaged into the relevant rate constants. Admittedly these are caricatures of the actual diseases, but they serve to illustrate the behavior of two important classes of STDs (and of diseases acting on two different timescales).

For bacterial and viral STDs, we assessed the accuracy with which frequency-dependent incidence represents pairing at different timescales (Figure 2), in the simplest case when all pairing and unpairing rates have the same value $k=l=1/D$. We describe the timescale of pair dynamics in terms of the mean partnership duration D, which in this case is also the mean time between partnerships. In all cases recall that durations are exponentially distributed with mean D, so that many partnerships would last

considerably longer (or shorter) than D time units. For both classes of STD the full-system simulation approaches the timescale-approximated solution as pairing and unpairing rates ($1/D$) increase to infinity. Exact correspondence with frequency-dependent transmission thus is obtained only for instantaneous contact duration. For finite partnership (and between-partnership) durations, the epidemic always proceeds more slowly, since opportunities to transmit infection are more limited (Kretzschmar 2000).

The results are radically different for the two disease classes, owing to their distinct intrinsic timescales. For the faster bacterial STDs, the epidemics predicted by the full-system simulation and by frequency-dependent transmission diverge rapidly as D increases (Figure 2A). For mean durations of 1 day or less (extremely fast partner change) the epidemics are roughly equivalent—differing by less than 10% in final prevalence and 25% in time to half-maximum prevalence. When D increases to just 1 week the full-system curve is hardly recognizable. The viral STD (Figure 2B), with its slower disease processes and higher reproductive number, is more forgiving of slower pair dynamics. For mean partnership durations of up to 1 month the full-system epidemic is roughly equivalent (as defined above) to the frequency-dependent approximation, and even with D=3 months the epidemics are qualitatively similar (though their rates of growth differ considerably). As a general rule, we have found that the agreement between the full-system and frequency-dependent models depends on both the relative timescales of disease and pairing dynamics and the reproductive number of the disease: slower disease dynamics and higher values of R_0 permit longer-term partnerships to be modeled accurately by frequency-dependence.

4. Infection-induced changes in pairing behavior

For scenarios where the timescale approximation is reasonable (discussed in Section 5), we now extend our derivation in keeping with strong evidence that the infection status of individuals influences their contact behavior. We introduce different rates k_S and k_I at which susceptible and infected individuals enter partnerships, and different rates l_{SS}, l_{SI} and l_{II} at which the three types of partnerships dissolve (see Equation (A2) for implementation details). In the appendix, we derive analytic solutions for the steady-state SI pair density P_{SI}^* from the resulting equations. As above we can use these in conjunction with Equation (2) to express the total incidence rate in closed-form expressions. By defining the dimensionless proportions $s=S/N$ and $i=I/N$, and representing our contact parameters by the vector $\boldsymbol{\kappa} = (k_S, k_I, l_{SS}, l_{SI}, l_{II})$ we find that all solutions share a generalized frequency-dependent form:

$$P_{SI}^* = \phi_{\boldsymbol{\kappa}}(s,i)\frac{SI}{N} \tag{4}$$

where $\phi_{\boldsymbol{\kappa}}(s,i)$ is a function of time-varying values of s and i (though note $s+i=1$ for all time) and the pair formation and dissolution rates $\boldsymbol{\kappa}$. The $\phi_{\boldsymbol{\kappa}}(s,i)$ for four different behavioral cases are shown in Table 1, encompassing situations where individual infection status influences either pair-formation rates or partnership durations or both, as well as the baseline case of no behavioral shifts.

The $\phi_{\boldsymbol{\kappa}}(s,i)$ are independent of total density N, as expected since the pairing process can be described entirely in terms of frequencies (by dividing both sides of Equation (A2) by N). Hence, as for the simple case in Section 2, we find that $\varepsilon_{\text{rapid pairing}} \propto SI/N \propto N$, i.e. that the total incidence increases linearly with N. Therefore the per

capita disease risk (the force of infection) is not influenced by N, so even these behaviorally complex models exhibit the density-independence characteristic of the frequency-dependent formulation.

Incorporating infection-induced changes in pairing behavior leads to non-linearities in the dependence of the incidence rate on the relative proportions of susceptible (s) and infectious (i) individuals. The incidence rate will thus vary from the standard frequency-dependent pattern ($\varepsilon_{\text{rapid pairing}} \propto siN$) during the course of an epidemic. As an example consider Case 2 in Table 1: if infected individuals have reduced pairing rates (i.e. $k_I < k_S$, and hence $\pi_I < \pi_S$), the value of $\phi_{\kappa}(s,i)$ increases as i gets larger. The incidence rate will accelerate as the epidemic progresses, relative to what is expected based on homogeneous pairing behavior (Case 1). This somewhat counterintuitive finding arises because the less numerous type limits the rate of SI pair formation, so k_S has increasing influence as i approaches 1. Such dynamics could be significant for the current discussion of the effect of epidemic phase on epidemiological studies of STDs (Blanchard 2002; Garnett 2002). Note that this effect is opposite to that usually treated in models of behavior-change at the population level, in which higher prevalence is assumed to lower transmission rates via reduced contacts or greater precautions (Hadeler and Castillo-Chavez 1995; Hyman and Li 1997; Hsieh and Sheu 2001).

To understand the effects pairing behavior being influenced by illness, we derive the fundamental epidemiological quantities for the SIS model with constant recruitment shown in Figure 1 (also see Equation (A4)). The basic reproductive number, R_0, is the expected number of secondary cases caused by an infectious individual in a wholly

susceptible population (Diekmann and Heesterbeek 2000). The threshold condition for disease invasion is $R_0>1$, since the first infection must more than replace itself for the disease to become established. The calculation of R_0 is described in the appendix—remarkably, for all four behavioral cases we obtain the same result:

$$R_0 = \frac{\beta_{\text{pair}}}{\sigma+\mu}\left(\frac{k_{\text{I}}}{k_{\text{I}}+l_{\text{SI}}}\right) \tag{5}$$

provided $k_S>0$ (and k_I and l_{SI} are replaced by k and l in the appropriate cases.)

This outcome is surprising for several reasons. First we note that equation (5) is identical to the standard result for frequency-dependent transmission, $R_0 = \frac{c_{\text{FD}}p_{\text{FD}}}{\sigma+\mu}$, if we equate $c_{\text{FD}}p_{\text{FD}} = \beta_{\text{pair}}\frac{k_{\text{I}}}{k_{\text{I}}+l_{\text{SI}}}$ (Hethcote 2000). From the arguments in Section 2, this equivalence requires that c_{FD} be interpreted as the rate at which infected individuals acquire new sexual partners. Some authors use this convention (e.g. McCallum et al. 2001) but many do not (e.g. Anderson et al. 1989; Hethcote 2000).

Second, R_0 is independent of several key parameters of the model. The rate at which susceptible individuals enter contact partnerships, k_S, does not affect the ability of the disease to invade a population (provided $k_S>0$). While not obvious, this can be understood by noting that the early formation of SI pairs will be rate-limited by the supply of infectious individuals. R_0 is also independent of l_{SS} and l_{II}, and hence of the durations of SS and II partnerships. Again this is unexpected, since individuals in SS (or II) partnerships are effectively vaccinated (or case isolated) due to their removal from the mixing population, and consideration of such partnerships is expected to counter the spread of an STD (Dietz and Hadeler 1988; Kretzschmar 2000). The number of II pairs at the time of invasion is negligibly small, however, and furthermore

the timescale approximation precludes the existence of long-lived SS (or II) pairs, and hence eliminates their potential vaccinating (or isolating) effect.

When R_0>1, the disease is able to invade and we are interested in its equilibrium properties. The calculation of i_∞, the steady-state endemic prevalence, is described in the appendix. Results are shown in Table 1, where we see that even the heterogeneous behavior cases are modifications of the classical solution $i_\infty = 1 - \frac{1}{R_0}$ (Anderson and May 1991). The dependence of i_∞ on the pair formation and dissolution rates is shown in Figure 3A, and in contrast to the R_0 result these trends immediately match our expectations. As k_S or k_I is increased (and hence susceptible or infectious individuals enter partnerships more quickly) the steady-state prevalence rises. The infected proportion drops as l_{SI} is raised, since SI pairs break up more quickly and the window for transmission grows shorter. As l_{SS} or l_{II} are increased, i_∞ increases due to the briefer protection offered by the effective vaccine and isolation of SS and II partnerships. We see i_∞ go to zero—failure of the disease to persist—only as k_I and l_{SI} change, because these are the only pairing parameters which influence R_0.

To demonstrate the potential dynamical effects of contact behavior varying with infection status, we simulated epidemics of bacterial and viral STDs for varying values of k_I (Figures 3B, 3C). (Note that to apply the timescale approximation for a bacterial STD, we model an extremely promiscuous population with mean partnership durations of 1 day.) The solid curves show the base case where disease does not influence behavior, while the broken lines show the outcomes when the infection causes individuals to decrease (or increase) their pair-formation rate by the proportions shown. For the bacterial STD (Figure 3B), diminished values of k_I lead to drastically altered

epidemics, since they bring R_0 close to, and even below, the threshold value of one. In the base case, R_0=2.48. When k_I is diminished by 50% (yielding R_0=1.65) the epidemic increases roughly half as fast (as measured by the time to half-maximum prevalence), but a 70% drop (R_0=1.14) slows it by more than ten-fold. A 90% reduction causes the disease to die out, since R_0<1. For the viral STD (Figure 3C), R_0=8.33 in the base case (considerably further from the threshold range), so reductions in k_I have much weaker effects than for the bacterial disease. Even when pairing activity of infected individuals is reduced by 70%, R_0 is still nearly 4 and the epidemic progresses at almost half its speed in the base case. Only when k_I is reduced ten-fold, yielding R_0~1.5, do dramatically slower dynamics result. For both diseases increases in k_I cause only incremental changes, since R_0 is already significantly greater than one.

5. Discussion

We have analysed the dynamics of pair-based disease transmission for populations with rapid partner exchange. Beginning with a pair-formation model, we presented a mechanistic derivation of the classical frequency-dependent incidence. We used this derivation in two ways: first as a formal framework to assess how accurately frequency-dependent incidence portrays the pair-based transmission known to underlie STD dynamics, then as a platform from which we derived extensions of classical epidemiological theory that include effects of illness on contact behavior.

Many assumptions were required in our derivation. Unless another mechanistic link between frequency-dependent incidence and pair-based transmission can be found, it seems that all of these assumptions are made implicitly whenever STD dynamics are

modeled using frequency-dependent transmission. Some of these assumptions are widely recognized: partnerships form by random mixing, with no memory of past contacts or consideration of social clustering. Individuals are sequentially monogamous, and mixing is proportionate with respect to disease status. For mathematical simplicity we did not subdivide the population by sex, age, or sexual activity class, but such structure is often included in frequency-dependent STD models.

The most challenging assumption, however, is the timescale approximation, wherein we assert that pairing processes are sufficiently faster than epidemic dynamics that we can assume they are at equilibrium. From our simulation results (Figure 2), we see that predictions obtained using this approximation equal those from a full pair-dynamic/epidemic model only in the limit of instantaneous partnerships. When partnerships have finite duration, the approximated epidemic always grows faster and reaches a higher final prevalence than the full system. The range of pairing timescales for which the two epidemics are similar, though, depends strongly on the disease in question, and our findings support Garnett's division of STDs into two functional groups by natural history (Garnett 2002). Suppose we say the timescale approximation is satisfied when the solutions differ by less than 10% in final prevalence and 25% in time to half-maximum prevalence. For relatively fast, transient STDs such as chlamydia (Figure 2A), only very promiscuous populations change partners rapidly enough that the approximation is satisfied—D=1 day implies roughly 180 sexual partners per year. (Note that a model specific to gonorrhea would have β_{pair} about twice as high as that used in Figure 2A, and thus even faster partner change would be

required.) For slower-moving, chronic STDs such as HIV, the approximation is satisfied for partnerships with mean duration of a month or more.

Results derived from the timescale approximation—including the classical frequency-dependent incidence and our results for infection-induced behavior changes—must therefore be used advisedly. For fast-moving, bacterial STDs, the results will be accurate only for a small subset of human populations. Failure to recognize this danger can lead to catastrophically poor predictions, as shown in Figure 2A where partnerships more than a few days long generate a simulated epidemic which differs drastically from that predicted by the classical model. For chronic, viral STDs, expressions derived from the timescale approximation are more widely applicable (see relationship data in Anderson et al. 1989), but great caution is required in extending models beyond core groups to general human populations. Note that concurrent sexual relationships, not considered here, effectively increase the rate of partner change and would soften these restrictions somewhat. Clearly the frequency-dependent models could also be applied to STDs in animal populations (Lockhart et al. 1996), and are particularly suitable for the many species whose mating systems feature short-lived, monogamous pairings.

For situations where the timescale approximation is reasonable, we extend our mechanistic derivation to develop new tools for disease modeling. There is rich evidence in humans and animals that sickness influences contact behavior, yet this phenomenon is largely overlooked in the theory of disease spread. We provide expressions for the incidence rate, basic reproductive number (R_0), and endemic prevalence (i_∞), for four cases in which pairing behavior is influenced increasingly by

infection. These expressions can be incorporated directly into SI or SIS epidemic models, and can be generalized readily to any SEIR-type model when only infectious (I) individuals display altered behavior.

We found a common expression for R_0 for all cases we considered, revealing that the invasion ability of a disease depends on the rate at which infectious individuals enter partnerships (k_I) and the duration of SI pairs ($1/l_{SI}$), but not on other pairing parameters. Concordant SS and II partnerships do not influence the disease's threshold properties—despite their expected roles as effective vaccine and case isolation—though they do influence i_∞ as expected. Our general expression for R_0 is identical to the classical result for frequency-dependent epidemics, subject to careful interpretation of the contact rate (c_{FD}) as the rate at which infected individuals acquire sexual partners. This finding accords with the one previous study (that we know of) that explicitly linked pair formation and break-up to the infection status of the individuals involved. Dietz & Hadeler (1988) presented a threshold analysis of a pair-formation model with parameters which differ by infection status, and found (as we did) that the classical result appears in the limit of fast pair dynamics.

It is important to distinguish our findings from previous models of "behavior change" in STD epidemics (Hadeler and Castillo-Chavez 1995; Hyman and Li 1997; Hsieh and Sheu 2001). In these studies, susceptible individuals may choose to change their rate of acquiring new sexual partners (c_{FD}) or probability of transmission per partnership (p_{FD}), typically as a result of public education campaigns or community awareness of disease risk. This is certainly an important topic in current STD epidemiology, but it is fundamentally different from the individual behavioral effects

addressed here. Our treatment considers the direct effects of individual infection status on individual contact behavior: we allow per capita pair-entry rates to differ between susceptible and infected individuals, and break-up rates to differ among SS, SI, and II pairs. This approach hinges on considering the distinct contributions of both individuals in forming a pair, and leads to nonlinear dependence of the transmission rate on the proportions of S and I individuals (beyond the usual bilinear *SI* form). A recent study found such non-linearity to be a common feature of incidence data (Fenton et al. 2002); disease effects on contact behavior are an unexplored mechanism for this trend. Simulations in which individual pairing rates were reduced (or increased) by infection demonstrated that these behavioral effects can have striking impacts on epidemic progression (Figure 3B-C).

These results may also be applicable to diseases transmitted by non-sexual contact (the results of Heesterbeek & Metz (1993) also apply to this problem). Diseases of casual contact increasingly are being modeled using frequency-dependent incidence, and it is reasonable to speculate on the effect of disease symptoms on casual contact behavior. The timescale approximation (of rapid contact dynamics) is likely to be valid for such diseases, but the assumption that partnerships are exclusive breaks down—higher-order groupings and simultaneous contacts may play a significant role. This question merits additional investigation, which would also be relevant to concurrent partnerships in STD models.

Links between analytic theory and complex simulation models are of vital importance, serving to distill insights and unify rapidly-growing fields of research. This point has been emphasized in the context of epidemiology (Anderson and Garnett 2000;

McCallum et al. 2002), and a similar goal is being pursued using correlation techniques for network models of disease (Bauch and Rand 2000; Ferguson and Garnett 2000; Eames and Keeling 2002). In exploring the relation between partnership dynamics and the classical frequency-dependent treatment of transmission, we have understood more clearly both the meaning of the classical model and the proper limits of its application. We have also extended theoretical results to include changes of contact behavior due to illness, so that this well-recognized effect can be incorporated into simple, analytic models when biological evidence requires it.

Table 1. Results for epidemics with infection-induced changes in behavior

Case	Rates, κ	$\phi_\kappa(s,i)$	i_∞
1	$k_S=k_I=k$ $l_{SS}=l_{SI}=l_{II}=l$	$\frac{k}{k+l}$	$1-\frac{1}{R_0}$
2	$k_S\neq k_I$ $l_{SS}=l_{SI}=l_{II}=l$	$\frac{\pi_S\pi_I}{\pi_S s+\pi_I i}$	$1-\frac{1}{R_0\left(\frac{\pi_S}{\pi_I}+\left(1-\frac{\pi_S}{\pi_I}\right)\right)}$
3	$k_S=k_I=k$ $l_{SS}\neq l_{SI}\neq l_{II}$	$\frac{\pi}{\frac{1}{2}+\frac{1}{2}\sqrt{1-4a\pi^2 si}}$	$1-\frac{1-a\pi^2/R_0}{R_0-a\pi^2/R_0}$
4	$k_S\neq k_I$ $l_{SS}\neq l_{SI}\neq l_{II}$	$\frac{\pi_S\pi_I}{\frac{1}{2}\left(\pi_S s+\pi_I i+\sqrt{\left(\pi_S s+\pi_I i\right)^2-4a\left(\pi_S\pi_I\right)^2 si}\right)}$	$1-\frac{1-a\pi_S\pi_I/R_0}{R_0\left(\frac{\pi_S}{\pi_I}+\left(1-\frac{\pi_S}{\pi_I}\right)\right)-a\pi_S\pi_I/R_0}$

Here $s=S/N$, $i=I/N$, $\pi_S=\frac{k_S}{k_S+l_{SI}}$ and $\pi_I=\frac{k_I}{k_I+l_{SI}}$, and

$a=\frac{l_{SI}}{k_I}\left(1-\frac{l_{SI}}{l_{SS}}\right)+\frac{l_{SI}}{k_S}\left(1-\frac{l_{SI}}{l_{II}}\right)+\left(1-\frac{l_{SI}^2}{l_{SS}l_{II}}\right)$. If $k_S=k_I$, then $\pi_S=\pi_I\equiv\pi$.

R_0 is shown in Equation (5).

Figure captions

Figure 1

STD epidemic model. The fast pairing dynamics and slow epidemic dynamics are linked by a timescale approximation, as described in the text. *Fast pairing dynamics*: X_S and X_I are the densities of susceptible and infectious single individuals, respectively, while P_{SS}, P_{SI}, and P_{II} represent partnerships between individuals of the disease states shown in the subscripts. Single individuals in the X_S and X_I pools enter partnerships at per capita rates k_S and k_I, independent of population density. Pairings of type P_{yz} break up at rate l_{yz} (for y,z=S or I), and hence have lifetimes which are distributed exponentially with mean $1/l_{yz}$. For the simple case described in Section 2, we set $k_S=k_I=k$ and $l_{SS}=l_{SI}=l_{II}=l$. When individuals of type y form partnerships, a proportion m_{yz} will be with individuals of type z (hence we always have $\sum_z m_{yz}=1$). *Slow epidemic dynamics*: Transmission occurs within SI pairs (at quasi-steady-state density P_{SI}^*) at rate β_{pair}, which is the product of the within-partnership rate of sex acts times the probability of transmission per sex act. Infected individuals recover at rate σ. There is constant influx λ into the population, and individuals leave the population (by death, emigration, or entering long-term relationships) at rate μ. We assume that the population density has reached an equilibrium value $N=S+I$, and hence $\lambda=\mu N$.

Figure 2

Accuracy of timescale approximation for pair dynamics on different timescales. (A) Epidemic of bacterial STD. The heavy line shows the timecourse predicted using the frequency-dependent incidence rate derived with the timescale approximation

(Equations 3 and A4); the lighter lines show the results of simulating the full system (Equation A1). Pairing timescales are described by D, which represents the mean duration of partnerships and between-partnership periods. Disease parameters: β_{pair}=0.15 day^{-1}, σ=0.03 day^{-1}. R_0=2.48 under the timescale approximation. (B) Epidemic of viral STD. As in (A), except β_{pair}=0.005 day^{-1} and σ=0, and therefore R_0=8.33. In all cases μ=0.0003 day^{-1}, and at t=0, S_0=0.99, I_0=0.01. Results were obtained using a second-order modified Rosenbrock method for stiff ODEs (algorithm ode23s in MATLAB v6.1 (Mathworks, Natick MA)).

Figure 3

Effect of infection-induced changes in behavior. (A) Dependence of the steady-state endemic prevalence, i_∞, on pairing rate parameters, calculated using the expressions shown in Table 1. The parameters were varied one at a time, while the non-varying parameter values were as for (C), below. (B) Epidemic timecourses for a bacterial STD. In the base case (solid line) k_S=k_I=1 day^{-1}, while in other cases (dotted lines) the pairing rate for infectious individuals (k_I) is reduced or increased by the proportion shown. (C) Epidemic timecourses for a viral STD. Disease parameters and initial conditions as in Figure 2, and l_{SS}=l_{SI}=l_{II}=1 day^{-1}.

Figure 1

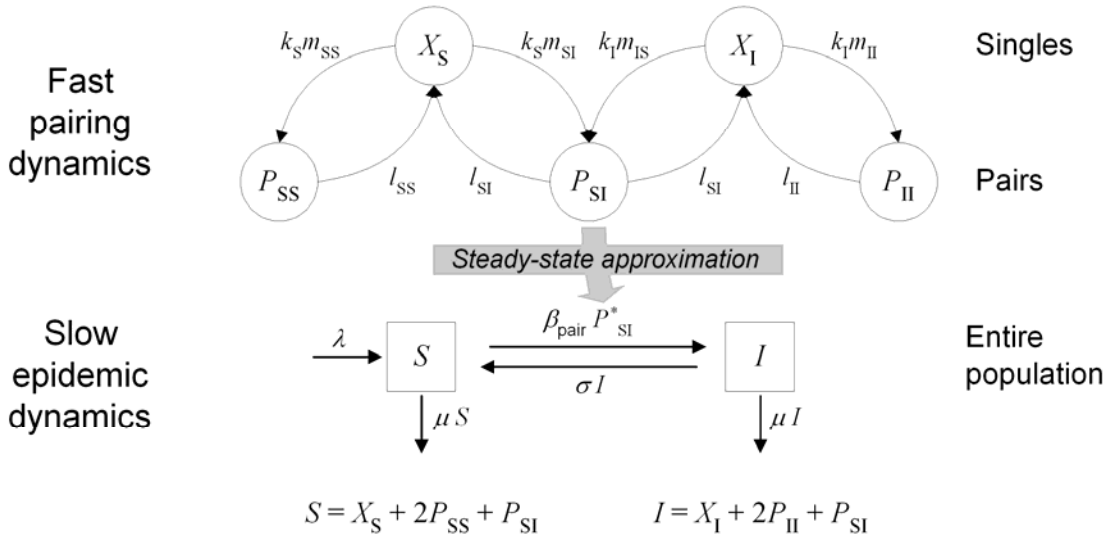
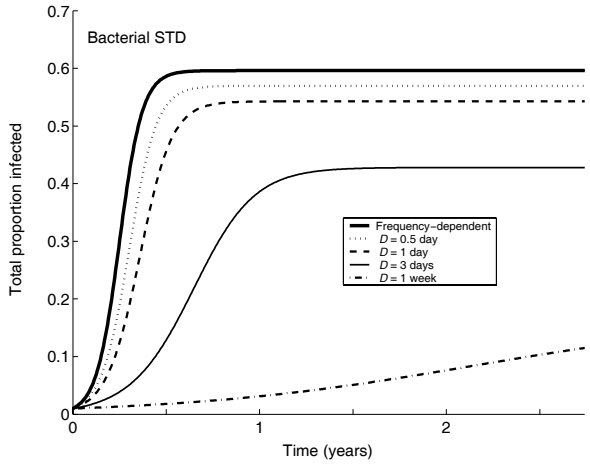

Fast pairing dynamics
$k_S m_{SS}$
X_S
$k_S m_{SI}$
$k_I m_{IS}$
X_I
$k_I m_{II}$
Singles
P_{SS}
l_{SS}
l_{SI}
P_{SI}
l_{SI}
l_{II}
P_{II}
Pairs
Steady-state approximation
Slow epidemic dynamics
λ
S
$\beta_{pair} P^*_{SI}$
σI
I
μS
μI
Entire population
$S = X_S + 2P_{SS} + P_{SI}$
$I = X_I + 2P_{II} + P_{SI}$

Figure 2

A)


Bacterial STD
Total proportion infected
Time (years)
Frequency−dependent
D = 0.5 day
D = 1 day
D = 3 days
D = 1 week
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0
1
2

B)

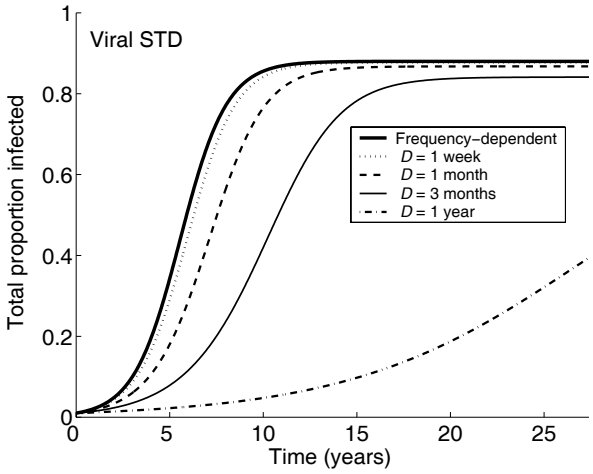
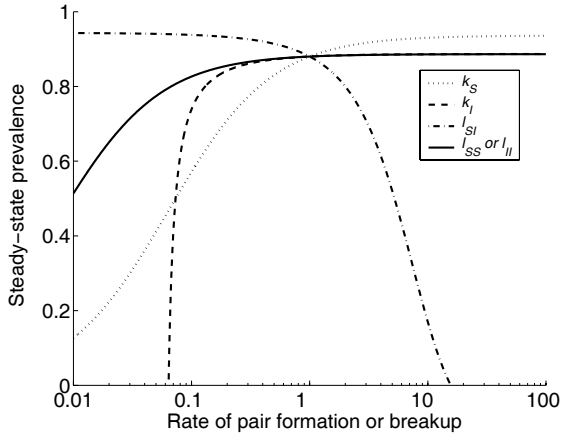
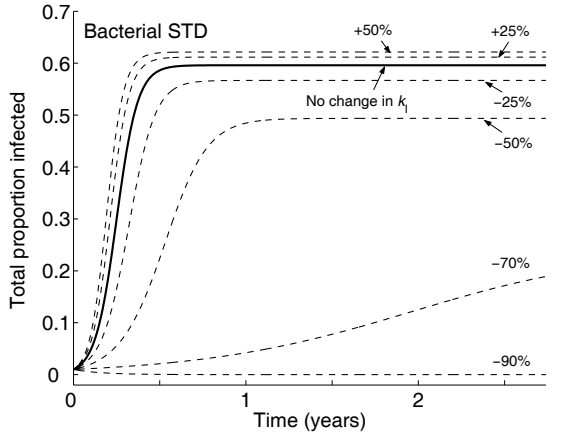

Viral STD
Total proportion infected
Time (years)
Frequency−dependent
D = 1 week
D = 1 month
D = 3 months
D = 1 year
0
0.2
0.4
0.6
0.8
1
0
5
10
15
20
25

Figure 3

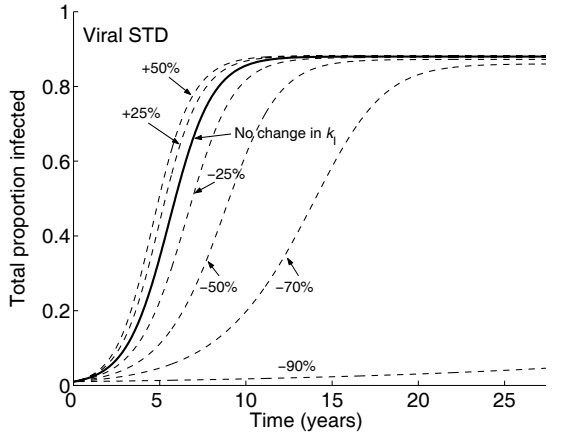
A)


Steady−state prevalence
Rate of pair formation or breakup
k_S
k_I
I_{SI}
I_{SS} or I_{II}

B)


Bacterial STD
Total proportion infected
Time (years)
+50%
+25%
No change in k_I
−25%
−50%
−70%
−90%

C)


Viral STD
Total proportion infected
Time (years)
+50%
+25%
No change in k_I
−25%
−50%
−70%
−90%

Appendix

Derivation of SI pair density

Here we present a general derivation of the steady-state SI pair density P_{SI}^{*}, including the results shown in Sections 2 and 4. We begin with the standard one-sex formulation of a pair-formation/epidemic model, extended such that pair-entry rates k_y and break-up rates l_{yz} (where y,z=S or I) can vary as a function of infection status (c.f. Dietz and Hadeler 1988):

$$\begin{aligned}
\frac{dX_{\mathrm{S}}}{dt} &= -k_{\mathrm{S}}X_{\mathrm{S}} + 2l_{\mathrm{SS}}P_{\mathrm{SS}} + l_{\mathrm{SI}}P_{\mathrm{SI}} + \sigma X_{\mathrm{I}} + \lambda - \mu X_{\mathrm{S}} + \mu\left(2P_{\mathrm{SS}} + P_{\mathrm{SI}}\right) \\
\frac{dX_{\mathrm{I}}}{dt} &= -k_{\mathrm{I}}X_{\mathrm{I}} + 2l_{II}P_{II} + l_{\mathrm{SI}}P_{\mathrm{SI}} - \sigma X_{\mathrm{I}} - \mu X_{\mathrm{I}} + \mu\left(2P_{\mathrm{II}} + P_{\mathrm{SI}}\right) \\
\frac{dP_{\mathrm{SS}}}{dt} &= \tfrac{1}{2}k_{\mathrm{S}}m_{\mathrm{SS}}X_{\mathrm{S}} - l_{\mathrm{SS}}P_{\mathrm{SS}} + \sigma P_{\mathrm{SI}} - 2\mu P_{\mathrm{SS}} \\
\frac{dP_{\mathrm{SI}}}{dt} &= \tfrac{1}{2}k_{\mathrm{S}}m_{\mathrm{SI}}X_{\mathrm{S}} + \tfrac{1}{2}k_{\mathrm{I}}m_{\mathrm{IS}}X_{\mathrm{I}} - l_{\mathrm{SI}}P_{\mathrm{SI}} - \beta_{\mathrm{pair}}P_{\mathrm{SI}} - \sigma P_{\mathrm{SI}} + 2\sigma P_{\mathrm{II}} - 2\mu P_{\mathrm{SI}} \\
\frac{dP_{\mathrm{II}}}{dt} &= \tfrac{1}{2}k_{\mathrm{I}}m_{\mathrm{II}}X_{\mathrm{I}} - l_{\mathrm{II}}P_{\mathrm{II}} + \beta_{\mathrm{pair}}P_{\mathrm{SI}} - 2\sigma P_{\mathrm{II}} - 2\mu P_{\mathrm{II}}
\end{aligned} \qquad \text{(A1)}$$

Parameters and variables are defined in the caption to Figure 1. The factors of ½ reflect that it takes two individuals to make a partnership (i.e. one unit of P_{yz} is equivalent to two units of X_{S} or X_{I}), and the factors of 2 in μ and σ terms reflect events that arise from transitions undergone by either member of a pair. Note that mixed partnerships (P_{SI}) are formed both by X_{S} individuals "choosing" X_{I} individuals and vice versa; the two separate terms are essential to maintaining constant population size. We assume that population density has reached equilibrium, and set $\lambda=\mu N$.

This derivation pertains to populations where pair formation and dissolution occur on faster timescales than disease and demographic processes (i.e. pairing rate parameters k_y and l_{yz} are significantly greater than epidemic rates β_{pair}, σ and μ). We

therefore approximate that disease states are constant on the timescale of pairing processes, and separate the fast pairing dynamics from the slow epidemic dynamics as described in the main text and shown in Figure 1. The pairing dynamics are then described by:

$$\text{Fast pairing dynamics} \quad \begin{cases} \dfrac{dX_{\mathrm{S}}}{dt} = -k_{\mathrm{S}} X_{\mathrm{S}} + 2l_{\mathrm{SS}} P_{\mathrm{SS}} + l_{\mathrm{SI}} P_{\mathrm{SI}} \\ \dfrac{dX_{I}}{dt} = -k_{\mathrm{I}} X_{\mathrm{I}} + 2l_{\mathrm{II}} P_{\mathrm{II}} + l_{\mathrm{SI}} P_{\mathrm{SI}} \\ \dfrac{dP_{SS}}{dt} = \tfrac{1}{2} k_{\mathrm{S}} m_{\mathrm{SS}} X_{\mathrm{S}} - l_{\mathrm{SS}} P_{\mathrm{SS}} \\ \dfrac{dP_{SI}}{dt} = \tfrac{1}{2} k_{\mathrm{S}} m_{\mathrm{SI}} X_{\mathrm{S}} + \tfrac{1}{2} k_{\mathrm{I}} m_{\mathrm{IS}} X_{\mathrm{I}} - l_{\mathrm{SI}} P_{\mathrm{SI}} \\ \dfrac{dP_{II}}{dt} = \tfrac{1}{2} k_{\mathrm{I}} m_{\mathrm{II}} X_{\mathrm{I}} - l_{\mathrm{II}} P_{\mathrm{II}} \end{cases} \tag{A2}$$

We consider the slower dynamics of disease transmission and demographic processes at the whole-population scale. First we collect the total densities (in and out of partnerships) of susceptible and infectious individuals into variables S and I:

$$\begin{aligned} S &= X_{\mathrm{S}} + 2P_{\mathrm{SS}} + P_{\mathrm{SI}} \\ I &= X_{\mathrm{I}} + 2P_{\mathrm{II}} + P_{\mathrm{SI}} \end{aligned} \tag{A3}$$

The epidemic can now be represented with a standard SEIR-type compartmental model; as an example we treat an SIS epidemic with constant recruitment rate λ. The total incidence rate is $\beta_{\mathrm{pair}} P_{\mathrm{SI}}^{*}$ (Equation 2), the mortality rate μ is independent of disease status, and the recovery rate is σ. Thus:

$$\text{Slow epidemic dynamics} \quad \begin{cases} \dfrac{dS}{dt} = \lambda - \beta_{\mathrm{pair}} P_{\mathrm{SI}}^{*} + \sigma I - \mu S \\ \dfrac{dI}{dt} = \beta_{\mathrm{pair}} P_{\mathrm{SI}}^{*} - (\sigma + \mu) I \end{cases} \tag{A4}$$

As described in Section 2, our goal is to find an expression for the steady-state density of mixed SI partnerships, P_{SI}^*, in terms of the population densities *S* and *I*. The first step is to specify the mixing matrix for pair formation. The matrix element m_{yz} is the proportion of partnerships formed by *y*-type individuals which will be with individuals of type *z*. In this study we assume proportionate mixing, hence the m_{yz} are simply the fractional contributions of each group to the total pair formation rate:

$$\begin{aligned} m_{SS} &= m_{IS} = \frac{k_S X_S}{k_S X_S + k_I X_I} \\ m_{SI} &= m_{II} = \frac{k_I X_I}{k_S X_S + k_I X_I} \end{aligned} \tag{A5}$$

Note that $k_S\, m_{SI}\, X_S = k_I\, m_{IS}\, X_I$ under this assumption.

Substituting equations (A3) and (A5) into system (A2) , we get:

$$\begin{aligned} \frac{dP_{SS}}{dt} &= \frac{\frac{1}{2}k_S^2\left(S-2P_{SS}-P_{SI}\right)^2}{k_S\left(S-2P_{SS}-P_{SI}\right)+k_I\left(I-2P_{II}-P_{SI}\right)} - l_{SS}P_{SS} \\ \frac{dP_{SI}}{dt} &= \frac{k_S k_I\left(S-2P_{SS}-P_{SI}\right)\left(I-2P_{II}-P_{SI}\right)}{k_S\left(S-2P_{SS}-P_{SI}\right)+k_I\left(I-2P_{II}-P_{SI}\right)} - l_{SI}P_{SI} \\ \frac{dP_{II}}{dt} &= \frac{\frac{1}{2}k_I^2\left(I-2P_{II}-P_{SI}\right)^2}{k_S\left(S-2P_{SS}-P_{SI}\right)+k_I\left(I-2P_{II}-P_{SI}\right)} - l_{II}P_{II} \end{aligned} \tag{A6}$$

We let the pairing dynamics go to steady-state by setting the right-hand sides of (A6) equal to zero. This leads to a system of three quadratic equations in the three unknowns P_{SS}^*, P_{SI}^*, P_{II}^*, which we wish to solve for P_{SI}^*. This system was simplified using Mathematica (Wolfram Research, Champaign IL), yielding the following quadratic equation in P_{SI}^*:

$$a\left(P_{SI}^{*}\right)^{2}+bP_{SI}^{*}+c=0 \quad \text{where}$$
$$a=\frac{l_{SI}}{k_I}\left(1-\frac{l_{SI}}{l_{SS}}\right)+\frac{l_{SI}}{k_S}\left(1-\frac{l_{SI}}{l_{II}}\right)+\left(1-\frac{l_{SI}^{2}}{l_{SS}l_{II}}\right)$$
$$b=-\left(\frac{\pi_S S+\pi_I I}{\pi_S\pi_I}\right) \qquad \text{(A7)}$$
$$c=SI$$

and $\pi_S=\frac{k_S}{k_S+l_{SI}}$ and $\pi_I=\frac{k_I}{k_I+l_{SI}}$.

We consider four cases, in which infection status has varying influence on pairing behavior (i.e. different sets of the pair formation and dissolution rates k_S, k_I, l_{SS}, l_{SI} and l_{II} have distinct values). When pair dissolution rates are equal ($l_{SS}=l_{SI}=l_{II}=l$), $a=0$ in Equation (A7) and $P_{SI}^{*}=-c/b$. Otherwise the quadratic formula was used to find P_{SI}^{*}. For $a<0$ there is only one positive solution, $P_{SI}^{*}=\frac{1}{2a}\left(-b-\sqrt{b^2-4ac}\right)$. When $0<a<b^2/4c$ both solutions are real and positive, but only $P_{SI}^{*}=\frac{1}{2a}\left(-b-\sqrt{b^2-4ac}\right)$ remains bounded on the $(0,N/2)$ interval (for all numerical tests we have conducted). By reorganizing terms in b^2-4ac into a difference-of-terms squared plus some positive terms, it can be shown that $b^2-4ac>0$ always (for $k_S,k_I,l_{SS},l_{SI},l_{II}>0$) so solutions are always real. Exact solutions for P_{SI}^{*} in all four cases can be expressed in the form shown in Equation (4), and full expressions are shown in Table 1. All solutions have been checked numerically to ensure their validity.

Calculation of R_0 and i_∞

The basic reproductive number, R_0, is the expected number of secondary cases caused by a typical infectious individual in a wholly susceptible population. As such, it

can be calculated as the product of the total rate of transmission per I individual times the expected duration of infectiousness, in the limit $S \to N$ (Anderson and May 1991). From Equation (A4), with P_{SI}^{*} substituted from Equation (4), we find for our model:

$$\begin{aligned} R_0 &= \lim_{S \to N}\left(\text{transmission rate per I individual} \times \text{duration of infectiousness}\right) \\ &= \lim_{S \to N}\left(\beta_{\text{pair}} \phi_{\kappa}(s,i) \frac{S}{N} \times \frac{1}{\sigma + \mu} \right) \\ &= \frac{\beta_{\text{pair}}}{\sigma + \mu} \lim_{S \to N}\left(\phi_{\kappa}(s,i)\right) \end{aligned} \tag{A8}$$

From the expressions $\phi_{\kappa}(s,i)$ in Table 1, it is readily shown that $\lim_{S \to N}\left[\phi_{\kappa}(s,i)\right] = \pi_{\mathrm{I}} = \frac{k_{\mathrm{I}}}{k_{\mathrm{I}} + l_{\mathrm{SI}}}$ for all cases. Therefore R_0 takes the same form for all four levels of infection-induced behavioral shifts, as shown in Equation (5).

We can also calculate the equilibrium density of infectives, I_{∞}, by finding the non-zero solution to $\left(\frac{dI}{dt}\right)_{I^* \neq 0} = 0$. From this we can calculate the steady-state endemic prevalence, $i_{\infty} = \frac{I_{\infty}}{N}$. Results are shown in Table 1.

Calculation of steady-state fraction in partnerships

When disease does not influence pair-formation dynamics, we can simply model the density of unpartnered individuals, X, and of partnerships, P:

$$\begin{aligned} \frac{dX}{dt} &= -kX + 2lP \\ \frac{dP}{dt} &= \tfrac{1}{2}kX - lP \end{aligned} \tag{A9}$$

Since the total population density (N) is constant, one of these equations is redundant. The steady-state is found by setting the right-hand side equal to zero and substituting $X=N-2P$. We can then solve for the steady-state density of partnerships:

$$P^* = \left(\frac{k}{k+l}\right)\frac{N}{2} \tag{A10}$$

Chapter Three

Curtailing transmission of severe acute respiratory syndrome within a community and its hospital

J.O. Lloyd-Smith, A.P. Galvani, and W.M. Getz

1. Introduction

When a previously unknown infectious disease emerges, initial options for control are limited. In the absence of an effective treatment or vaccine, policies are restricted to case management (such as isolation of known cases and quarantine of their contacts) and contact precautions for identifiable high-risk groups. Due to limitations in public health resources and concern about personal freedoms and economic impacts, policy-makers face decisions regarding the relative importance of such efforts. This paper identifies priorities and tradeoffs for control of the newly identified coronavirus that is causing an epidemic of Severe Acute Respiratory Syndrome (SARS) worldwide, particularly in parts of Asia. Our analysis focuses on the role of hospitals and healthcare workers (HCWs) in disease transmission by dividing the threatened population (city or region) into two groups: a hospital community and the community-at-large.

Hospitals have been widely recognized as the highest-risk settings for SARS transmission (Drazen 2003, Lee 2003). Currently HCWs have comprised roughly 63% of SARS cases in Hanoi, 51% in Toronto, 42% in Singapore, 22% in Hong Kong, and 18% in mainland China (Booth et al. 2003, Hong Kong Department of Health 2003, Leo et al 2003, Twu et al 2003, World Health Organization 2003a). Healthcare settings and HCWs are thus an obvious focus for SARS control efforts, with particular concern for preventing leakage of the disease from hospitals back into surrounding communities.

We present a model of a nascent SARS outbreak in a community and its hospital, addressing the relative benefits of case management and contact precautions for containing the disease. This situation holds great relevance because travel from

regions with on-going epidemics has the potential to seed new outbreaks worldwide (Twu et al. 2003). Of greatest concern are countries with poor health infrastructure, and the possibility of a seasonal re-emergence of SARS in the Northern hemisphere's next winter. Furthermore, although our model is parametrized for SARS, the lessons learned will be useful in future outbreaks of novel infectious diseases or pandemic flu strains.

Our analysis is based on a stochastic model, because chance events can greatly influence the early progression of an outbreak. We pay particular attention to capturing realistic distributions for the incubation and symptomatic periods associated with SARS, because the natural history of disease progression is less likely to vary between epidemics than are the mixing and transmission patterns. Disease control measures are described by explicit parameters, so their respective contributions to containment can be teased apart (see Table 1). Rates of hospitalizing infected community members and HCWs are denoted by h_c and h_h, respectively, and q is the rate of quarantining exposed individuals in the community following contact tracing. The transmission rate for case-isolated individuals is modified by a factor κ, reflecting measures such as respiratory isolation and negative pressure rooms, and transmission by quarantined individuals is modified by γ. Hospital-wide contact precautions, such as the use at all times of sterile gowns, filtration masks, and gloves, modify within-hospital transmission by a factor η. A final parameter ρ describes efforts to reduce contact between off-duty HCWs and the community. Control parameters are thus divided into those describing case management measures (h_c, h_h, q, γ, κ) and those describing contact precautions (η and ρ).

The most difficult process to characterize in any epidemic is disease transmission. This can be divided into a contact process and the probability of transmission given contact. The former varies dramatically between communities (due to, for example, usage patterns of public transport, or housing density) and between diseases (due to different modes of transmission); the latter also depends both on the disease (proximity required for transmission) and the community (cultural mores relating to intimacy of contact and hygiene). Thus transmission rates are both disease- and community-dependent and will vary from country to country, as well as between cities within a region (Galvani et al. 2003). This geographic variation translates directly into variations in the basic reproductive number R_0 of the epidemic, which is the average number of secondary cases generated by a "typical" infectious individual in a completely susceptible population, in the absence of control measures (Diekmann & Heesterbeek 2000). For instance, SARS is likely to have different R_0 values in Beijing and Toronto, due to differences in cultural practices, environmental conditions, and population density. Such differences are confirmed by Galvani et al (2003), who report widely varying doubling times for SARS outbreaks in six affected regions.

Recent analyses of SARS incidence data from Hong Kong (Riley et al 2003) and Singapore and other settings (Lipsitch et al 2003) report R_0 to be 2.7 (95% CI: 2.2-3.7) and about 3 (90% CI: 1.5-7.7), respectively. As acknowledged by these authors, estimation of R_0 is complicated by healthcare practices in place before outbreaks are recognized—measuring a disease's rate of spread in the true absence of control is rarely possible. In the case of SARS, non-specific control measures may have helped or hindered early epidemic growth: hospitalization of symptomatic cases reduced

transmission to the general community, but put HCWs at risk due to unprotected medical procedures, possibly contributing to so-called super-spreading events. (Note that the R_0 reported by Riley et al (2003) excludes such events.) Estimates of R_0 from data including non-specific control measures therefore could be biased in either direction. Taking note of the confidence intervals cited above, R_0 values associated with SARS in different parts of the world could easily vary from 1.5 to 5—roughly the same range as has been estimated for influenza (e.g. Hethcote 2000, Ferguson et al. 2003). SARS is thought to be primarily transmitted via large-droplet contact, compared to airborne transmission via droplet nuclei for influenza, but fecal-oral and fomite transmission are suspected in some circumstances (Riley et al 2003, Seto et al. 2003, Wenzel & Edmond 2003).

To obtain general results we evaluate control strategies for scenarios reflecting R_0 values from 1.5 to 5, with particular emphasis on R_0~3, the current most-likely estimate for Hong Kong and Singapore. Because the feasibility of implementing different control measures varies from country to country (due to public health infrastructure, for instance, or concerns about civil liberties), we evaluate the extent to which one control measure can compensate for another. We distinguish the impact of two types of delay in the control response: the delay in isolating (or quarantining) particular individuals, and the delay in implementing a systemic control policy after the first case arises. Finally, we consider the pathways of transmission in our model, to obtain direct insight into the role played by HCWs in containing the epidemic.

2. Model description

Our model divides the population into a healthcare worker (HCW) core group and the general community (denoted by subscripts h and c, respectively); infected individuals may enter case management (subscript m), representing quarantine or case isolation. We categorized disease classes as susceptible (S_c, S_h), incubating (E_c, E_h, E_m), symptomatic (I_c, I_h, I_m), and removed due to recovery or death (R_c, R_h). The model updates in one-day timesteps, representing the smallest interval on which people's activities can be thought to be equivalent. Substructures associated with daily movements are summarized in Figure 1, with details given in the caption. Stochastic transitions of individuals between classes, based on the probabilities shown in Figure 1, were implemented by Monte Carlo simulation.

The baseline transmission rate of symptomatic individuals in the community, β, was chosen to produce the range of R_0 values used in our different scenarios (see below). Given the possibility of pre-symptomatic transmission of SARS (Cyranoski & Abbott 2003, Donnelly et al. 2003), our model allows transmission at a reduced rate $\varepsilon\beta$ by incubating individuals. Disease transmission occurs via the following pathways, with contributions weighted as shown in Figure 1D. The hazard rate of infection for susceptibles in the community, represented by λ_c, contains contributions from unmanaged incubating and symptomatic community members (E_c and I_c), as well as off-duty healthcare workers (E_h and I_h). Quarantined individuals (E_m) transmit to S_c at a reduced rate, reflecting household contacts and breaches of quarantine. The infection hazard rate for HCWs, λ_h, contains contributions from workplace transmission risks, from isolated patients (I_m) and unmanaged incubating and symptomatic co-workers (E_h

and I_h), as well as from off-duty time in the community. To describe case management and control of the epidemic, we further modified transmission (Figure 1D) in terms of parameters γ, κ, η and ρ introduced above (see Table 1). Specifically, in the healthcare environment, transmission occurs at a rate $\eta\beta$ due to hospital-wide precautions, and transmission by isolated patients (I_m) is further modified by the factor κ. Transmission rates between HCWs and community members are modified by the factor ρ. Quarantine reduces contact rates by a factor γ, such that the net transmission rate of quarantined incubating individuals (E_m) is $\gamma\varepsilon\beta$.

When all other parameters are fixed at values representing no control measures, each R_0 from our assumed range of 1.5 to 5 uniquely determines a value of β to be used in our simulations. Throughout this paper, unless otherwise stated, we assume that incubating individuals transmit at one-tenth the rate of symptomatic individuals (i.e. ε=0.1). The effects of control are explored by calculating the values of the case management rates (h_c, h_h, q) and transmission-reduction parameters (ρ, η, γ, κ) required to reduce the effective reproductive number R below 1, where R is the expected number of secondary cases generated from an average infection when a given control policy is in place. Calculation of reproductive numbers R_0 and R for our model is described in the appendix. For a two-pool model such as this, the system-wide reproductive number is the dominant eigenvalue of a 2×2 next-generation matrix (Diekmann & Heesterbeek 2000); individual elements of this matrix (R_{ij}, where $i,j=c$ or h) give insight into the potential for disease spread within and between the c and h pools.

Each simulation is initiated with a single infection in the community. Unless otherwise stated, we model the spread and control of SARS in a population of 100,000

individuals and a hospital of 3,000 individuals (c.f. Dwost et al 2003). In the appendix we assess sensitivity of our results to absolute population size, and also to the relative size of the two pools.

3. Results and Discussion

Stochastic epidemics and the reproductive number

Epidemics with reproductive number R less than 1 tend to fade out, because on average each infection does not replace itself. When R is greater than 1, the epidemic is expected to grow, although if the number of cases is small then random events can lead to fadeout of the disease, particularly if R is close to 1. (Note that these statements apply equally to the basic reproductive number, R_0, and the effective reproductive number under a control strategy, R). Sample simulations of our model exhibit this basic trend (Figure 2A), as we see fadeout for four of five simulations corresponding to R=1.2, two of five for R=1.6 and one of five for R=2. Also note the variability in epidemic timing and rate of growth between realizations of our stochastic model.

Because fadeout is an imprecise concept, we frame our results in terms of "epidemic containment", which we define as eradication of the disease within 200 days of the first case, subject to the additional criterion that fewer than 1% of the population ever become infected. (This criterion is needed because a highly virulent disease can pass through a population within 200 days and still infect a large proportion of individuals before extinguishing itself.) The probability of containment in our model decreases with increasing R (Figure 2B), but is still significantly larger than zero for R~5. (This relationship can be defined precisely for stochastic models simpler than

ours—see Diekmann & Heesterbeek (2000).) We note from Figure 2B that even control measures which do not reduce R below 1 can have a substantial probability of succeeding, provided they are imposed when the number of cases is small.

Effect of case isolation

To identify control strategies sufficient to contain a SARS outbreak, we consider parameter combinations which reduce the effective reproductive number to 1. We plot R=1 contours for a range of R_0 values in 2-dimensional parameter spaces (Figures 2C-E); parameter regions to the left of the lines give R<1. First we explore whether case isolation alone can control a SARS outbreak in a community and its hospital, as a function of the daily probability that a symptomatic individual in the community will be identified and isolated (h_c), and the factor by which isolation modifies the transmission rate (κ). We compare scenarios where hospital-wide transmission occurs at the same rate (η=1, Figure 2C) versus half the rate (η=0.5, Figure 2D) as in the general community. Unsurprisingly, the extent of measures required to control an outbreak is strongly dependent on R_0. If R_0~3 (as reported for SARS in Hong Kong and Singapore) and no general contact precautions are taken in the hospital (Figure 2C, black lines), we see that an outbreak can be controlled only if case isolation reduces transmission at least four-fold (κ<0.25), and the mean onset-to-hospitalization time is less than 3 days (h_c>0.3). More stringent infection control (lower κ) allows slightly slower hospitalization to be effective. When general contact precautions cut hospital transmission by half (Figure 2D), case isolation has a considerably greater chance of success, though rapid hospitalization is still imperative.

These results agree qualitatively with Riley et al (2003), who conclude that the observed reduction of mean onset-to-hospitalization time from 4.84 days to 3.67 days could not solely have controlled the Hong Kong outbreak (with their assumed value of κ=0.2). Different model structures complicate quantitative comparisons (we model HCWs explicitly, while Riley et al consider spatial coupling), but to test changes in R due case isolation we approximate their model by setting $h_h=h_c$ and $\rho=\eta=1$. Increasing h_c from 1/4.84 to 1/3.67 then reduces R by 11%, in precise agreement with the result obtained in the Hong Kong study.

Delays in initiating isolation of symptomatic individuals are shown by three contours of each color (delays increase from right to left). Such delays are relatively unimportant when R_0 is low but become critical as R_0 increases, since in higher-transmission settings symptomatic individuals are more likely to reproduce their infection in just a few days of unrestricted mixing, requiring disproportionate efforts on other control measures if isolation is delayed. For example, when R_0=5 and η=0.5 (Figure 2D, purple lines), if case isolation is initiated after the first day of symptoms then a six-fold reduction in transmission (κ~0.17) and 40% daily isolation probability (h_c=0.4) is sufficient to contain the outbreak (point marked *), whereas delaying isolation for two more days requires near-perfect isolation practices (κ<0.1, h_c>0.7) to assure containment. Note that preliminary evidence suggests variation in viral load throughout the symptomatic period (Peiris et al. 2003), which if correlated with infectiousness could affect our conclusions regarding the importance of immediate case isolation.

In all cases (Figure 2C, 2D), we identify a sharp threshold in the interaction between κ and h_c. For example, in an outbreak with R_0=2.5 and η=0.5 (Figure 2D, red lines), making improvements in isolation practices (decreasing κ) has little effect on R if current control measures place the system at point A, but show dramatic benefits if the system is at point B. Conversely, increasing h_c significantly boosts control from point A but has negligible effects from point B. This threshold arises because even if all cases are isolated immediately (h_c=1) the epidemic will not be contained unless κ is sufficiently low. Conversely, even if isolation stops transmission entirely (κ=0), the outbreak will not be contained unless a sufficient proportion of cases are isolated soon enough. The sharpness of this threshold arises in part because the proportion of individuals not isolated by the n^{th} day is $(1-h_c)^n$, where values of n~10-20 are pertinent because individuals with SARS often remain symptomatic for an extended period. The threshold softens as R_0 increases, since when transmission rates are higher the critical values of n are smaller (i.e. individuals must be isolated sooner, on average, in order to keep R<1). These findings highlight that pushing blindly to upgrade any given control measure may not advance the fight against SARS; with limited resources, the best approach is to identify where the current policy is failing (e.g. inadequate case identification versus ineffective isolation practices) and make targeted improvements.

Contact tracing and quarantine

For situations where modern case isolation facilities are not available, we explore the extent to which quarantine can compensate as a control measure (Figure 2E). For fixed case isolation probabilities, we investigate the tradeoff between q, the

daily probability that an incubating individual will be traced and quarantined, and κ, the degree to which case isolation reduces transmission. We assume transmission in quarantine is reduced by half, on average, because experiences in Singapore and elsewhere have indicated that many people are not compliant (Mandavilli 2003). For each value of R_0, we plot R=1 contours for two speeds of contact tracing (i.e. delays before quarantining begins).

The potential for quarantining to aid SARS containment increases markedly with R_0. In low-transmission settings (Figure 2E, dark blue lines) there is little difference between immediate quarantine and none at all—a small change in κ or η would be more effective than instituting a quarantine policy. In contrast, for higher R_0 quarantining can aid control substantially. In a setting with R_0=5 (Figure 2E, purple lines), in the absence of quarantining a value κ<0.06 is required to contain an outbreak. If daily quarantining probabilities are at least 0.3, however, then κ~0.25 is sufficient to achieve containment. Thus even a partially effective quarantine policy creates a significant opportunity to assure containment (i.e. to bring R below 1) where little existed before. In developing-world settings, implementing moderate quarantine programs may be far more tractable than attaining near-perfect case isolation.

Quarantining aids disease containment in two ways in our model. First, individuals in quarantine (E_m) are assumed to proceed immediately to case isolation (I_m) when they develop symptoms. Second, lower contact rates of quarantined individuals reduce transmission during the incubation period (only relevant when $\varepsilon \neq 0$). The broken lines in Figure 2E show cases where no transmission occurs during incubation (i.e. ε=0); the lines are shifted rightwards because this decreases R_0 by a small amount.

Sensitivity to quarantining rate (shown by the curvature of the lines) is only slightly diminished, though, indicating that the effect of quarantine is due primarily to rapid isolation once symptoms develop.

The maximum benefit of quarantining is realized when contact tracing begins on the first day following exposure (rightmost solid line of each color). Rapid gains are made as q increases from zero, but this effect saturates at relatively low daily probabilities. This saturation occurs because the effect of quarantine is due primarily to faster case isolation, so the important quantity is the proportion of cases traced before they progress to symptoms. The proportion traced by the n^{th} day is $1-(1-q)^n$, which for $n\sim5$ (the median incubation period for SARS) approaches 1 quickly as q increases. If contact tracing is delayed such that no individuals are quarantined until five days following exposure (leftmost lines), the contribution of quarantine is considerably reduced even if q is high. This follows because five days is the median incubation period, so that half of cases will have already developed symptoms before quarantining begins. Thus it is essential for contact tracing to begin quickly, even if coverage is initially low: tracing just a few exposed contacts quickly can have a large effect. As for case isolation, the impact of delays is greater for higher-transmission settings because unmanaged cases (in this case, the untraced individuals who become symptomatic) can do greater damage.

Note that in settings where individuals remain quarantined for some portion of the symptomatic period instead of entering isolation immediately, quarantining aids control efforts by reducing mixing rates during the early phase of the highly infectious

period. This contribution to containment is not included in our model, and would increase the impact of quarantine measures on epidemic growth

Contact precautions

Comparison of Figures 2C and 2D show the dramatic impact of hospital-wide contact precautions (η), suggesting that measures to reduce transmission among the high-risk HCW population may be a powerful complement to case management. In Figure 2F, we directly compare the individual influence of all four transmission-reduction parameters (κ, γ, ρ, and η) on R, under fixed probabilities of quarantine and isolation. Hospital-oriented measures (η and κ) are the most potent by far. Hospital-wide precautions (η) will always be stronger than specific case isolation measures (κ), since both factors contribute equally to reducing transmission by isolated patients while only η affects transmission by not-yet-identified HCW cases (see Figure 1D). Reduction in HCW contact with the community (ρ) has a weaker effect on R, and the effectiveness of quarantine (γ) plays a minimal role (though this will increase if transmission during incubation is higher, or in settings where individuals remain in quarantine through some of the highly infectious period).

Contact precautions in the hospital setting (η and κ) are thus critical to successfully controlling a SARS outbreak, due to the deliberate importation of highly-infectious symptomatic cases into hospitals. In the appendix we test the robustness of this conclusion to changes in case management scenarios and R_0. As expected, the role of η and κ in reducing R is diminished as hospitalization rates become very low. However, in every scenario considered the contribution to R due to poor contact

precautions in the hospital ($\eta \to 1$) is higher than that for any other failure of transmission control, particularly if screening of HCWs for symptoms is poor. Some degree of hospital-wide contact precautions is thus essential to combating a SARS outbreak.

Control strategies and delays in implementation

Having assessed the importance of various control measures alone or in pairs, we now consider the effects of integrated control strategies on SARS outbreaks. We treat a scenario with R_0=3, similar to outbreaks in Hong Kong and Singapore. The median and 50% confidence intervals (i.e. the 25th and 75th percentile values) of cumulative incidence indicate that such an epidemic is likely to spread rapidly through the population if uncontrolled (Figure 3A, black lines) . Control strategies emphasizing contact precautions (Figure 3A, green lines) or quarantine and isolation (Figure 3A, red lines) both reduce the effective reproductive number to R=1.5, thereby substantially slowing the epidemic's rate of growth and increasing the probability of containment. A combined strategy of contact precautions and case management measures reduces R below 1 (R=0.84 in this case—blue lines in Figure 3A), thereby leading to rapid containment of the outbreak in 85% of simulations. Considering the elements of the next-generation matrix (see figure caption), we see that control is finally achieved because simultaneous lowering of κ and η brought R_{ch} and R_{hh} below 1. In all cases R_{cc}<1, thus transmission involving the high-risk HCW pool is required to sustain the uncontrolled outbreaks.

We consider the repercussions of delaying implementation of these three control strategies (Figure 3B). In all cases, a prompt public health response is critical to containing a SARS outbreak. This is particularly true for strategies which do not reduce R below 1, since the possibility of stochastic fadeout (see Figure 2B) falls drastically if the disease has time to spread beyond the initial few cases. We also considered the impact of implementation delays on the most effective, combined strategy from Figure 3A, in scenarios where R_0 equals 2, 3 and 4 (Figure 3C). Again timing is critical, especially for the most-transmissible case when the effective R after control is greater than 1. In all cases in Figures 3B and C, a window of opportunity exists at the beginning of a SARS epidemic when a given control strategy has the greatest chance of success. This critical period is lengthened by more effective control strategies, lower-transmission settings, and, potentially, by a smaller hospital pool (see appendix). As implementation is delayed, the probability of containing the epidemic is reduced to levels obtained in the absence of control; for instance if R_0=3, the weaker strategies lose their possible impact on containment within a few weeks, and even the highly-effective combined strategy must be implemented within two months to show any gains over doing nothing. Note that this result pertains strictly to epidemic containment (as we define it), and is oriented toward minimizing overall morbidity and mortality. We emphasize that implementation of stringent control measures is all the more essential if containment has already been lost and a full-scale epidemic is in progress.

Preventing generalized community transmission

The preceding discussion has focused on reducing R and containing the outbreak without regard to the distribution of cases between the hospital and community. Preventing SARS from entering general circulation in the community is an important goal for at least two reasons. First, restricting the outbreak to close contacts of known cases and hospitals facilitates contact tracing and surveillance, and greatly reduces the probability of an uncontrolled epidemic. Second, as evidenced by the recent furor in Toronto, reports of generalized community transmission can have dire economic impacts.

We now consider measures directed at restricting transmission of SARS from the hospital back to the community, in particular the effect of precautions taken by HCWs to reduce their community contacts during off-duty time, as measured by the parameter ρ. In Figure 4A we plot the daily incidence of new cases, broken down by the pathway of transmission, for one realization of a stochastic epidemic in which HCWs behave normally when off-duty (ρ=1). The pie-chart inset shows the average proportions of infections along each pathway for 500 such epidemics. Note that h-to-c transmission—describing infections in the community caused by off-duty HCWs—accounts for 15% of all infections in this scenario, and that a significant number of cases arise in the community each day. In Figure 4B we plot results for an identical scenario, except that the contact of HCWs with community members is reduced ten-fold (i.e. ρ=0.1). This control measure notably decreases the growth rate of the epidemic (R=1.4 instead of 1.6, as in Figure 4A), and even more dramatically alters the patterns of spread: h-to-c transmission drops from 15% to just 2% of all infections. Furthermore

there is minimal ongoing *c*-to-*c* transmission—since R_{cc} is virtually unchanged from Figure 4A (see figure caption), we attribute this to the reduction in re-seeding from the HCW pool (note that R_{hc} drops from 0.23 to 0.02 due to the decrease in ρ).

Despite the relatively small contribution of ρ to the effective reproductive number (Figure 2F), our results indicate that reducing HCW-community mixing can play a critical role in preventing escape of SARS into the general population via the next-generation matrix element R_{hc}. These results obtain despite high case isolation probabilities for HCWs (h_c=0.9), which act independently of ρ to limit *h*-to-*c* transmission (since case-isolated HCWs no longer contact the community). Within-hospital contact precautions, however, were not assumed to be highly effective (η=0.5, κ=1). Clearly the importance of ρ diminishes when hospital conditions minimise infection of HCWs.

4. General discussion

Hospitals have been focal points of SARS transmission in all affected areas for which data are available. Our model examines a SARS outbreak in a community and its hospital, to identify tradeoffs and interactions between the limited suite of control measures available for a novel viral disease transmitted by casual contact. We consider a range of R_0 values, reflecting variation between cultural settings, but focus on scenarios with R_0~3 in keeping with best estimates for the Hong Kong (Riley et al 2003) and Singapore (Lipsitch et al 2003) outbreaks. Certain results hold true regardless of the precise value of R_0, and hence are relevant to SARS containment efforts

everywhere. These robust conclusions also provide guidance for future outbreaks of emerging pathogens, particularly those exhibiting a tendency for nosocomial spread.

Our analysis identifies healthcare workers (HCWs) as critical targets for control efforts. Their status is analogous to high-activity "core groups" in sexually-transmitted disease epidemics, but the deliberate importation of SARS cases from the background community adds a twist to this established paradigm. HCWs are thus exposed to a local prevalence much higher than that in the community-at-large, and as a result measures which reduce transmission within the hospital have the greatest impact on the reproductive number (R) of the epidemic. Hospital-wide precautions (η) have the strongest effect on R, followed by specific precautions for isolating SARS patients (κ).

This finding is bolstered by the detailed account of Dwosh et al. (2003) of a comprehensive and effective response by a community hospital near Toronto. A dedicated SARS ward was established (under negative air pressure to prevent aerosol spread to other parts of the hospital) and private rooms provided for each SARS patient (also under negative pressure); these measures correspond to reduced κ in our model. Intensive barrier and contact precautions were practiced by all hospital staff at all times (corresponding to $\eta<<1$). Hospital staff were screened at least twice daily for SARS symptoms, increasing h_h. Further measures included voluntary quarantine of all staff presumed exposed, restriction of visitor and patient numbers, and prevention of patient or staff transfers to other institutions. The hospital outbreak was contained, without a single further infection after contact precautions were imposed.

Such a thorough response is unattainable in many regions, particularly when facilities for respiratory isolation are not available. Fortunately measures which reduce

η are simple and inexpensive—masks, gowns and hand-washing significantly reduced transmission of SARS to HCWs in Hong Kong (Seto et al. 2003). For a broad range of scenarios our model indicates that high values of η (i.e. poor precautions) contribute more to epidemic growth than any other parameter, thus some degree of hospital-wide contact precautions is essential to combating a SARS outbreak. Whenever within-hospital measures are insufficient to stop infection of HCWs, however, it becomes critical to reduce leakage of the infection back into the community. Reducing contacts of off-duty HCWs with community members can accomplish this goal, but can produce the perverse effect that good HCW precautions lead to a higher proportion of SARS cases being HCWs. This results not from a higher incidence among HCWs, but from preventing infection from escaping back into the general community. HCWs in Hanoi effectively sealed themselves off from the outside world, resulting in the fastest containment of any significant SARS outbreak but also one in which 63% of cases were HCWs (Reilley et al 2003, Twu 2003).

Our model identifies minimum case management measures required to contain SARS outbreaks in different settings. For R_0 values reported for Hong Kong and Singapore we show that control is assured only if case isolation reduces transmission at least four-fold and the mean onset-to-hospitalization time is less than 3 days. There is a threshold in the interaction between hospitalization rates and isolation efficacy, beyond which further improvements contributed virtually nothing to containment. This highlights the need to understand the reasons why particular control strategies are failing before rushing to improve control in any way possible. Contact tracing and quarantine can compensate to some extent for inadequate isolation facilities, making an

increasingly significant contribution as R_0 rises. The impact of quarantine is due primarily to rapid isolation of cases once symptoms develop, and we show that it is essential for contact tracing to begin immediately, even if coverage is initially low, since tracing just a few exposed contacts quickly can have a large effect.

In general, our results indicate that delays in initiating quarantine or isolation undermine the effectiveness of other control measures, with increasing impact for greater R_0. More harmful still is delaying the implementation of control after emergence of the first case; this is an acknowledged hazard from earlier disease outbreaks (e.g. Keeling et al 2001). For particular control strategies, our model identifies critical windows of opportunity beyond which measures lose almost all ability to contribute to containment. The original SARS outbreak in Guangdong province, not officially acknowledged for over 5 months, serves as a tragic example of the hazards of delaying disease control efforts.

Epidemic modelers must always approximate the social structures of interest, and as with all models the approach presented here has potential shortcomings. First, within the hospital and community pools mixing is assumed to be random. This assumption ignores pockets of the population which do not share common contacts, but is reasonable if attention is restricted to low prevalence levels such as below 1% as in our analysis. Models with network structure tend to predict lower initial rates of spread, due to correlations in infection status that develop between neighbors, but they are better suited to diseases transmitted by intimate contact (such as needle-sharing or sex) or static hosts. When contacts are dynamic and transmission more casual, these correlations decay and system behavior approaches the random-mixing case (Keeling

1999). Household or network-structured models allow a direct treatment of contact tracing, rather than the *ad hoc* approach used here, but at the cost of additional parameters which are difficult to estimate.

Second, the observed number of secondary cases per index case of SARS has been very heterogeneous, and alternative modes of transmission have been postulated (including airborne, fomite, and fecal-oral spread (Wenzel & Edmond 2003, Riley et al 2003)). So-called superspreading events (SSEs), in which single individuals generated an extraordinary number of secondary cases, played an important role in the early evolution of several SARS outbreaks (Riley et al 2003, Leo et al 2003, Lipsitch et al 2003). Whether SSEs are rare epidemiological exceptions or represent the tail of a highly overdispersed distribution is subject to debate (Dye & Gay 2003). In our model, heterogeneity in secondary cases arises due to stochastic effects but SARS transmission is assumed to be a homogeneous process (in that the baseline transmission rate is represented by a single parameter, β). Lipsitch et al (2003) found that increasing variance in the distribution of secondary cases leads to reduced probability of disease invasion. Since we have not explicitly incorporated such heterogeneity in our model, our assessments of containment probabilities will be conservative to the extent that SSEs are a normal part of SARS epidemiology.

Third, the hospital pool is considered to include HCWs and SARS cases, but other patients are not modeled explicitly. Infection of other patients has played a significant role in some outbreaks, though it will be less important in hospitals which eliminate non-essential procedures while SARS remains a significant risk (e.g. Dwost et al 2003, Maunder et al. 2003), and for regions that have opened dedicated SARS

hospitals or wards. Future work on hospital-community SARS outbreaks could incorporate patient dynamics, and could also evaluate the effect of staff reductions (Maunder et al. 2003) or of mass quarantine of hospital staff following diagnosis of the first case (as reported by Dwost et al. 2003).

Some caution is required in identifying R_0 in our model with that obtained from incidence data for particular outbreaks. As discussed above, reproductive numbers derived from data inevitably incorporate some degree of control due to routine healthcare practices. We calculate R_0 from its formal definition, however, as the expected number of secondary cases in the absence of control measures (i.e. without hospitalization or any contact precautions). While it is uncertain whether routine health practices help or hinder the spread of SARS, we suspect that estimates of R_0 under our strict definition would be somewhat higher than those reported for Hong Kong and Singapore (which incorporate some measures, such as hospitalization, from the outset). Of course this will depend on the details of non-specific healthcare practices in each setting, on assumptions regarding their effect on SARS spread, and on how R_0 is calculated (particularly the treatment of SSEs).

The most successful examples of quickly controlling SARS outbreaks (e.g. Hanoi and Singapore) show common features of stringent within-hospital contact precautions, and success in preventing leakage of infection from hospitals back to the general community. The continuing difficulty that Toronto health officials face in containing their SARS outbreak, meanwhile, testifies to the disease's potential for spread despite the implementation of intensive control strategies. Unprecedented human mobility means that emerging infectious diseases can rapidly impact public

health around the world. To contain outbreaks of SARS, or other pathogens for which vaccines or treatment are not available, requires aggressive case management measures complemented by contact precautions to reduce transmission in healthcare settings.

Table 1. Summary of transmission and case management parameters, including baseline values used throughout the study and the three control strategies depicted in Figure 3.

Parameter	**Symbol**	**Range examined**	**Fig 3 (1)**	**Fig 3 (2)**	**Fig 3 (3)**
Baseline transmission rate (day^{-1})	β	0.08-0.26 (R_0=1.5-5)	0.15 (R_0=3)	0.15 (R_0=3)	0.15 (R_0=3)
Factors modifying transmission rate, due to:					
Pre-symptomatic transmission	ε	0-0.1	0.1	0.1	0.1
Hospital-wide contact precautions	η	0-1	0.5	0.9	0.5
Reduced HCW-community mixing	ρ	0-1	0.5	1	0.5
Case isolation	κ	0-1	1	0.5	0.5
Quarantine	γ	0-1	0.5	0.5	0.5
Daily probability of:					
Quarantining of E_c individual	q	0-1	0	0.5	0.5
Isolation of I_c individual	h_c	0-1	0.3	0.9	0.9
Isolation of I_h individual	h_h	0.9	0.9	0.9	0.9

Figure captions

Figure 1

Flow diagram of the transmission dynamics of a SARS epidemic within a hospital coupled to that in a community. **(A)** We modeled SARS transmission as an SEIR process (*S*: susceptible, *E*: incubating, *I*: symptomatic, *R*: removed) structured into a healthcare worker (HCW) core group (subscript *h*), a community-at-large (subscript *c*), and a case managed group (subscript *m*) of quarantined (E_m) and isolated (I_m) individuals. **(B)** Incubating individuals in all three groups (E_x, where x=*c*,*h*,*m*) were further structured into ten disease-age classes, with daily probabilities p_i of progressing to the symptomatic phase. Values of p_i were interpolated linearly between p_1=0 and p_{10}=1, yielding a distribution of incubation periods consistent with data (see appendix). **(C)** Symptomatic individuals in all three groups (I_x, where x=*c*,*h*,*m*) were structured into two initial disease-age and three subsequent disease-stage classes (in which individuals have a probability *r* of moving to the next stage each day). Individuals leaving the final symptomatic class go to R_c or R_h, according to their group of origin. Independent of the disease progression described in (B) and (C), individuals can enter case management with daily probabilities *q* for quarantine, or h_c and h_h for isolation (in the community and HCW groups respectively). Individuals must already be in class I_c or I_h to be isolated, so the soonest an un-quarantined individual can be isolated is after the first day of symptoms; individuals in quarantine are assumed to move directly into isolation when symptoms develop. **(D)** The transmission hazard rates for susceptible individuals S_j are denoted by λ_j (j=*c*,*h*), and depend on weighted contributions from community and HCW sources as described in the text (and Table 1). The discrete-time stochastic

formulation of our model allows for the possibility of multiple infectious contacts within a timestep, so for a susceptible individual subject to total hazard rate λ_j the probability of infection on a given day is $1-exp(-\lambda_j)$. (Note that the units of β are day^{-1}.) We assume density-independent contact rates and random mixing within each pool, so hazard rates of infection are dependent on the transmission rate for each infectious class multiplied by the proportion of the population in that class. Specifically, defining the effective number of individuals in the hospital mixing pool as $N_h=S_h+E_h+I_h+R_h+I_m$ and in the community mixing pool as $N_c=S_c+E_c+I_c+R_c+\rho(S_h+E_h+I_h+R_h)$, the total hazard rates are $\lambda_c=[\beta(I_c+\varepsilon E_c)+\rho\beta(I_h+\varepsilon E_h)+\gamma\beta\varepsilon E_m]/N_c$ and $\lambda_h=\rho\lambda_c+\eta\beta(I_h+\varepsilon E_h+\kappa I_m)/N_h$. In simulations, the number of infection events in each timestep is determined by random draws from binomial(S_j, $1-exp(-\lambda_j)$) distributions ($j=c,h$). Equations describing the model are given in the appendix.

Figure 2

(A) Sample output from the model, showing cumulative numbers of cases. Five realizations of the stochastic model are shown for three values of R, to highlight variability in outcomes and the increased probability of fadeout for lower R. Several epidemics died out immediately and cannot be resolved: 1 for R=2, and 2 each for R=1.6 and R=1.2. **(B)** Probability of epidemic containment (as defined in the text) as a function of R, for a population of 100,000 with a single initial case. We set ε=0.1, and β was varied to give the desired R values, with no control measures imposed. Probabilities were calculated from 100 runs per R value. **(C)-(E)** Threshold control policies for containment of the epidemic. Lines show R=1 contours for scenarios where

R_0=1.5 (green), 2 (blue), 2.5 (red), 3 (black), 4 (light blue) and 5 (purple); parameter regions to the left of the lines give R<1. Not all cases appear because some are off the scale. Panels **C** and **D** show, for η=1 and η=0.5 respectively, the effect of varying h_c (the daily probability that symptomatic SARS cases in the community will be isolated) and κ (the modification to transmission due to case isolation procedures) on the threshold where R=1. From right to left, three lines of each color show the effects of increasing delays in case isolation (i.e. each symptomatic individual has no possibility of being isolated for 1, 2, or 3 days respectively, but a constant daily probability (h_c) thereafter). Points in (D) marked A, B, and * are described in the text. We assume no quarantining (q=0) and a fixed strategy of case isolation of symptomatic hospital workers (h_h=0.9) starting after their first day of symptoms. Also ρ=1, ε = 0.1. Panel **E** shows the extent to which contact tracing and quarantine can substitute for imperfect case isolation. Here η=0.5, ρ=1, γ=0.5, h_c=0.3 and h_h=0.9, so the case isolation strategy is fixed (and assumed to commence after the first day of symptoms), but the degree to which transmission is reduced by isolation (κ) varies. From right to left, two lines of each color represent 1-day and 5-day delays in contact tracing before quarantining begins. Solid lines show the case ε = 0.1, when transmission can occur during the incubation period. The dashed lines show the case ε=0 (note that R_0=2.44, 2.92 and 3.90, rather than 2.5, 3 and 4). **(F)** Sensitivity of effective reproductive number R to the four transmission-control parameters. In all cases R_0=3, ε=0.1, q=0.5, h_c=0.3, and h_h=0.9; all parameters κ, γ, ρ, η were set to 0.5, then varied one at a time.

Figure 3

(A) Increase in cumulative cases over 200 days, for uncontrolled outbreak compared with three control strategies. For each day we plot the median value (solid line) and 25th and 75th percentiles (dashed lines) of 500 simulations. Probabilities of containment (and 95% confidence intervals) for each scenario are shown in the legend. Control strategies are implemented 28 days after the first symptomatic case. In all instances ε =0.1 and R_0=3. Strategy 1 (Contact precautions): h_c=0.3, h_h=0.9, η=0.5, ρ =0.5, κ=1, γ=1, q=0, yielding R=1.5 once implemented ($\{R_{cc},R_{ch},R_{hc},R_{hh}\}$=(0.56, 1.04, 0.12, 1.36)). Strategy 2 (Quarantine and isolation): h_c=0.9, h_h=0.9, η=0.9, ρ =1, κ=0.5, γ=0.5, q=0.5, yielding R=1.5 ($\{R_{cc},R_{ch},R_{hc},R_{hh}\}$=(0.06, 1.12, 0.23, 1.34)). Strategy 3 (Combined measures): h_c=0.9, h_h=0.9, η=0.5, ρ =0.5, κ=0.5, γ=0.5, q=0.5, yielding R=0.84 ($\{R_{cc},R_{ch},R_{hc},R_{hh}\}$= (0.06, 0.63, 0.12, 0.74)). **(B)** Probability of containing the outbreak as a function of increasing delay in implementing the control strategies, for the three strategies from Figure 3A. **(C)** Probability of containment versus delay in implementation for strategy 3 from Figure 3A, for R_0 values of 2, 3 and 4, which yield effective R values of 0.56, 0.84, and 1.12 after control is implemented. Horizontal lines correspond to the probability of fadeout in the absence of control (determined through simulation). In 3B and 3C, each probability (p_{est}) is estimated from 500 simulations; error bars represent 95% confidence intervals on p_{est}, given by $2\times[p_{est}(1-p_{est})/500]^{1/2}$.

Figure 4

Importance of HCW mixing restrictions in preventing SARS spread to the community. **(A)** and **(B)** show two stochastic epidemics with identical disease parameters and

control measures, differing only in HCW-community mixing precautions (ρ=1 in A and ρ=0.1 in B). Daily incidence is shown, broken down by route of transmission within or between the hospital and community pools. Inset, pie-charts show average contributions of the different routes of infection for 500 stochastic simulations of each epidemic (standard errors for these proportions were estimated by jackknifing the simulation results, but in all cases were less than one percentage point). Note that *c*-to-*h* transmission includes hospitalized community-members infecting the HCWs caring for them. R_0=3 in both cases, and other parameters are from Scenario 1 of Figure 3: ε=0.1, κ=1, η=0.5. q=0, h_c=0.3, h_h=0.9, yielding R=1.60 in (A) ($\{R_{cc},R_{ch},R_{hc},R_{hh}\}$=(0.55, 1.05, 0.23, 1.37)) and R=1.39 in (B) ($\{R_{cc},R_{ch},R_{hc},R_{hh}\}$= (0.57, 1.04, 0.02, 1.36)). The control policy is implemented 14 days into the outbreak.

Figure 1

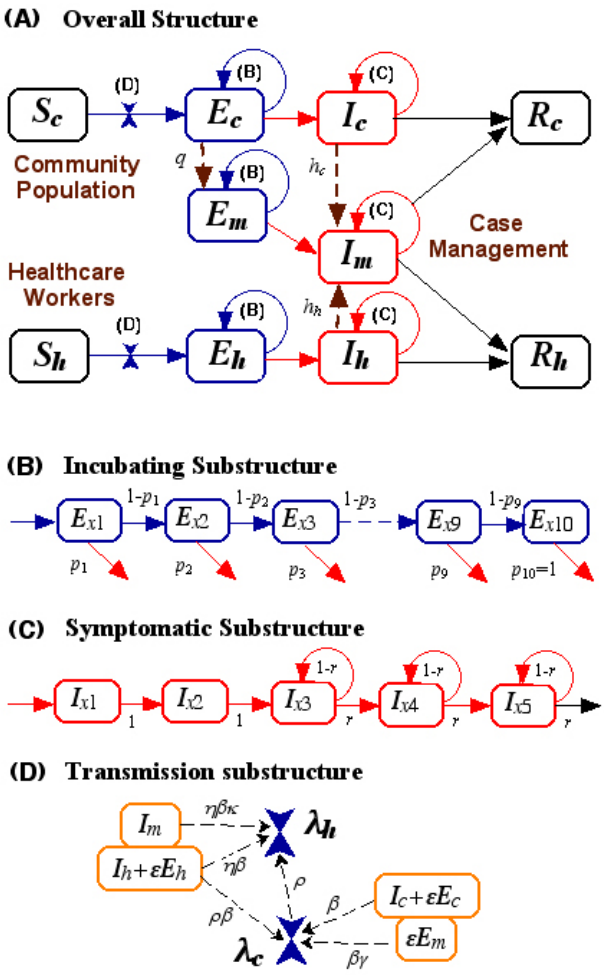

(A) Overall Structure
Community Population
Healthcare Workers
Case Management
(B) Incubating Substructure
(C) Symptomatic Substructure
(D) Transmission substructure

Figure 2

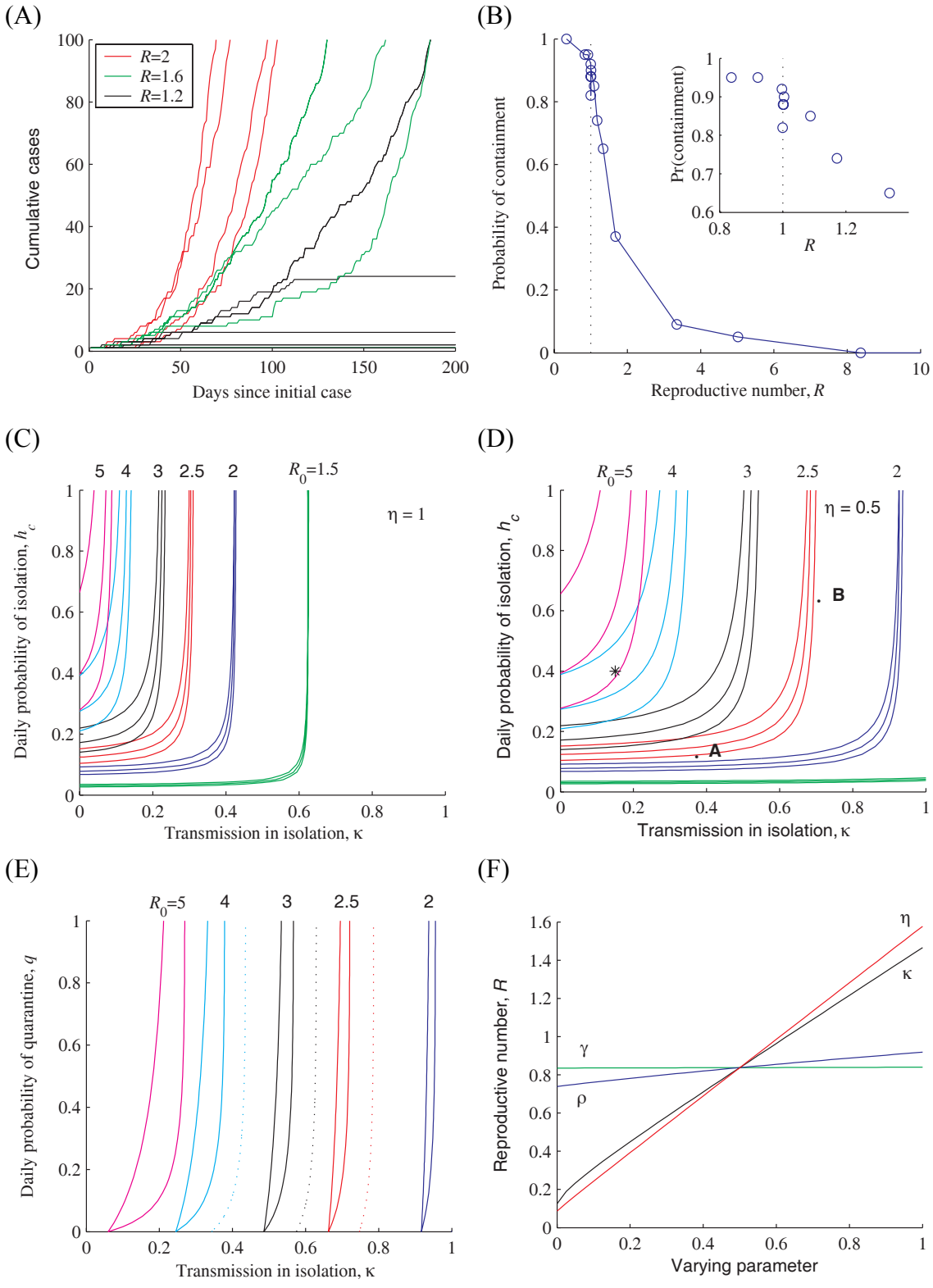
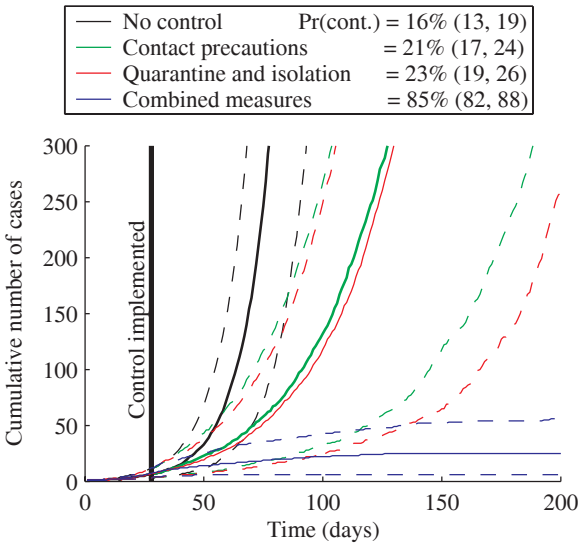
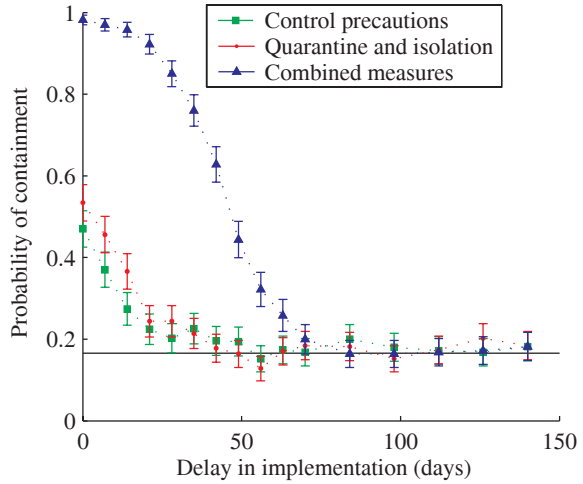

(A)
Cumulative cases
Days since initial case
R=2
R=1.6
R=1.2
(B)
Probability of containment
Reproductive number, R
Pr(containment)
R
(C)
Daily probability of isolation, h_c
Transmission in isolation, κ
5 4 3 2.5 2
R_0=1.5
η = 1
(D)
Daily probability of isolation, h_c
Transmission in isolation, κ
R_0=5 4 3 2.5 2
η = 0.5
B
A
(E)
Daily probability of quarantine, q
Transmission in isolation, κ
R_0=5 4 3 2.5 2
(F)
Reproductive number, R
Varying parameter
η
κ
γ
ρ

Figure 3
(A)


No control Pr(cont.) = 16% (13, 19)
Contact precautions = 21% (17, 24)
Quarantine and isolation = 23% (19, 26)
Combined measures = 85% (82, 88)
Control implemented
Cumulative number of cases
Time (days)
0
50
100
150
200
250
300

(B)


Control precautions
Quarantine and isolation
Combined measures
Probability of containment
Delay in implementation (days)
0
0.2
0.4
0.6
0.8
1
50
100
150

(C)

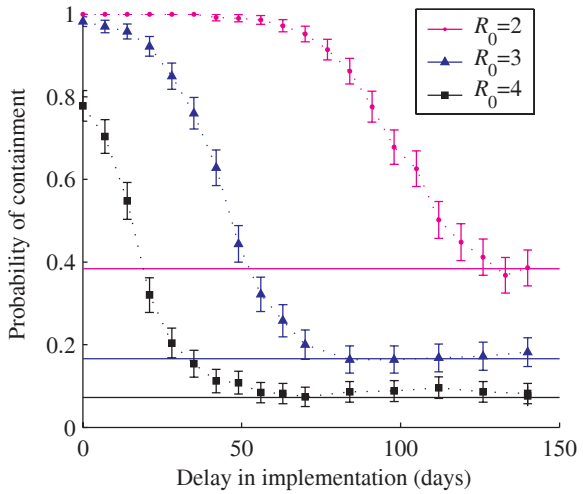
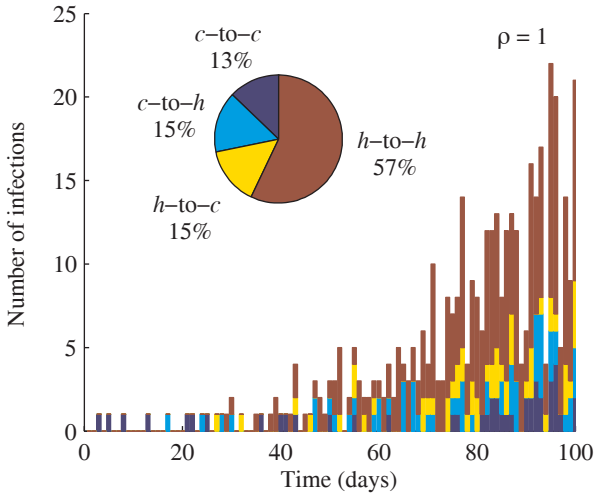
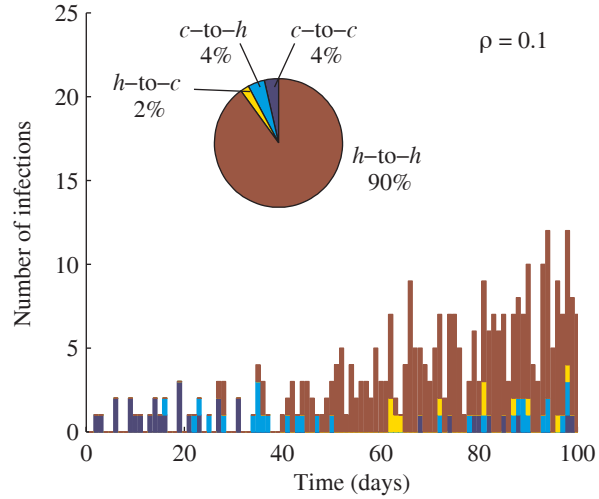

R_0=2
R_0=3
R_0=4
Probability of containment
Delay in implementation (days)
0
0.2
0.4
0.6
0.8
1
50
100
150

Figure 4
(A)


ρ = 1
c–to–c
13%
c–to–h
15%
h–to–h
57%
h–to–c
15%
Number of infections
Time (days)
0
5
10
15
20
25
0
20
40
60
80
100

(B)


ρ = 0.1
c–to–h
4%
c–to–c
4%
h–to–c
2%
h–to–h
90%
Number of infections
Time (days)
0
5
10
15
20
25
0
20
40
60
80
100

Appendix

Sensitivity to population size

We tested the sensitivity of key model results to both absolute and relative changes in pool sizes. Figure S1 shows results obtained when both pools are reduced ten-fold in size (i.e. for a HCW pool of 300 individuals, and community pool of 10,000). Comparing these figures to those in the main text, we see that changes in system scale do not qualitatively alter our findings. This is not surprising, since we treat contact rates as density-independent and restrict our attention to the invasion phase when overall prevalence is less than 1%.

It is less clear whether our results will be sensitive to changes in the relative size of the two pools, since this will alter the weighting of different transmission pathways. In Figure S2 we present the same analyses when the HCW pool contains 1000 individuals (compared to 3000 throughout the main text), and community pool is still 100,000 individuals. Results pertaining to the reproductive number (Figures S2A-B) are not significantly changed, again due to our assumption of density-independent contact rates. When the evolving epidemic is simulated, though, slight differences emerge. A smaller HCW pool seems to slightly extend the window of time within which the combined control strategy contributes to outbreak containment (Figure S2C), perhaps due to slower initial spread. This effect is subtle but persists in all our simulations. The possibility that smaller hospital size reduces the risk of outbreaks is intriguing and has implications for health policy, and merits further investigation. In Figure S2D we see some changes in proportional routes of transmission, but the

essential result remains that reducing HCW-community contacts can prevent leakage of the infection from the hospital.

Robustness of transmission-reduction results

A major finding of this study is that hospital-oriented contact precautions, such as wearing masks and gowns at all times and respiratory isolation of identified patients, are the most potent measures for combating an incipient SARS outbreak. Figures S1B and S2B show that this conclusion is robust to absolute and relative changes in pool sizes. We now explore the sensitivity of this result to different case management scenarios and R_0 values, by plotting analogues of Figure 2F to show the effect of each transmission-reduction parameter on R.

We first consider a scenario with no quarantining (Figure S3A), which leads to a greater proportion of symptomatic individuals spending their initial days of symptoms mixing freely with the community. This reduces the contribution of hospital-based transmission to R, and accordingly we see a smaller relative contribution of η and κ to determining the effective reproductive number. Reinforcing this point, a scenario with less efficient case isolation and no quarantining (Figure S3B) exhibits still weaker dependence of R on the values of η and κ, and thus greater relative sensitivity to ρ. The three measures are almost equivalent as the parameters approach zero—we see that stopping HCW-community transmission ($\rho \to 0$) has a roughly equal effect to perfect case isolation ($\kappa \to 0$) and almost as great an effect as eliminating within-hospital transmission entirely ($\eta \to 0$). Strikingly, though, note that the cost of poor hospital-wide contact precautions ($\eta \to 1$) is much greater now that the rate of isolating

symptomatic HCWs is low. Indeed, the adverse effect of $\eta \to 1$ is always higher than any other failure of transmission-control measures. Some degree of hospital-wide contact precautions is thus essential to combating a SARS outbreak.

Finally, considerating the original case management strategy but raising R_0 to 5 (Figure S3C) shows that the overall transmissibility acts only to scale the lines from Figure 2F, but does not alter their relation to one another.

Model equations

For ease of presentation, the following equations show a deterministic analogue of our model. All terms shown here as products of a probability and a state variable are generated in our simulations by drawing binomial random variables. The community pool is described as follows, where all variables and parameters are as described in Figure 1 of the main text:

$$S_c(t+1) = \exp\big(-\lambda_c(t)\big)S_c(t)$$
$$E_{c,1}(t+1) = \big[1-\exp\big(-\lambda_c(t)\big)\big]S_c(t)$$
$$E_{c,i}(t+1) = (1-p_{i-1})(1-q_{i-1})E_{c,i-1}(t) \qquad i = 2,...,10$$
$$I_{c,1}(t+1) = \sum_{i=1}^{10} p_i(1-q_i)E_{c,i}(t)$$
$$I_{c,2}(t+1) = (1-h_{c,1})I_{c,1}(t)$$
$$I_{c,3}(t+1) = (1-h_{c,2})I_{c,2}(t) + (1-r)(1-h_{c,3})I_{c,3}(t)$$
$$I_{c,j}(t+1) = r(1-h_{c,j-1})I_{c,j-1}(t) + (1-r)(1-h_{c,j})I_{c,j}(t) \quad j = 4,5$$
$$R_c(t+1) = R_c(t) + rI_{c,5}(t) + \big[rI_{m,5}(t)\big]^*$$

Daily probabilities of quarantine (q_i) or hospitalization ($h_{c,i}$) are subscripted by i because they can vary between subcompartments (in the analysis presented here they vary only between 0 and a fixed value, to describe delays in contact tracing or case identification). The final term in the $R_c(t+1)$ equation is marked with an asterisk because only those

individuals in $I_{m,5}$ who were originally from the community pool (i.e. community members who have been hospitalized) move to the R_c pool upon their recovery. Individuals in $I_{m,5}$ who began in the HCW pool progress to R_h upon recovery (indicated below with another asterisk). The equations for the HCW pool are:

$$S_h(t+1) = \exp\big(-\lambda_h(t)\big)S_h(t)$$
$$E_{h,1}(t+1) = \big[1-\exp\big(-\lambda_h(t)\big)\big]S_h(t)$$
$$E_{h,i}(t+1) = (1-p_{i-1})E_{h,i-1}(t) \qquad i = 2,...,10$$
$$I_{h,1}(t+1) = \sum_{i=1}^{10} p_i(1-q_i)E_{h,i}(t)$$
$$I_{h,2}(t+1) = (1-h_{h,1})I_{h,1}(t)$$
$$I_{h,3}(t+1) = (1-h_{h,2})I_{h,2}(t) + (1-r)(1-h_{h,3})I_{h,3}(t)$$
$$I_{h,j}(t+1) = r(1-h_{h,j-1})I_{h,j-1}(t) + (1-r)(1-h_{h,j})I_{h,j}(t) \quad j = 4,5$$
$$R_h(t+1) = R_h(t) + rI_{h,5}(t) + \big[rI_{m,5}(t)\big]^*$$

As described in the caption of Figure 1 (main text), the total hazard rates are $\lambda_c=[\beta(I_c+\varepsilon E_c)+ \rho\beta(I_h+\varepsilon E_h)+\gamma\beta\varepsilon E_m]/N_c$ and $\lambda_h=\rho\lambda_c+\eta\beta(I_h+\varepsilon E_h+\kappa I_m)/N_h$, where E_j and I_j represent sums over all sub-compartments in the incubating and symptomatic classes for pool j. The effective number of individuals in the hospital mixing pool is $N_h=S_h+E_h+I_h+R_h+I_m$, and in the community mixing pool is $N_c=S_c+E_c+I_c+R_c+\rho(S_h+E_h+I_h+R_h)$. In simulations, the number of infection events in each timestep is determined by random draws from binomial(S_j, $1-exp(-\lambda_j)$) distributions ($j=c,h$).

Finally, the equations describing the case-managed pool (quarantined and case-isolated individuals) are as follows:

$$E_{m,i}(t+1) = (1-p_{i-1})\left[q_{i-1}E_{c,i-1}(t) + E_{m,i-1}(t)\right] \qquad i = 2,\ldots,10$$

$$I_{m,1}(t+1) = \sum_{i=1}^{10} p_i\left[q_i E_{c,i}(t) + E_{m,i}(t)\right]$$

$$I_{m,2}(t+1) = h_{c,1}I_{c,1}(t) + h_{h,1}I_{h,1}(t) + I_{m,1}(t)$$

$$I_{m,3}(t+1) = h_{c,2}I_{c,2}(t) + h_{h,2}I_{h,2}(t) + I_{m,2}(t) + (1-r)\left[h_{c,3}I_{c,3}(t) + h_{h,3}I_{h,3}(t) + I_{m,1}(t)\right]$$

$$I_{m,j}(t+1) = r\left[h_{c,j-1}I_{c,j-1}(t) + h_{h,j-1}I_{h,j-1}(t) + I_{m,j-1}(t)\right] + (1-r)\left[h_{c,j}I_{c,j}(t) + h_{h,j}I_{h,j}(t) + I_{m,j}(t)\right] \quad j = 4,5$$

Calculation of the reproductive number

The progression of each infected individual through incubating and symptomatic stages of the disease, and possibly through case management stages, can be described by a stochastic transition matrix. When the removed state is included, the infectious lifetime of each individual can be represented as an absorbing Markov chain (where "absorption" corresponds to the end of the infectious period). For a given set of transition probabilities (i.e. disease progression parameters and probabilities of entering case management from each disease stage), the expected residence time in each pre-absorption stage can be calculated from the fundamental matrix of the Markov chain (Caswell 2000).

Since case management probabilities may vary between the community and hospital pools, we define d_j (for $j=c$ or h) as a vector of expected residence times in the states (E_j, I_j, E_m, I_m), i.e. the length of time a "typical" individual infected in pool j will spend in each of those disease classes. We the define b_{jk} as vectors of transmission rates from pool j to pool k for each disease state. In particular, from the above description we have $b_{cc}=(\varepsilon\beta/N_c, \beta/N_c, \gamma\varepsilon\beta/N_c, 0)$, $b_{ch}=(\rho\varepsilon\beta/N_c, \rho\beta/N_c, \rho\gamma\varepsilon\beta/N_c, \kappa\eta\beta/N_h)$, $b_{hc}=(\rho\varepsilon\beta/N_c, \rho\beta/N_c, \gamma\varepsilon\beta/N_c, 0)$, and $b_{hh}=(\rho^2\varepsilon\beta/N_c+\eta\varepsilon\beta/N_h, \rho^2\beta/N_c+\eta\beta/N_h, \rho\gamma\varepsilon\beta/N_c, \kappa\eta\beta/N_h)$.

The two terms in the first two elements of b_{hh} represent community and workplace exposure risks for healthcare workers, respectively. The factors of ρ^2 reflect that community transmission between HCWs depends on the community-contact precautions of both HCWs.

For a susceptible individual in pool k, the total hazard of infection due to the index case is thus $\lambda_{jk}=d_j \cdot b_{jk}$, so the probability of exposure is $1-\exp(-\lambda_{jk})$. If there are S_k susceptibles in pool k, then the expected number of secondary infections in pool k due to an index case who is infected in pool j is $R_{jk}=[1-\exp(-\lambda_{jk})]S_k$. We then define the next-generation matrix:

$$\mathbf{R} = \begin{bmatrix} R_{cc} & R_{ch} \\ R_{hc} & R_{hh} \end{bmatrix}$$

If **R** is primitive, then its dominant eigenvalue is the reproductive number for the entire system (Diekmann & Heesterbeek 2000). When the population is entirely susceptible and no control measures are in place this is the basic reproductive number, R_0; otherwise it is the effective reproductive number R. The eigenvalues of **R** can be expressed:

$$\lambda_{1,2} = \tfrac{1}{2}\left\{\left(R_{cc}+R_{hh}\right) \pm \sqrt{\left(R_{cc}+R_{hh}\right)^2 - 4\left(R_{cc}R_{hh}-R_{ch}R_{hc}\right)}\right\}$$

Because all elements of **R** are non-negative, both eigenvalues are real and positive, and without loss of generality we have:

$$\begin{aligned} R_0 &= \tfrac{1}{2}\left\{\left(R_{cc}+R_{hh}\right) \pm \sqrt{\left(R_{cc}+R_{hh}\right)^2 - 4\left(R_{cc}R_{hh}-R_{ch}R_{hc}\right)}\right\} \\ &= \tfrac{1}{2}\left\{\left(R_{cc}+R_{hh}\right) + \sqrt{\left(R_{cc}-R_{hh}\right)^2 + 4R_{ch}R_{hc}}\right\} \end{aligned}$$

Figure 2A of the main text shows the probability of epidemic containment as a function of the reproductive number, which displays the qualitative behavior expected

for a stochastic epidemic: the probability is nearly one for $R<1$, then diminishes as R increases (but remains significantly greater than zero up to $R\sim5$).

To build intuition about the next-generation matrix and reproductive number, we present an analytic treatment of a simplified version of our model. We assume that no transmission occurs during the incubation period ($\varepsilon=0$), that multiple exposures within a timestep can be neglected, and that initially $S_k \approx N_k$. The elements of the next-generation matrix are now given by $R_{jk}=d_j \cdot b_{jk} N_k$. We define $d_{I,j}$ and $d_{M,j}$ as the expected residence times in the I_j and I_m states for an individual infected in pool j (for $j=c$ or h). Finally, we re-scale time such that $\beta=1$, and we can write down the elements of the next-generation matrix which emphasize the dependence on the control parameters ρ, κ, and η:

$$\mathbf{R} = \begin{bmatrix} R_{cc} & R_{ch} \\ R_{hc} & R_{hh} \end{bmatrix} = \begin{bmatrix} d_{I,c} & \rho d_{I,c} + \kappa\eta d_{M,c} \\ \rho d_{I,h} & \rho^2 d_{I,h} + \eta d_{I,h} + \kappa\eta d_{M,h} \end{bmatrix}$$

These expressions show clearly how the control measures represented by ρ, κ, and η affect the different pathways of transmission in our model. Recalling that smaller values of these parameters represent more effective measures, we see directly how the four elements of the next-generation matrix can be diminished by control efforts. Note also that the total expected duration of the symptomatic period will be equal in the two pools (i.e. $d_{I,c}+d_{M,c}=d_{I,h}+d_{M,h}$=constant), and is determined by the natural history of the disease. The durations of unmanaged mixing ($d_{I,c}$ and $d_{I,h}$) will be diminished by increasing efforts in contact tracing and quarantine (q) and case identification and isolation (h_c and h_h). The precise relationships for our model are complex, but roughly speaking we have $d_{I,c} \propto 1/h_c$ and $d_{I,h} \propto 1/h_h$, and also $d_{I,c} \propto 1/q$ since quarantined individuals go directly into case isolation.

Finally, from the expression given above we can write out the effective reproductive number of the system in terms of our control parameters:

$$R = \tfrac{1}{2}\left\{d_{I,c} + \rho^2 d_{I,h} + \eta d_{I,h} + \kappa\eta d_{M,h} + \ldots \right.$$
$$\left. \sqrt{\left(\rho^2 d_{I,h} + \eta d_{I,h} + \kappa\eta d_{M,h} - d_{I,c}\right)^2 + 4\rho d_{I,h}\left(\rho\, d_{I,c} + \kappa\eta d_{M,c}\right)}\right\}$$

Incubation and symptomatic periods

The incubation period is modeled with ten subcompartments as shown in Figure 1B of the main text. Each sub-compartment represents one day, and an individual in their i^{th} day since infection has a probability p_i of progressing to the symptomatic phase of the disease. The number of sub-compartments and values of p_i were chosen to be consistent with clinical data from 42 patients in Toronto with a single known contact with a SARS case. For these cases, the mean incubation period was 5 days, with a median of 4 days and a range from 2 to 10 days (Health Canada 2003); similar numbers are reported for 21 point-exposure cases in Singapore (Leo et al. 2003). We selected the most parsimonious model which was consistent with these data: 10 subcompartments with p_i interpolated linearly from p_1=0 to p_{10}=1. Figure S4A shows the distribution of incubation periods obtained from this model, which has a mean period of 4.5 days, a median period of 4 days, and a range from 2 to 10 days. Other researchers have presented a distribution of incubation periods which includes longer durations (Donnelly et al. 2003), but experts assembled by the World Health Organization continue to assert a maximum incubation period of 10 days (World Health Organization 2003b).

The symptomatic period is modeled with two disease-age subcompartments and three disease-stage subcompartments. After each day individuals automatically

progress through the age sub-compartments, and progress through the stage subcompartments with probability r. We include the initial disease-age subcompartments to allow assessment of the importance of beginning case isolation following day 1, 2 or 3 of symptoms. We assume that individuals are symptomatic for at least 5 days. From clinical reports of 23 patient histories we estimated that the distribution of symptomatic period has a mean of 16.2 days (with standard deviation of 7.9 days) and a median 16 days (Poutanen et al. 2003, Tsang et al. 2003). Figure S4B shows the distribution of symptomatic periods obtained from our model (with r=0.21), which has a mean period of 16.3 days, a median period of 15 days, and a standard deviation of 7.3 days.

While our modeled distribution is roughly consistent with data, we note that estimation of the symptomatic period poses a difficult challenge. We are seeking to capture the period of high infectiousness (which we call the symptomatic period to distinguish it from the incubation period, during which we assume individuals may be slightly infectious), but this is difficult to gauge because infectiousness is not readily observable. Our estimated symptomatic period—or highly infectious period—falls between those used in the two first modeling analyses of SARS outbreaks. Riley et al (2003) use hospitalization periods as a surrogate, and present a range of mean symptomatic periods from 27 to 41 days. (These include a symptomatic, not-yet-hospitalized period with mean duration of 3.67-4.84 days, and a symptomatic, hospitalized period with mean duration of 23.5 or 35.9 days depending on clinical outcome. Transmission by hospitalized individuals is reduced by a factor of 0.2, analogous to our κ.) Lipsitch et al (2003) do not model the symptomatic period directly

but instead assume an "average duration of infectiousness" of 5 days (range: 1-5 days). This is markedly shorter than the symptomatic periods used in our model (and that of Riley et al), but the difference results from their assumption that case isolation is absolutely effective, so an individual's "infectious period" lasts only until he or she is hospitalized. In contrast, our approach is to keep the biological phenomenon of infectiousness separate from the control-mediated phenomenon of transmission, leading to a longer total symptomatic period with transmission weighted by control parameters depending on case management practices.

Our model can still be consistent with the serial interval data presented by Lipsitch et al. (The serial interval is the time from onset of symptoms in an index case to onset of symptoms in a subsequent case infected by the index case. If the transmission rate is constant and the population is well-mixed, this equals the sum of the mean incubation period and the mean infectious period. The serial interval for SARS in Singapore before full-scale control policies were implemented was 10 days—subtracting the mean incubation period of 5 days yields the estimated 5-day infectious period.) Most simply, an exponentially-distributed period of uniform infectiousness with a mean duration of 5 days (as modeled by Lipsitch et al) could be approximated in our model by setting h_c=0.2 and κ=0, though in our model the tail of the distribution would be truncated by disease recovery. A more likely depiction of events in Singapore would be a higher hospitalization rate and non-zero κ, such that the weighted mean of all infectious periods (before and after case isolation) was 5 days. By separating the biological and control-mediated aspects of transmission, our model naturally portrays this or any other control scenario.

We therefore wish to characterize the natural history of the disease accurately. The duration of hospitalization is a plausible surrogate for the symptomatic period, but for a disease as virulent as SARS it is likely to be an overestimate, since patients must recover from severe lung damage and are not discharged from hospital until several days after all symptoms are resolved (Lee et al 2003). The most direct measurement of SARS infectious periods are the viral load measurements of Peiris et al (2003), which show that mean viremia (for 75 patients) peaks roughly 10 days after onset of symptoms, and after 15 days has dropped below its level after 5 days of symptoms. This is attributed to onset of IgG seroconversion, which begins as early as 10 days after onset of symptoms (with mean of 20 days).

These results indicate that symptomatic periods in our model, as shown in Figure S4B, probably characterize the period of high infectiousness quite adequately. Should there be any inaccuracies, our strategy of considering scenarios with different values of R_0 would largely buffer the impact on our results, since reproductive numbers estimated for particular outbreaks can be compared to model epidemics with the same net growth rate. This would entail a slight skew in parameter values: for instance, if we had underestimated the duration of infectiousness, for each R_0 scenario we would overestimate the baseline transmission rate, β. Simulations would show slightly faster epidemic growth than is justified, and hence slightly greater reductions in efficacy due to delaying control measures. A change in β has no effect on the relative importance of different routes of transmission, however, or on the impacts of control measures focused on contact precautions versus case management. The major findings of this study therefore should be robust to misestimation of the distribution of symptomatic periods.

Figure captions – Appendix

Figure S1

Testing sensitivity to the absolute size of the system. Selected results are presented for both HCW and community pools ten-fold smaller than in the main text (HCW pool has 300 individuals and community pool has 10,000 individuals). For each figure, all parameter values other than population sizes are as described in the main text. (A) Analogue of Figure 2D. (B) Analogue of Figure 2F. (C) Analogue of Figure 3B. (D) Analogues of pie-charts from Figures 4A-B.

Figure S2

Testing sensitivity to the relative size of the HCW pool. Results are presented for a HCW pool of 1000 individuals (compared to 3000 throughout the main text), and community pool of 100,000 individuals. Again, all other parameters are as given in the main text. (A) Analogue of Figure 2D. (B) Analogue of Figure 2F. (C) Analogue of Figure 3B. (D) Analogues of pie-charts from Figures 4A-B.

Figure S3

Robustness of conclusions regarding sensitivity of R to transmission-reduction parameters. All details are as given in Figure 2F except as noted. (A) No quarantine: q=0. (B) No quarantine, and limited case isolation: h_c=0.1, h_h=0.1, q=0. (C) Case management as in Figure 2F, but R_0=5. Also note Figures S1B and S2B, which show the insensitivity of these results to absolute and relative size of the two pools.

Figure S4

Distribution of (A) incubation periods and (B) symptomatic periods used in the model, each generated from 10,000 Monte Carlo simulations using the stage progression rules outlined in the text.

Figure S1

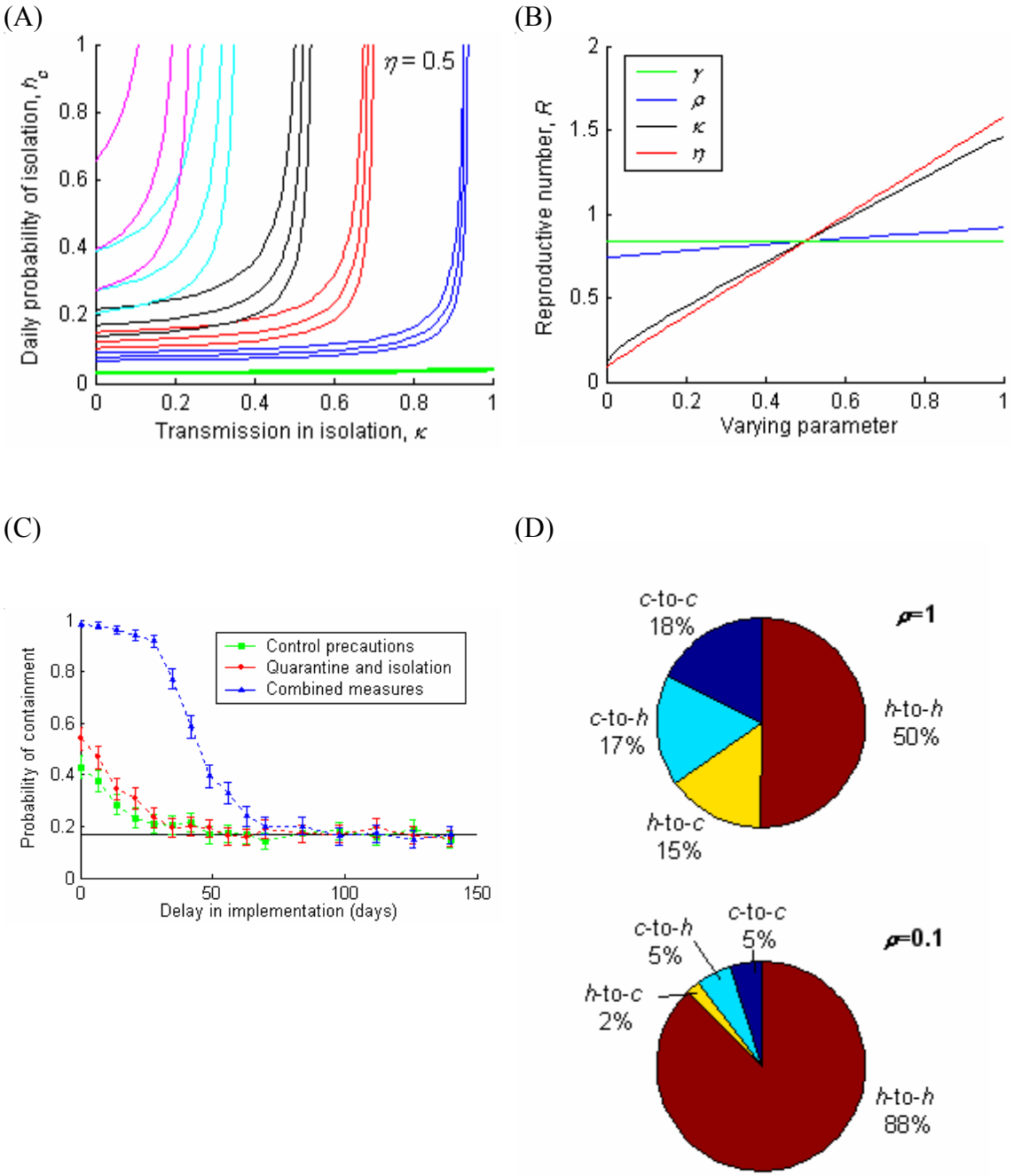

(A)
Daily probability of isolation, h_c
Transmission in isolation, κ
$\eta = 0.5$
(B)
Reproductive number, R
Varying parameter
γ
ρ
κ
η
(C)
Probability of containment
Delay in implementation (days)
Control precautions
Quarantine and isolation
Combined measures
(D)
ρ=1
c-to-c 18%
h-to-h 50%
c-to-h 17%
h-to-c 15%
ρ=0.1
c-to-c 5%
c-to-h 5%
h-to-c 2%
h-to-h 88%

Figure S2

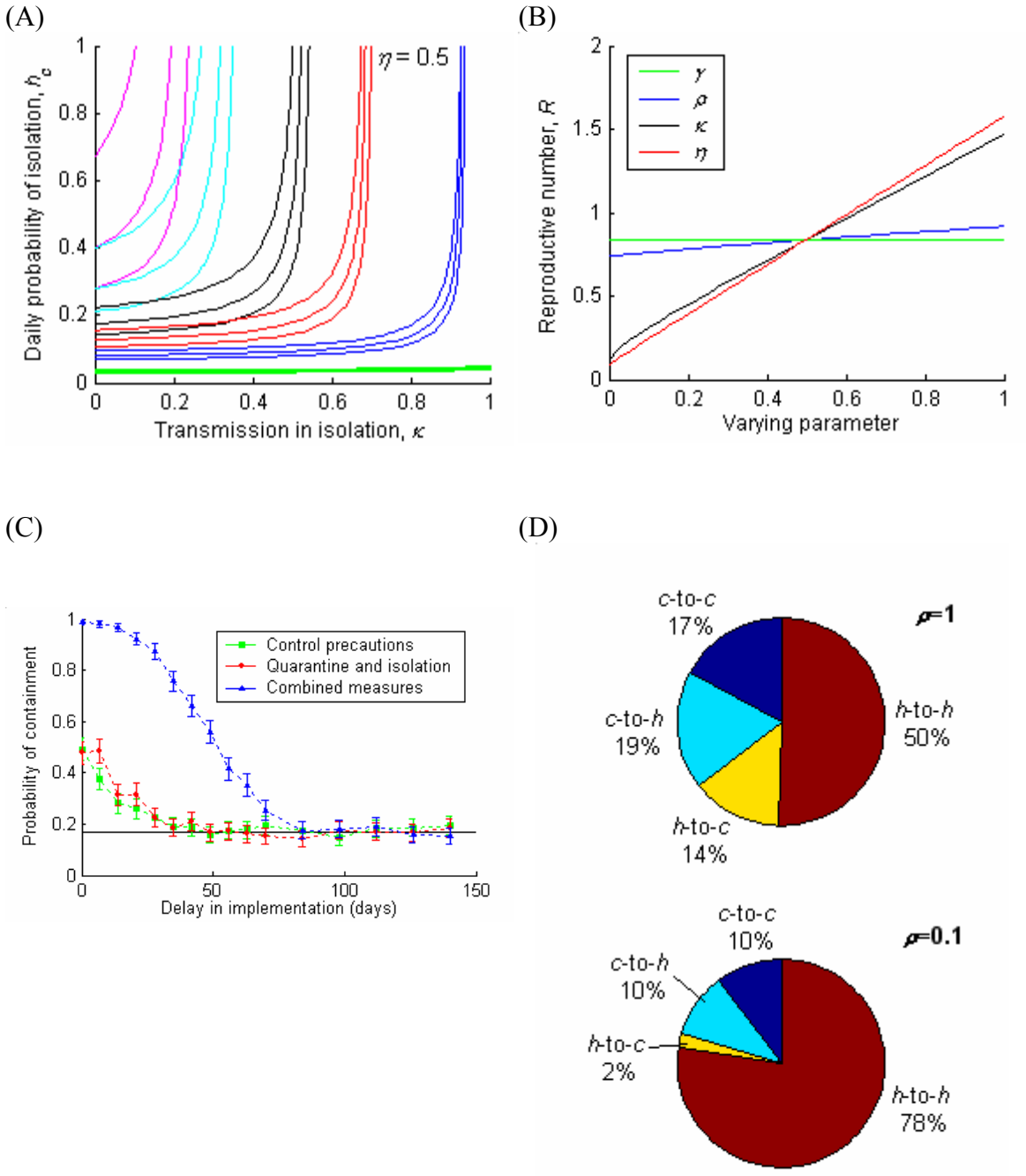
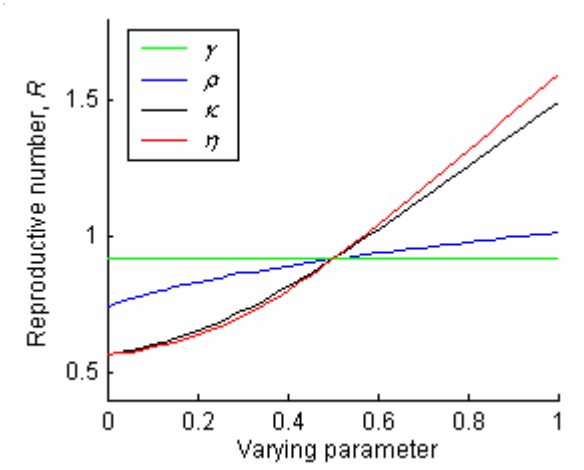
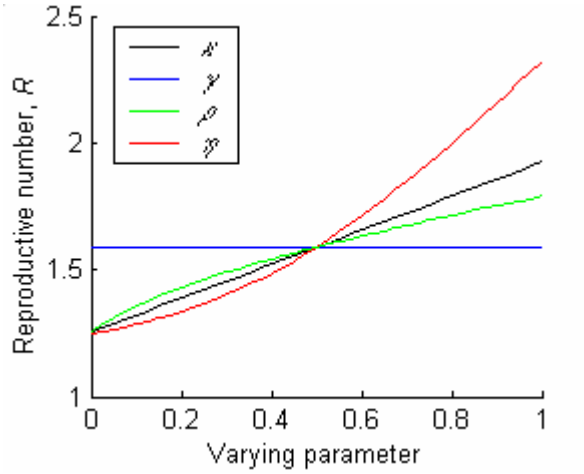

(A)
(B)
(C)
(D)
Daily probability of isolation, h_c
Transmission in isolation, κ
$\eta = 0.5$
Reproductive number, R
Varying parameter
γ
ρ
κ
η
Probability of containment
Delay in implementation (days)
Control precautions
Quarantine and isolation
Combined measures
c-to-c 17%
h-to-h 50%
c-to-h 19%
h-to-c 14%
ρ=1
c-to-c 10%
c-to-h 10%
h-to-c 2%
h-to-h 78%
ρ=0.1

Figure S3

(A)



(B)



(C)

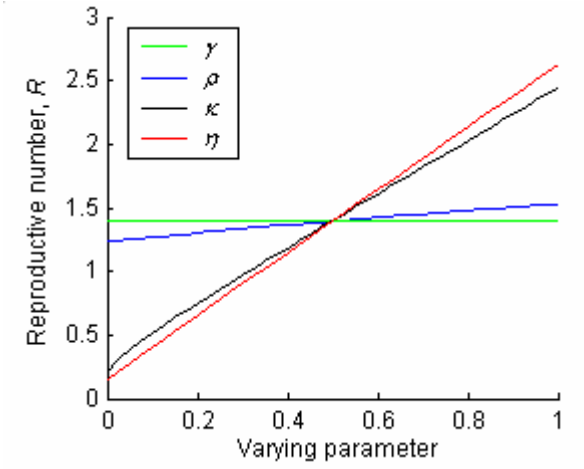
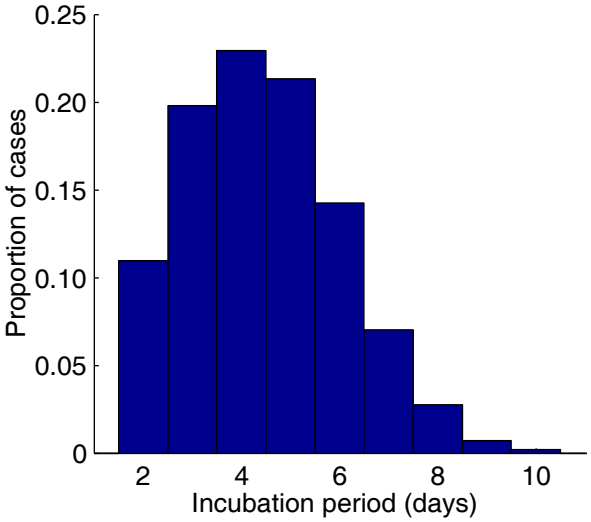
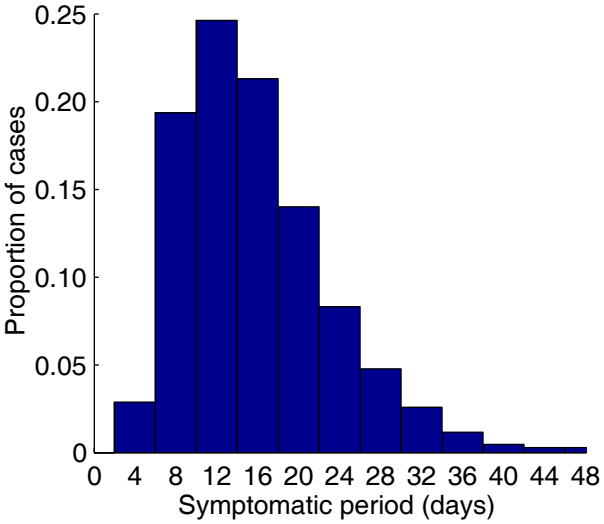


Figure S4

(A)


0.25
0.20
0.15
0.10
0.05
0
Proportion of cases
2
4
6
8
10
Incubation period (days)

(B)


0.25
0.20
0.15
0.10
0.05
0
Proportion of cases
0 4 8 12 16 20 24 28 32 36 40 44 48
Symptomatic period (days)

Chapter Four

Superspreading and the impact of individual variation on disease emergence

J.O. Lloyd-Smith, S.J. Schreiber, P.E. Kopp, and W.M. Getz

1. Introduction

During the global emergence of severe acute respiratory syndrome (SARS) in 2003, numerous "superspreading events" (SSEs) were reported in which individuals infected large numbers of secondary cases (Leo et al. 2003; Shen et al. 2004). These SSEs focused attention on the important role of individual variation in disease invasion dynamics, and on the shortcomings of existing epidemiological theory in treating heterogeneity in infectiousness (Dye and Gay 2003; Lipsitch et al. 2003; Riley et al. 2003), a long-recognized challenge (Anderson and May 1991; Levin et al. 1997; Wallinga et al. 1999; Diekmann and Heesterbeek 2000). Previous work on this topic has focused on vector-borne and sexually-transmitted diseases (STDs), and includes heterogeneous infectiousness only via contact rates between hosts. Based on these considerations, a recent study asserted a general "20/80 rule" whereby 20% of individuals are responsible for 80% of transmission of STDs and vector-borne diseases (Woolhouse et al. 1997). In contrast, there has been little empirical work on individual-level variation in transmission for diseases of casual contact (DCC) such as SARS (though see (Gani and Leach 2004)); only the distribution of infectious periods has been analyzed in depth, with the conclusion that conventional models probably incorporate too much individual variation relative to data (Keeling and Grenfell 1998; Lloyd 2001). In this paper, we show statistical evidence of strong individual variation in infectiousness for ten different DCC, and with an integrated theoretical model we demonstrate important effects of this variation on stochastic dynamics of disease outbreaks, with clear implications for understanding and improving disease control

policies. We conclude that measures of both mean and variance of infectiousness are required to characterize the dynamics of an emerging infectious disease.

The basic reproductive number, R_0, is defined as the mean number of secondary infections per infected individual in a wholly susceptible population, and is central to our current understanding of infectious disease dynamics (Anderson and May 1991; Diekmann and Heesterbeek 2000). Heterogeneity is typically included (if at all) by dividing modeled populations into identifiable sub-groups, each of which is itself homogeneous (Anderson and May 1991; Diekmann and Heesterbeek 2000). To account for individual variation in infectiousness, we introduce the "individual reproductive number", ν, which is the expected number of secondary cases caused by a particular individual in the course of their infection. Values of ν are drawn from a continuous probability distribution with population mean R_0 where the distribution of ν enfolds all aspects of variability in infectious histories of individuals including properties of the host, pathogen, and environment. We therefore treat SSEs not as exceptional events (Riley et al. 2003) but as important realizations from the tail of a right-skewed distribution of ν (Dye and Gay 2003; Anderson et al. 2004), enabling an incisive analysis of their crucial effects on outbreak dynamics. We propose a general definition for superspreading events—currently lacking (c.f. (Leo et al. 2003; Riley et al. 2003; Shen et al. 2004; Wallinga and Teunis 2004))—based on separating the contributions of stochasticity and individual heterogeneity to observed variation in transmission. We review 37 published accounts of SSEs for 11 DCC and discuss trends and common features. To enable further empirical study of individual reproductive numbers, we describe datasets sufficient for our analysis and propose a simple new

quantity p_0, the proportion of cases not transmitting, to be tabulated routinely alongside R_0 in outbreak reports.

Due to its self-evident existence and suspected importance, heterogeneity in infectiousness has been incorporated in previous theoretical and simulation studies of DCC dynamics. Lipsitch et al. (Lipsitch et al. 2003) used a branching process as a heuristic tool to demonstrate the increased extinction probability due to individual variation for SARS, and superspreading individuals were included in a network simulation of SARS (Masuda et al. 2004). A separate analysis assumed an exponential distribution of transmission rates for SARS (Chowell et al. 2004). Gani & Leach (Gani and Leach 2004) showed that pneumonic plague transmission is described better by a geometric than a Poisson distribution, and explored resulting impacts on control measures. "Epidemic trees" reconstructed from the 2001 outbreak of foot-and-mouth disease in Britain allowed direct estimation of farm-level reproductive numbers, and emphasized the importance of variance in ν (Haydon et al. 2003). Observed prevalence patterns of *Escherichia coli* O157 in Scottish cattle farms were explained better by models incorporating individual-level variation in transmission than those with farm-level differences L. Matthews et al., in preparation.. Chain binomial models incorporating various assumptions about infectiousness have been used to study stochastic outbreaks in finite populations (such as households) in great detail (Bailey 1975; Becker 1989), and analysis of a multitype branching process yielded important insight into group-level variation in both infectiousness and susceptibility (Becker and Marschner 1990). Our study builds on this work, presenting empirical evidence

integrated with theoretical modeling to demonstrate the universality and practical relevance of individual variation in infectiousness.

2. Evidence of individual-level variation

Let Z represent the number of secondary cases caused by a given infectious individual in an outbreak. Z is a random variable with an "offspring distribution" $\Pr(Z=k)$ that incorporates influences of stochasticity and variation in the population. Stochastic effects in transmission are modeled using a Poisson process, as is conventional (Diekmann and Heesterbeek 2000), with intensity given by the individual reproductive number, ν i.e. $Z\sim\text{Poisson}(\nu)$. In conventional models neglecting individual variation, all individuals are characterized by the population mean, yielding $Z\sim\text{Poisson}(R_0)$. Another common approach, motivated by models with constant recovery or death rates and homogeneous transmission rates, is to assume that ν is exponentially distributed, yielding $Z\sim\text{geometric}(R_0)$. (For clarity we express all distributions using the mean as the scale parameter; see appendix A for correspondence with standard notation.) We introduce a more general formulation, in which ν follows a gamma distribution with mean R_0 and shape parameter k, yielding $Z\sim$negative binomial(R_0,k) (henceforth abbreviated $Z\sim\text{NegB}(R_0,k)$) (Taylor and Karlin 1998). The negative binomial model includes the Poisson ($k\to\infty$) and geometric ($k=1$) models as special cases.

Realized offspring distributions can be determined from detailed contact tracing of particular outbreaks, or from surveillance data covering multiple introductions of a disease. The above candidate models can then be challenged with these data using

model selection techniques, lending insight into the distribution of ν underlying the observed transmission patterns. For SARS outbreaks in Singapore and Beijing, the negative binomial model is unequivocally favored (Figure 1A, Table 1). Stochasticity alone (or combined with exponential variation in infectious period) cannot account for the observed variation in *Z*. For Singapore, the maximum-likelihood estimate of the shape parameter ($\hat{k}_{mle}$) is 0.16, indicating an underlying distribution of ν that is highly overdispersed (Figure 1A, inset). By this analysis, the great majority of SARS cases in Singapore were barely infectious (73% had ν<1) while a small proportion were highly infectious (6% had ν>8), consistent with field reports from Singapore, Beijing and elsewhere (Leo et al. 2003; Shen et al. 2004).

This analysis was applied to outbreak data for ten different DCC, revealing that the influence of individual variability differs in degree between diseases and outbreak settings (Table 1; data and analysis described in appendices A and B). Maximum-likelihood estimates ($\hat{k}_{mle}$) were obtained by fitting the full distribution of *Z* when such data were available, or else from the mean number of secondary cases and the proportion of individuals who did not transmit (i.e. from R_0 and $p_{0,}$ yielding $\hat{k}_{pz}$). The latter data are far less labor-intensive to collect, and $\hat{k}_{pz}$ is close to $\hat{k}_{mle}$ in most cases (Figure S1). For all outbreaks shown in Table 1, data were restricted to periods before specific outbreak control measures were imposed, although in several cases populations were highly vaccinated. SARS exhibits high variation (low values of *k*) in both traced outbreaks, in keeping with its reputation for frequent SSEs. Monkeypox and smallpox (both Variola major and V. minor) exhibit intermediate variation, with $\hat{k}_{mle}$<1 but

bootstrap 90% confidence intervals encompassing k=1 in most cases. Pneumonic plague appears slightly less variable, and Ebola hemorrhagic fever still less so. An unusual outbreak of hantavirus (the first ever reported with human-to-human transmission) exhibited intermediate variation in Z, while data from avian influenza show extreme heterogeneity because only one case was suspected to transmit the disease. For rubella, one well-traced outbreak shows extreme variability ($\hat{k}_{mle}$=0.032, but note confidence interval) that is almost certainly atypical, but epidemiological reports indicate that high variation in infectiousness and explosive outbreaks are a common pattern for rubella (Hattis et al. 1973; Langmuir 1980). Finally, surveillance data for measles in highly vaccinated populations show high variation, probably resulting from rare outbreaks in non-immunized communities.

These heterogeneous distributions of infectiousness can be summarized by plotting the proportion of all transmission due to a given proportion of the population, where the population is ranked in decreasing order of infectiousness (Figure 1B). For a homogeneous population, where all individuals have ν=R_0, this relation is a straight line. For all diseases analyzed here, the line is bowed due to variability in ν—strongly for diseases such as SARS and avian influenza, and moderately for diseases with less heterogeneity. We can compare our findings for DCC to the general 20/80 rule proposed for STDs and vector-borne diseases (Woolhouse et al. 1997), by plotting the proportion of transmission due to the most infectious 20% of the population (t_{20}) versus the shape parameter k (Figure 1C, Table 1). No consensus value of t_{20} emerges so a general 20/80 rule does not hold, but the core insight that transmission is distributed unevenly is certainly supported. Even a disease such as pneumonic plague, with

relatively mild variation in ν, shows roughly half of all transmission due to the most infectious 20% of cases.

3. Superspreading events

Detailed outbreak or surveillance datasets such as those analyzed above are rare, but further evidence for variation in individual reproductive number ν is found in numerous reports of super-spreading events (SSEs) for many diseases. We have reviewed 37 published accounts of SSEs for 11 DCC (Figure 2A; Table S1), and discerned that unrecognized or misdiagnosed illness (perhaps arising from atypical presentation or co-morbidities) is the most common cause of superspreading, followed by high contact rates and coinfections aiding transmission (see appendices A and B). High pathogen load or shedding rates are implicated in several instances, but usually are not measured. SSEs are associated with a broad range of DCC, but were conspicuously common for SARS (Figure 2A). While some predictors of SARS superspreading have been postulated (Shen et al. 2004), reasons for the abundance of SSEs reported for SARS are not fully understood. Undiagnosed cases and misguided medical procedures played a strong role (Kamps and Hoffmann 2003), especially early in the epidemic, but other possible factors include high variation in host-pathogen interactions, variability in viral load with time, or reporting bias due to intense surveillance.

A consistent and general definition of SSEs is currently lacking—for SARS an SSE has been defined arbitrarily as $Z \geq 8$ (Shen et al. 2004), $Z \geq 10$ (Leo et al. 2003), $Z > 10$ (Wallinga and Teunis 2004) or "many more than the average number" (Riley et al. 2003), but clearly none of the first three definitions can apply to measles with R_0 as high

as 18 (Anderson and May 1991), while the latter is not sufficiently precise. We propose the following general protocol for defining an SSE: (1) set the context for transmission by estimating the effective reproductive number, R, for the disease and setting in question, including immunization levels; (2) construct a Poisson distribution with mean R, representing the expected range of Z due to stochastic effects in the absence of individual variation; (3) define an SSE as any case who infects more than $Z^{(n)}$ others, where $Z^{(n)}$ is the n^{th} percentile of the Poisson(R) distribution. If the 99$^{\text{th}}$ percentile is used, as we do here, an SSE is defined as any case that leads to more secondary cases than would occur in 99% of infectious histories in a homogeneous population.

The observed proportion of cases that caused SSEs, denoted Ψ_{obs}, can be compared to expected values and confidence intervals under the Poisson and NB models for Z (Figure 2B). In all datasets for which SSEs were observed, Z~NegB(R,k) gives a closer estimate of Ψ_{obs} than Z~Poisson(R). Furthermore in five of nine instances Ψ_{obs} lies outside the 95% confidence interval for the homogeneous assumption, while in all cases it lies within the NB confidence interval. This extends the support for the NB model, showing that it provides a reasonable estimate of the high-Z tail of the offspring distribution. Further important information (accessible even when the total number of cases is unknown) is held by the size of SSEs, which often greatly exceeds R. For example, Z=84 for one measles SSE in a highly vaccinated school environment (Table S1). If R=20 (a very generous estimate given vaccination levels of US schoolchildren exceeding 80%), then $\Pr(Z \geq 84|Z\sim\text{Poisson}(R))=3.8\times10^{-15}$ under the assumption of homogeneous ν, even if R=40 then $\Pr(Z \geq 84|Z\sim\text{Poisson}(R))=9.0\times10^{-10}$. (In contrast, for crude estimates R=6 and k=0.5 then $\Pr(Z \geq 84|Z\sim\text{NegB}(R,k))=2.5\times10^{-4}$.) While we

cannot make inferences based on the probability of single events, the accumulated evidence is a strong case against the assumption of homogeneous infectiousness (Figure 2A).

For diseases with different degrees of individual variation we can now predict the frequency of SSEs (Figure 2C), a standing challenge in emerging disease epidemiology (*4*). Once the threshold number of cases $Z^{(99)}$ has been defined for a 99th-percentile SSE under effective reproductive number R, then for any k one can calculate from Z~NegB(R,k) the proportion of individuals $\Psi_{R,k}$ expected to generate $Z>Z^{(99)}$. In a homogeneous population ($k\rightarrow\infty$), $\Psi_{R,\infty}\leq 0.01$ by definition (where the less-than arises because the Poisson distribution is discrete (see appendix A)). When heterogeneity is accounted for, $\Psi_{R,k}>\Psi_{R,\infty}$ and varies strongly with both R and k, peaking between k=0.1 and k=1 for the low R values of interest for emerging diseases. Because the variance-to-mean ratio is fixed at 1 for the Poisson distribution but increases linearly with R for the NB model (see appendix A), for moderate k values $\Psi_{R,k}$ increases strongly with R as the relative density of $Z>Z^{(99)}$ increases. The influence of SSEs on a particular outbreak can be summarized intuitively by the superspreading load, L_{SSE}, defined as the factor by which the number of SSEs exceeds what is expected under homogeneity (For 99th-percentile SSEs, we suggest $L_{SSE}=\Psi_{R,k}/0.01$. Strictly, this is the factor by which the frequency of SSEs exceeds the 1% target frequency under homogeneity as laid out in our SSE definition. The true frequency of SSEs under homogeneity is $\Psi_{R,\infty}=0.01$, so $L_{SSE}=\Psi_{R,k}/\Psi_{R,\infty}$. $\Psi_{R,k}/\Psi_{R,\infty}$ is not a desirable summary measure because it varies erratically with R, as $\Psi_{R,\infty}$ fluctuates in response to integer changes in $Z^{(99)}$.).

4. Impact on disease invasion

Individual variation in infectiousness for DCC thus is supported by evidence from both outbreak reports and SSEs, and the degree of variation can differ markedly between diseases and outbreak settings. To assess the impact of individual variation on dynamics of disease invasion, we analyze a stochastic branching process model with a negative binomial offspring distribution (see appendix A). This corresponds to an underlying gamma distribution of the individual reproductive number, ν, encompassing a diverse family of distributions (including the exponential and constant cases) depending on the value of the shape parameter k (Figure 3A). The variance of ν is $R(1+R/k)$, so smaller values of k indicate greater heterogeneity.

We first consider the probability of stochastic extinction, q, following introduction of a single infected individual, as a function of R_0 (Figure 3B). For $R_0<1$ all invasions go extinct, as in standard models. For $R_0>1$, the degree of individual variation has a decisive influence on the probability of stochastic extinction. If $R_0=3$, as generally estimated for SARS worldwide (Lipsitch et al. 2003; Riley et al. 2003; Anderson et al. 2004; Wallinga and Teunis 2004), then $q=0.06$ under the assumption of homogeneous ν (or $q=0.33$ if $k=1$ so that ν is exponentially distributed); but $q=0.76$ if $k=0.16$ as estimated from SARS data (Table 1). This greater extinction risk arises chiefly from a higher $\Pr(Z=0)$ resulting from the overdispersed distribution of ν (Figures 1A, 3A, S2A). This effect can thwart invasion by diseases that are very potent spreaders on average: for arbitrarily high R_0, the theory predicts that $q\to 1$ as $k\to 0$ (Figure S2B). Considering the timing of disease extinction, our model predicts that for

low k extinction happens within the first few generations or not at all (Figure S2C). In the event of disease extinction, the expected number of cases in the outbreak is scarcely affected by variation (Figure S2D)—if the invasion dies out spontaneously, the potential variability of a low-k disease was probably not manifested (i.e. no SSEs) and it resembles a homogeneous population with lower R.

When the outbreak avoids stochastic extinction, the rate of epidemic growth is strongly affected by variation in individual reproductive number (Figure 3C). Diseases with low k display infrequent but explosive epidemics following introduction of a single case. This pattern is strongly reminiscent of SARS in 2003, for which many settings experienced no epidemic despite unprotected exposure to SARS cases (Ha et al. 2004; Park et al. 2004), while a few cities suffered explosive outbreaks. From these simulations, given our estimates $\hat{k}$=0.16 and $\hat{R}_0$=1.63 for SARS and hence $\Psi_{R,k}$=0.092, this difference can be explained simply by the presence or absence of high-ν individuals in the early generations of each outbreak (Shen et al. 2004). In contrast, conventional models (with k=1 or $k\to\infty$) cannot simultaneously generate frequent failed invasions and rapid growth rates without contrived elaborations of model structure. To characterize the dynamics of disease emergence requires both R_0 and a measure of individual variation (such as t_{20}, L_{SSE}, Ψ, p_0 or k), where the latter determines the relative frequency with which introductions lead to explosive outbreaks via SSEs.

5. Impact on control measures

Beyond improving our understanding of observed patterns of disease invasion, accounting for individual variation in infectiousness can usefully inform planning of

disease control interventions. Health measures, and public awareness of an outbreak, may increase or decrease individual heterogeneity. The population as a whole may alter social mixing and contact patterns (as in cities affected by the 2003 SARS outbreak), or governments may impose isolation, quarantine or infection control on individuals (either traced at random or targeted in groups more likely to produce SSEs). Due to limited facilities and the costs of control, authorities must seek to maximize curtailment of disease spread for a given degree of control effort.

For an outbreak with offspring distribution Z~NegB(R_0,k) before control, we consider the impact of control effort c, where c=0 reflects an uncontrolled outbreak and c=1 reflects complete blockage of transmission. Several idealized cases can be explored theoretically. Under homogeneous partial (HP) control, the infectiousness of every individual in the population is reduced by a factor c (i.e. $\nu_c^{\mathrm{HP}}=(1-c)\nu$ for all individuals). The ν-distribution is rescaled but its shape is not changed, so $R_c^{\mathrm{HP}}=(1-c)R_0$ and $k_c^{\mathrm{HP}}=k$ (see appendix A), and the offspring distribution under control is Z_c^{HP}~NegB(R_c^{HP},k_c^{HP}). Under random absolute (RA) control, a proportion c of infected individuals are traced and isolated completely such that they cause zero infections (i.e. $\nu_c^{\mathrm{RA}}=0$ for a proportion c of infected individuals). If controlled individuals are chosen at random, the offspring distribution under control can be specified exactly (see appendix A) or approximated well by Z_c^{RA}~NegB(R_c^{RA},k_c^{RA}), with $R_c^{\mathrm{RA}}=(1-c)R_0$ and k_c^{RA} the solution to $p_0+c(1-p_0)=\left(1+R_c^{\mathrm{RA}}/k_c^{\mathrm{RA}}\right)^{-k_c^{\mathrm{RA}}}$ where $p_0=(1+R_0/k)^{-k}$ (see appendix A). Note that k_c^{RA} decreases monotonically as c increases (Figure S3C). RA control raises the degree of heterogeneity in the outbreak, as

measured by the variance-to-mean ratio of Z, while HP control reduces it (see appendix A). The two control measures generate identical extinction probabilities q_c for c=0 and $c \geq 1-1/R_0$, but for intermediate values of c the RA approach always yields better control (Figure 4A,B), consistent with our finding in Figure 3B that for a given value of $R>1$, higher variation in Z favors disease die-out. Branching process theory confirms that $q_c^{\mathrm{RA}} > q_c^{\mathrm{HP}}$ whenever $c \in (0, 1-1/R_0)$ (see appendix A).

To assess the realism of these idealized control scenarios, we analyzed contact tracing data from periods before and after introduction of control measures for four outbreaks of DCC. In all cases, $\hat{k}_{c,mle} < \hat{k}_{mle}$ as predicted by the RA model, though small sample sizes limited our ability to establish non-overlapping confidence intervals (Table 2). This increased skew in transmission arises from individuals missed by control efforts or misdiagnosed, who continue infecting others (and even causing SSEs) while controlled cases infect very few. To further examine our control theories, for each dataset we calculated c from $\hat{R}_{c,mle}=(1-c)\,\hat{R}_{0,mle}$ then estimated k_c^{HP} and k_c^{RA} as described above. Once more, in all cases the RA model was closer to the data; twice, $\hat{k}_{c,mle}$ fell between the predictions of the two models indicating a possible combination of control mechanisms (Figure 4C). Larger datasets are required to definitively address this question.

Incorporating knowledge of individual variation into control efforts offers the possibility of increasing efficiency by targeting highly-infectious individuals. If individuals or situations with higher ν can be identified a priori—a significant challenge discussed below—then the probability of outbreak containment for a given level of

control effort can be increased greatly (Figure 4D). Targeting the most infectious 20% of cases so they are four times more likely to be controlled than those in the remaining 80% (and hence total control effort is divided equally between the two groups) allows the impact of control on reducing the chance of a major outbreak to be increased as much as three-fold. When k=0.1 or 0.5, containment is assured for targeted control levels roughly half the threshold level of $c=1-1/R_0$ under random control. For more intensely targeted control this gain can be even larger, but it saturates as greater proportions of the whole population are controlled (Figure S3D). The beneficial effect of targeting acts at lower coverage levels for diseases with more marked individual variation (e.g. k=0.1), but substantial gains are realized for diseases throughout the range of k estimated in Table 1. Again branching process theory generalizes these findings: for a given proportion c of individuals controlled absolutely, greater targeting of higher-ν individuals always leads to lower reproductive number R_c and higher extinction probability q_c (see appendix A).

6. Discussion

Individual variation in infectiousness has broad effects in emerging disease epidemiology. Explosive epidemics demand rapid action by health authorities, and can strain public health infrastructure. Higher probabilities of stochastic extinction indicate that disease introductions may be occurring more frequently than we suspect: for a disease like SARS with R_0~3, if we assume homogeneity we would expect major outbreaks following ~95% of instances when the virus jumps the species barrier to humans, whereas under our empirical estimate of heterogeneity only ~25% would

succeed (Figure 3B). For a given reduction in R, control measures focusing on particular individuals (and hence increasing variation) are preferable to those applying partial measures to everyone, and targeting the most infectious individuals can yield several-fold gains in control efficacy (Figure 4D).

To maximize the benefit of these insights, more detailed data collection on transmission patterns is required. Datasets presented here were collected from published literature, and may be subject to selection bias for unusual instances such as SSEs rather than typical disease behavior. The proportion of individuals who do not transmit (i.e. those with Z=0) strongly influences estimation of the shape parameter k, but is prone to misestimation due to missed sporadic cases in surveillance datasets and to individuals in traced outbreaks whose source case could not be identified. We strongly urge that detailed transmission tracing data be collected and published whenever possible, so that more can be learned about the important influence of individual variation on disease emergence. At the least, we propose a new measure to be included in outbreak reports alongside R_0 and the secondary attack rate: the proportion of cases not transmitting (p_0), which with R_0 is sufficient to estimate the degree of heterogeneity in ν. To be amenable to detailed tracing, a disease must be relatively rare in a population, be directly-transmitted, and have distinctive symptoms and few subclinical cases; to date, the intense effort required to collect such data has been expended only for diseases of particular public health interest. As more data become available, it will be interesting to learn whether trends emerge in the degree of variation present, e.g. for zoonotic versus human diseases, different modes of spread, or varying levels of virulence.

Our analysis has several other limitations. A fundamental assumption of branching process models is that cases are independent of one another, hence we have ignored the possibility that values of ν could be correlated within chains of transmission. We have highlighted the importance of variable infectiousness, but variable susceptibility (of an individual's contacts) is treated only as a potential factor influencing ν. Each individual's susceptibility and infectiousness will be correlated to the extent that each is driven by contact rate (100% correlation is assumed in previous work on STD dynamics (May and Anderson 1987)), in which case high-ν individuals will be more likely to become infected (see (Becker and Marschner 1990) for a group-level treatment of this scenario). Our analysis is founded on non-overlapping generations of transmission, and time is not modelled explicitly. This does not affect the main questions we address (of ultimate extinction probability or outbreak size), but could misrepresent the potential role of long infectious periods or disease carriers.

To successfully target high-ν individuals, we need to understand what factors contribute to variation in infectiousness, and potentially lead to SSEs. Variation in individual reproductive number arises due to a combination of host, pathogen and environmental effects (see additional references in appendices A and B). At the host level, distributions of contact rates are often skewed (Jezek and Fenner 1988; Woolhouse et al. 1997; Eubank et al. 2004; Shen et al. 2004) and index cases in SSEs are often noted to have high numbers of occupational or social contacts (Shen et al. 2004). Increased transmission is correlated with host activities that facilitate pathogen dispersion (such as food handling and singing), coinfections with other pathogens that aid spread or mask symptoms, and age (Shen et al. 2004); previous vaccination can lead

to lower infectiousness. Other relevant host factors may include hygiene habits, immunocompetence, norms regarding bodily contact, and tendency to seek treatment or comply with control measures. Host-pathogen interactions affect transmission rates via variation in pathogen load or shedding (Hamburger et al. 1945; Fenner et al. 1988) and in symptom severity (which may increase or decrease transmission: (Rao et al. 1968; Fenner et al. 1988; Shen et al. 2004)). At the pathogen level, evolution of highly-transmissible pathogen strains is possible, but should lead to observable correlations in Z within transmission chains. An open question is the extent to which pathogen biology influences the different degrees of heterogeneity observed in Table 1.

Environmental factors have a strong influence on transmission. Crowded or confined settings—such as schools, nightclubs, markets, and airplanes—often lead to multiple infections, as can funerals and hospitals for virulent diseases (see appendices A and B). Other important environmental factors are the susceptibility of an individual's contacts, due to age, illness (Shen et al. 2004) or lack of vaccination, and the state of medical knowledge, particularly for a novel disease such as SARS for which misguided procedures and missed diagnoses are inevitable. Imperfect disease control measures tend to increase variation in ν, since transmission is concentrated in a few missed cases or pockets of unvaccinated individuals: this pattern manifests strongly in data for SARS, smallpox, and measles (see appendix B), and is replicated in our theoretical analysis of random absolute control (Figure 4C). We emphasize that all of these host, pathogen and environmental factors join to comprise a case's infectious history, which in turn dictates the individual reproductive number ν.

Recognition of the universality of individual variation in infectiousness for DCC opens challenges for future work. Of central interest for disease emergence is the impact of individual variation on adaptive dynamics of host-pathogen evolution (Levin et al. 1997; Dieckmann et al. 2002), including mutual invasibility of strains with different k. Established theories of control, such as the critical vaccination coverage required to eradicate a disease (Anderson and May 1991), could be revisited in light of the distinction found between homogeneous partial and random absolute control, and the benefits of targeting high-ν individuals should be elaborated. Innovative recent work bases surveillance methods for diseases with $R<1$ on the distribution of total outbreak sizes (Farrington et al. 2003; Jansen et al. 2003), and proposes using a threshold number of cases as an indicator of possible genetic reassortment of avian influenza strains (Ferguson et al. 2004). These analyses assume constant or exponentially-distributed ν, incorporating less variation than we find for most diseases in our analysis. Reanalyzing these issues with gamma-distributed ν may yield further insights, including how to address the complications and potential false alarms (acknowledged by the authors (Ferguson et al. 2004)) caused by SSEs when using a threshold cluster size as a marker of flu adaptation. (Note that the avian influenza data reported in (Ferguson et al. 2004) yields $\hat{k}_{mle}=0.026$ (but note the 90% CI, $(0.026,\infty)$), while the geometric offspring distribution applied in its analysis assumes $k=1$.)

Given the litany of contributing factors outlined above, the marked individual heterogeneity evident in outbreak data is unsurprising. Almost all transmission datasets soundly rejected the Poisson model (Table 1), showing that describing the infectiousness of all individuals by the population mean R_0 is inconsistent with

observations for the diseases examined here. Furthermore the behavior of mean-based models can differ sharply from that predicted when data-driven levels of variation are incorporated (Figures 3B-C), and from observations in the field: real outbreaks often grow too fast, or die out too frequently, compared to predictions of models assuming homogeneous R_0. We argue that a continuously-distributed individual reproductive number, ν, is a logical and necessary extension to the concept of R_0. Data show that individual variation is a universal feature of disease transmission (Tables 1,2,S1,Figures 1B-C,2A-B,4C (Woolhouse et al. 1997)), if not always in a fixed 20/80 proportion, and epidemiological theory should reflect that reality.

Table 1. Parameter estimation from outbreak and surveillance data, and statistical support for the Poisson (P), geometric (G) and negative binomial (NB) models of the offspring distribution.

Datasets (full distribution of Z_{obs})	Model	ΔAIC_c	Akaike weight	$\hat{R}_{0,mle}$ (90% CI)	$\hat{k}_{mle}$ (90% CI)	t_{20} (90% CI)
SARS	P	250.4	0	1.63	0.16	0.88
Singapore 2003	G	41.2	0	(0.54,2.65)	(0.11,0.64)	(0.60,0.94)
N=57	NB	0	1			
SARS	P	49.2	0	0.94	0.17	0.87
Beijing 2003	G	10.6	0	(0.27,1.51)	(0.10,0.64)	(0.60,0.95)
N=33	NB	0	1			
Smallpox (V. major)[v80?]	P	129.3	0	3.19	0.37	0.71
Europe 1958-1973	G	7.4	0.02	(1.66, 4.62)	(0.26, 0.69)	(0.59,0.79)
N=32[s]	NB	0	0.98			
Smallpox (V. major)[v20-70]	P	13.0	0	0.80	0.32	0.74
Benin 1967	G	0.8	0.45	(0.32, 1.20)	(0.16,1.76)	(0.44,0.88)
N=25	NB	0	0.55			
Smallpox (V. minor)[v60?]	P	16.4	0	1.60	0.65	0.60
England 1966	G	0	0.71	(0.88,2.16)	(0.34,2.32)	(0.41,0.73)
N=25	NB	1.7	0.29			
Monkeypox[v70]	P	10.6	0	0.32	0.58	0.62
Zaire 1980-84	G	0	0.62	(0.22,0.40)	(0.32,3.57)	(0.36,0.74)
N=147[s]	NB	1.0	0.37			

Pneumonic plague	P	15.5	0	1.32	1.37	0.47
6 outbreaks	G	0	0.67	(1.01,1.61)	(0.88,3.53)	(0.37,0.54)
N=74	NB	1.5	0.33			
Avian influenza H5N1	P	2.2	0.17	0.06	0.026	1.00
Southeast Asia 2004	G	0.9	0.32	(0, 0.18)	(0.026,∞)[u,b]	(0.20,1.00)
N=33[s]	NB	0	0.51			
Rubella*[v50-70]	P	83.5	0	1.00	0.032	1.00
Hawaii 1970	G	25.4	0	(0.0,1.95)	(0.013,∞)	(0.20,1.00)
N=19	NB	0	1			
Hantavirus (Andes	P	1.0	0.31	0.70	1.66	0.45
virus)*[†]	G	0	0.52	(0.20,1.05)	(0.24,∞)	(0.20,0.80)
Argentina 1996	NB	2.3	0.17			
N=20						
Ebola HF[†]	P	0	0.56	1.50	5.10	0.34
Uganda 2000	G	1.4	0.28	(0.85,2.08)	(1.46,∞)	(0.20,0.46)
N=13	NB	2.4	0.17			

Datasets (mean and prop. of zeros)	Model	ΔAIC_c	Akaike weight	$\hat{R}_{0,pz}$ (95% CI)	$\hat{k}_{pz}$ (90% CI)	t_{20} (90% CI)
Smallpox (V. major)[v]	P	10.8	0	1.49[#]	0.72	0.58
W. Pakistan	G	0	0.71		(0.32,2.23)[b]	(0.41,0.74)
N=47[s]	NB	1.8	0.29			

Measles[v95]	P	27.8	0	0.63[#]	0.23	0.81
US 1997-1999	G	9.6	0.01	(0.47,0.80)[#]	(0.13,0.40)[b]	(0.70,0.92)
N=165[s]	NB	0	0.99			
Measles[v95?]	P	12.9	0	0.82[#]	0.21	0.83
Canada 1998-2001	G	3.5	0.15	(0.72,0.98)[#]	(0.09,0.52)[b]	(0.64,0.96)
N=49[s]	NB	0	0.85			

Table notes

Upper portion, results based on full distribution of Z; lower portion, only mean of Z and proportion of zeros known. Data and analysis described in appendices A and B.

ΔAIC_c: Akaike information criterion, modified for small sample size, relative to lowest score.

Akaike weight: approximate probability that model is the best of models considered.

$\hat{R}_{0,mle}$ and $\hat{k}_{mle}$: maximum likelihood estimates of mean and shape parameter of negative binomial distribution, from full distribution of Z_{obs}.

$\hat{R}_{0,pz}$ and $\hat{k}_{pz}$: estimates of mean and shape parameter of negative binomial distribution, from mean and proportion of zeros of Z_{obs}.

t_{20}: proportion of transmission due to most infectious 20% of cases, calculated from $\hat{k}$.

[vXX] significantly vaccinated population; XX = proportion vaccinated, ([?] value estimated or unknown).

[*] outbreak probably atypical of disease.

[†] incomplete or inaccurate tracing suspected.

N = number of infectious individuals in dataset.

[s] surveillance data.

90% CI: Bootstrap confidence intervals, based on 10000 resamples and bias-corrected percentile method.

[#] value calculated in source report; raw data not available. 95% confidence intervals reported.

[b] confidence interval determined directly from binomial distribution.

[u] upper limit of confidence interval determined by assumption that $\hat{k}_{pz} \to \infty$ for bootstrap datasets with no non-zero values of *Z*.

Table 2. Impact of control measures on offspring distribution (for Z~NegB(R,k))

Outbreak	Before control			During control		
	N	$\hat{R}_{0,mle}$ (90% CI)	$\hat{k}_{mle}$ (90% CI)	N	$\hat{R}_{c,mle}$ (90% CI)	$\hat{k}_{mle}$ (90% CI)
SARS Beijing 2003	33	0.94 (0.27, 1.51)	0.17 (0.10, 0.64)	43	0.28 (0,0.84)[b]	0.006 (0.006, ∞)[u]
SARS Singapore 2003	57	1.63 (0.54,2.65)	0.16 (0.11,0.64)	114	0.68 (0.17, 1.14)	0.071 (0.049, 0.41)
Pneumonic plague Mukden 1946	12	2.00 (1.08, 2.75)	2.63 (1.26, ∞)	27	0.41 (0.11, 0.63)	0.32 (0.12,2.15)
Smallpox (V. major)[v?] Kuwait 1967	4	2.75 (1.00, 4.00)	2.64 (2.23, ∞)	23	0.91 (0, 1.57)[b]	0.026 (0.012, ∞)[u,b]

Symbols as in Table 1.

Figure captions

Figure 1

Evidence for variation in individual reproductive number ν. (A) Transmission data from SARS outbreak in Singapore, 2003 (Leo et al. 2003). Bars show observed frequency of Z, the number of individuals infected by each case; lines show maximum likelihood fits for Z~Poisson (squares), Z~geometric (triangles), and Z~negative binomial (circles). (A, inset) Probability density function (solid) and cumulative distribution function (dashed) for gamma-distributed ν (corresponding to Z~negative binomial) estimated from Singapore SARS data. (B) Proportion of all transmission due to the most infectious cases for six diseases of casual contact, based on $\hat{k}_{mle}$ values in Table 1. (C) Proportion of transmission due to the most infectious 20% of cases, for data drawn from single outbreaks (circles), multiple outbreaks (squares), and long-term surveillance (triangles). Dashed lines show proportions expected under the 20/80 rule (top) and homogeneity (bottom). Superscript v indicates a significant level of vaccination in the affected population; * indicates an outbreak that is probably exceptional for rubella.

Figure 2

Superspreading events (SSEs). (A) Reported SSEs (diamonds) relative to estimated reproductive number R (squares) for twelve diseases of casual contact. Lines show 5-95 percentile range of Z~Poisson(R), and crosses show the 99th percentile proposed as threshold definition for SSEs. Stars represent SSEs with more than one source case. Other diseases plotted in the top row are: a, Streptococcus group A; b, Lassa fever; c, Mycoplasma pneumonia; d, pneumonic plague; e, tuberculosis. See data in Table S1.

(B) Proportion of cases causing 99th-percentile SSEs, Ψ, compared with predicted Ψ and 95% confidence intervals for Z~Poisson($\hat{R}_{0,mle}$) (circles and dashed lines) and Z~NegB($\hat{R}_{0,mle}, \hat{k}_{mle}$) (triangles and solid lines). Observed proportions (stars) were calculated from (# SSEs)/(# cases), as shown; for Ebola HF, Ψ_{obs}=0 so the star is not visible. Data from Table 1; smallpox 1,2,3 refer to data from Europe, Benin and England, respectively. (C) Expected proportion of cases causing 99th-percentile SSEs, $\Psi_{R,k}$, for outbreaks with Z~NegB(R,k). Values of R were selected such that Pr($Z \leq Z^{(99)}$|Z~Poisson(R))=0.01 (see appendix A).

Figure 3

Disease invasion dynamics with different degrees of individual variation in infectiousness. (A) Individual reproductive number ν is drawn from a gamma distribution with mean R_0 and shape parameter k. Probability density functions are shown for seven gamma distributions with R_0=1.5. (B) Probability of stochastic extinction of an outbreak, q, versus population-average reproductive number R_0, following introduction of a single infected case, for k=0.01 (red), k=0.1 (blue), k=0.5 (green), k=1 (black), k=4 (cyan), $k \to \infty$ (magenta). (C) Growth rate of simulated outbreaks with R_0=1.5 and one initial case, conditional on non-extinction. Boxes show median and interquartile range of the first disease generation with 100 cases; whiskers show 95-percentile range, and crosses show outliers. Percentages show the proportion of 10,000 simulated outbreaks that reached the 100-case threshold (i.e. roughly 1-q).

Figure 4

Impact of control measures. (A) Probability of stochastic extinction for diseases with different degrees of individual variation, k, under homogeneous partial (HP) control policies where the infectiousness of all individuals is reduced by a factor c. (B) Increase in extinction probability under random absolute (RA) control compared to HP control. In RA control, a randomly-selected proportion of c of infectious individuals have their infectiousness reduced to zero. In (A) and (B), outbreaks had R_0=3 and began with a single infectious case, and control was assumed to be present from the outset. (C) Estimates of R and k from outbreak datasets before and after control measures were initiated (joined by solid lines; Table 2), and post-control estimates of k_c generated by theoretical models of RA and HP control. (D) Effect of control measures targeting the most infectious individuals. The probability of outbreak containment (defined as never reaching the 100-case threshold) for four diseases with R_0=3 and k=0.1 (blue), k=0.5 (green), k=1 (black), or $k\rightarrow\infty$ (magenta). Control policies are HP (solid lines), RA (dotted lines), or targeted absolute (dashed lines) in which a proportion c of the population has infectiousness reduced to zero (as in RA control) but individuals in the top 20% of infectiousness are four-fold more likely to be controlled. Bars show the factor by which four-fold targeting increased the impact of control relative to RA control (see appendix A), when 20%, 40% or 60% of the total population is controlled. For $k\rightarrow\infty$, targeting has no effect so this factor is 1, and dotted and dashed lines overlay one another. Results are the mean of 10000 simulations, with control beginning in the second generation of cases.

Figure 1

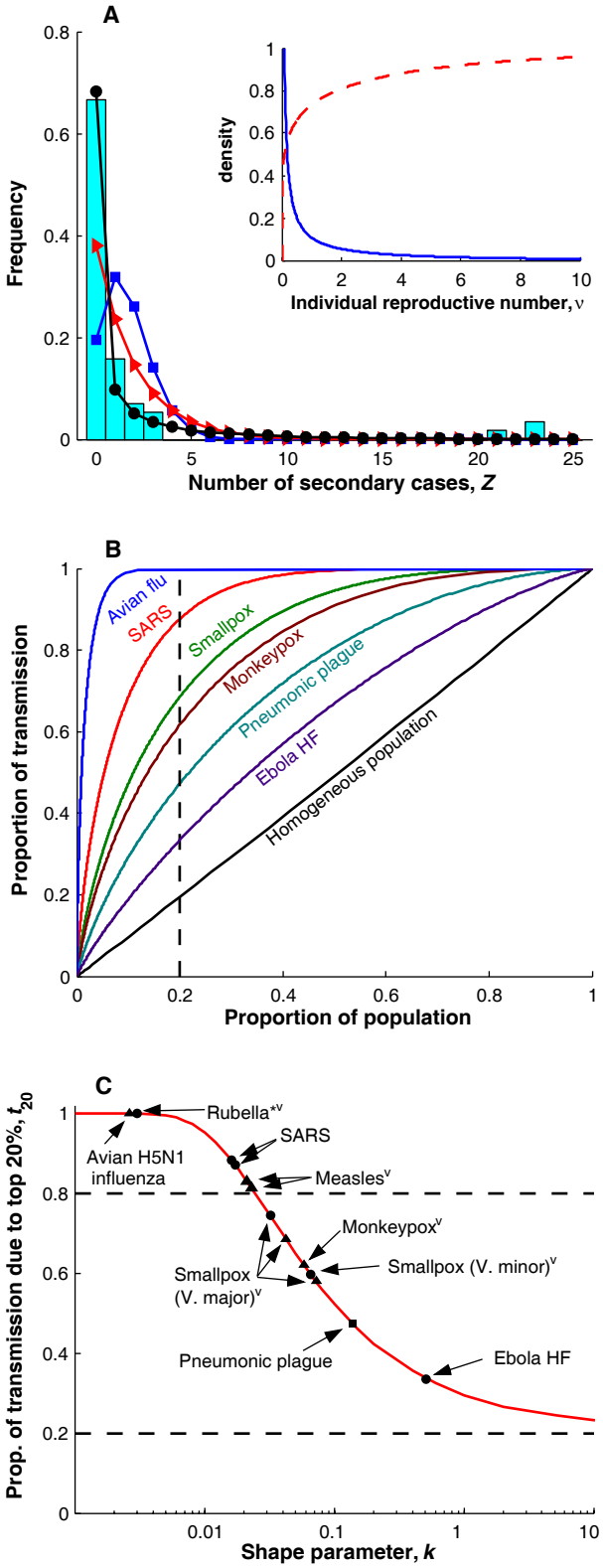

A
Frequency
Number of secondary cases, Z
density
Individual reproductive number, ν
B
Proportion of transmission
Proportion of population
Avian flu
SARS
Smallpox
Monkeypox
Pneumonic plague
Ebola HF
Homogeneous population
C
Prop. of transmission due to top 20%, t20
Shape parameter, k
Rubella*ν
SARS
Measlesν
Avian H5N1 influenza
Monkeypoxν
Smallpox (V. minor)ν
Smallpox (V. major)ν
Pneumonic plague
Ebola HF

Figure 2

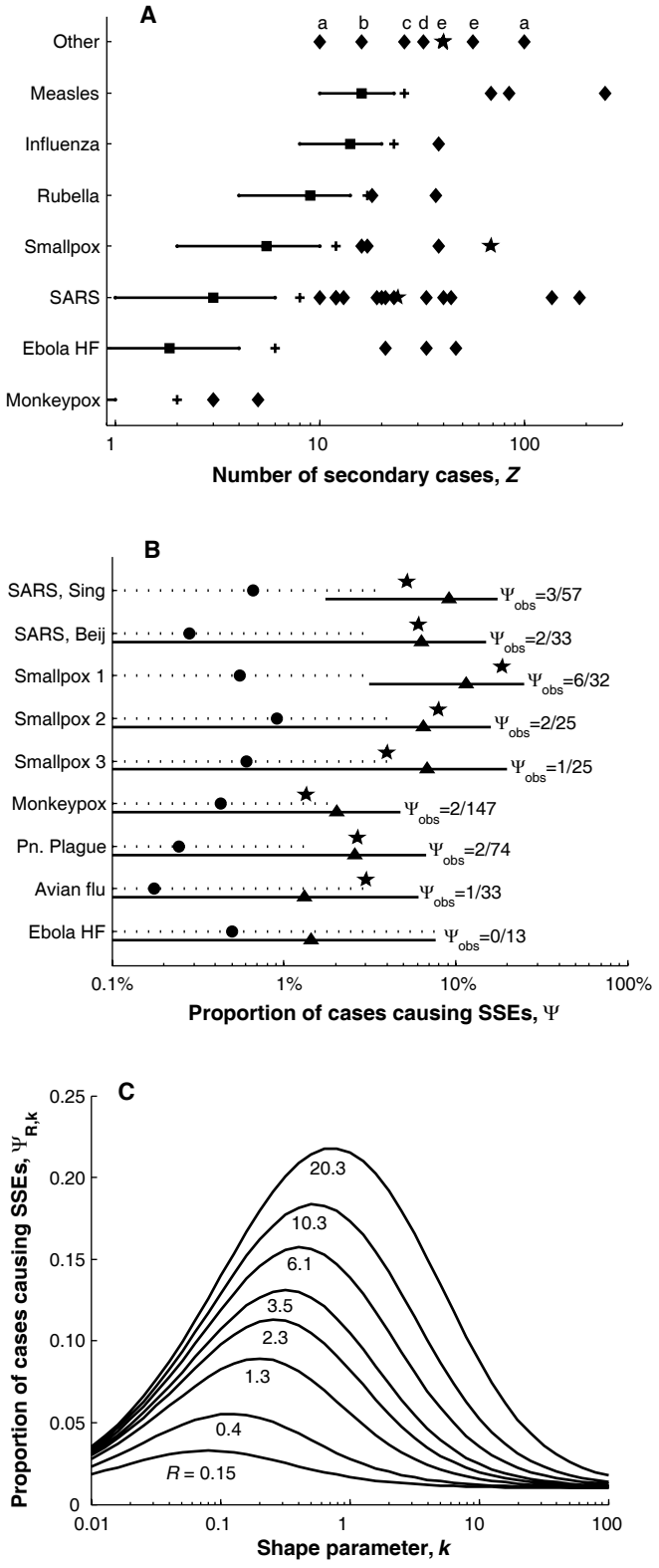

A
a b c d e e a
Other
Measles
Influenza
Rubella
Smallpox
SARS
Ebola HF
Monkeypox
1
10
100
Number of secondary cases, Z
B
SARS, Sing
SARS, Beij
Smallpox 1
Smallpox 2
Smallpox 3
Monkeypox
Pn. Plague
Avian flu
Ebola HF
Ψobs=3/57
Ψobs=2/33
Ψobs=6/32
Ψobs=2/25
Ψobs=1/25
Ψobs=2/147
Ψobs=2/74
Ψobs=1/33
Ψobs=0/13
0.1%
1%
10%
100%
Proportion of cases causing SSEs, Ψ
C
Proportion of cases causing SSEs, ΨR,k
0.25
0.20
0.15
0.10
0.05
0
20.3
10.3
6.1
3.5
2.3
1.3
0.4
R = 0.15
0.01
0.1
1
10
100
Shape parameter, k

Figure 3

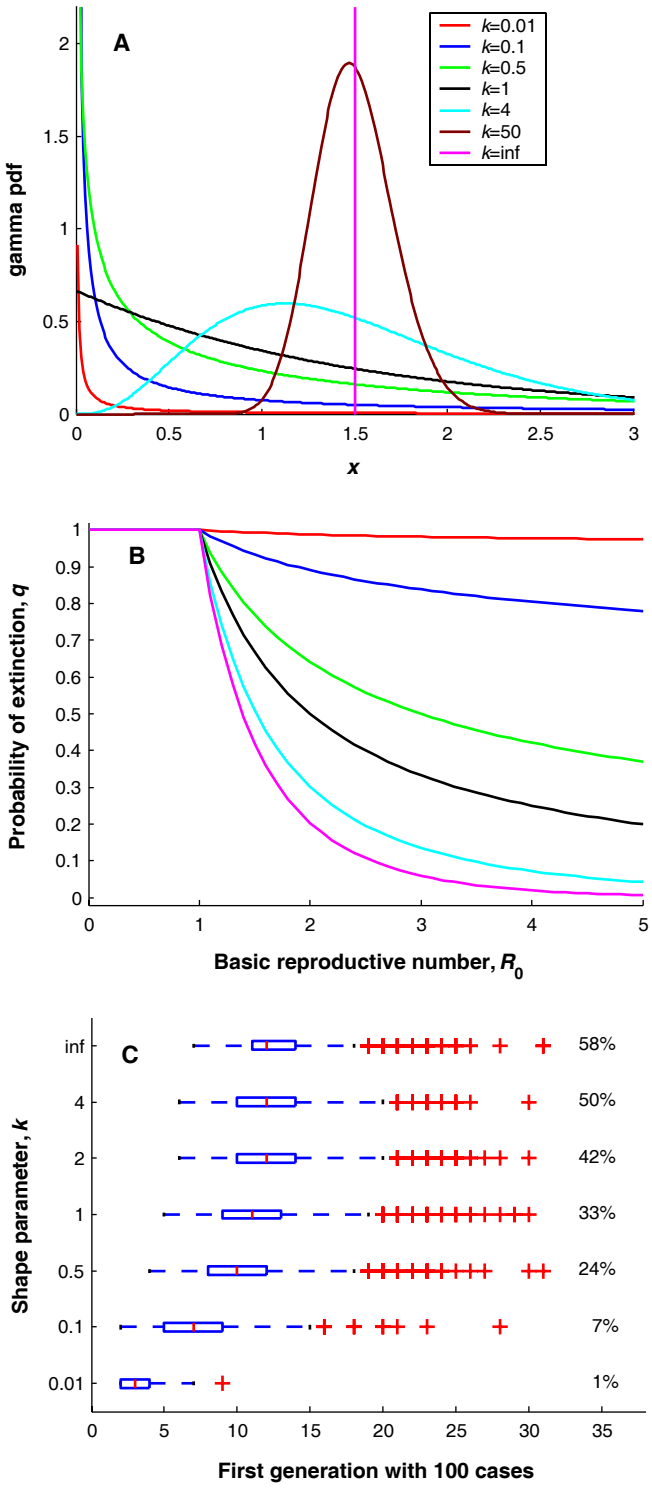
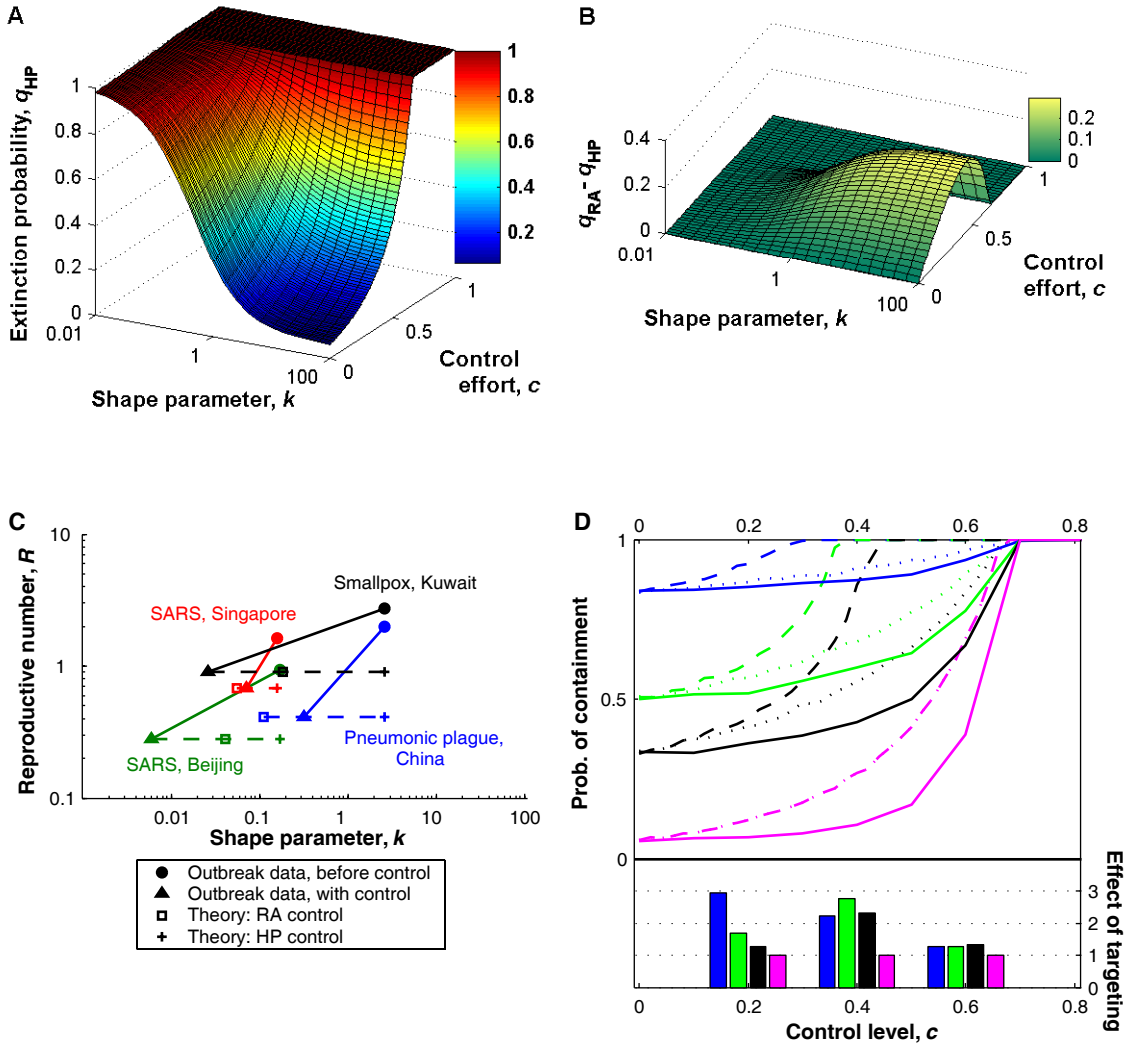

A
k=0.01
k=0.1
k=0.5
k=1
k=4
k=50
k=inf
gamma pdf
x
B
Probability of extinction, q
Basic reproductive number, R0
C
Shape parameter, k
inf
4
2
1
0.5
0.1
0.01
58%
50%
42%
33%
24%
7%
1%
First generation with 100 cases

Figure 4


A
Extinction probability, q_{HP}
Shape parameter, k
Control effort, c
B
q_{RA} - q_{HP}
Shape parameter, k
Control effort, c
C
Reproductive number, R
Shape parameter, k
Smallpox, Kuwait
SARS, Singapore
Pneumonic plague, China
SARS, Beijing
Outbreak data, before control
Outbreak data, with control
Theory: RA control
Theory: HP control
D
Prob. of containment
Effect of targeting
Control level, c

Appendix A – METHODS

Candidate models for the offspring distribution

The offspring distribution is the probability distribution for the number of secondary cases Z caused by each infectious individual. We modeled the offspring distribution using a Poisson process to represent the demographic stochasticity inherent in the transmission process, with intensity ν that could vary to reflect individual variation in infectiousness. The value of ν for a given individual is thus the expected number of secondary cases they will cause, i.e. their individual reproductive number. Note that ν is an expectation and can take any positive real value, while Z is necessarily a natural number (0,1,2,3,…).

The offspring distribution is given by Z~Poisson(ν). We consider three distinct treatments of the individual reproductive number, yielding three candidate models for the offspring distribution. To aid discussion of epidemiological matters, we denote the scale parameter of all offspring distributions by R_0; the relation to conventional notation is stated below.

1. If individual variation is neglected and the individual reproductive number for all cases is assumed to equal the population mean (all $\nu=R_0$), then the offspring distribution is Z~Poisson(R_0).
2. In models with constant per capita rates of leaving the infectious state (by recovery or death), the infectious period is exponentially distributed. If the transmission rate is assumed to be identical for all individuals, then the individual reproductive number is exponentially distributed (ν~exponential($1/R_0$)). Using this expectation in the Poisson process

representing transmission yields a geometric offspring distribution, Z~geometric(R_0) (Taylor and Karlin 1998). (Note: conventional notation is Z~geometric(p) where p=1/(1+R_0).)

3. To incorporate variation in individual infectious histories (from a range of sources), we introduce a more general formulation in which ν follows a gamma distribution with shape parameter k and mean R_0. As shown in Figure 3A, this includes ν=R_0 and ν~exponential(1/R_0) as special cases, and also allows enormous flexibility to fit real-world complexities (at the expense of an added parameter). A Poisson process with this gamma-distributed intensity yields a negative binomial offspring distribution with shape parameter k and mean R_0, Z~NegB(R_0,k) (Taylor and Karlin 1998). (Note: conventional notation is Z~NegB(p,k) where $p = \left(1 + R_0/k\right)^{-1}$.) When k=1 the NegB(R_0,k) distribution reduces to Z~geometric(R_0), and when $k \to \infty$ it reduces to Z~Poisson(R_0).

In all three candidate models, the population mean of the offspring distribution is R_0. The variance-to-mean ratio differs significantly, however, equaling 1 for the Poisson distribution, 1+R_0 for the geometric distribution, and 1+R_0/k for the negative binomial distribution.

Parameter estimation and model selection from full datasets

When full contact tracing information was available, the dataset consisted of a list of Z values for all infected individuals prior to the imposition of control measures. In some instances data from several outbreaks were merged, or surveillance data for the first generation of transmission for many disease introductions were combined.

Descriptions of all outbreaks and issues specific to each dataset are outlined later in the SOM.

For the Poisson, geometric and negative binomial models, the maximum likelihood estimate of the basic reproductive number (R_0 or R) was the sample mean (Rice 1995). For the negative binomial distribution, the shape parameter k was also estimated by maximum likelihood estimation (mle). A numerical algorithm was used for maximization, which required that an upper bound be set for the range of k sampled. This upper bound was set to k=1000, and when this bound was returned as the mle estimate then $\hat{k}_{mle}$ was set to infinity, because a NegB(R_0,k=1000) distribution is indistinguishable in practice from Poisson(R_0).

Having computed the maximum likelihood scores for each dataset, we compared the Poisson, geometric and negative binomial models using Akaike's information criterion (AIC) (Anderson et al. 2000):

$$\mathrm{AIC} = -2\ \ln\left(\mathrm{L}\left(\hat{\theta} \mid data\right)\right) + 2K$$

where $\ln\left(\mathrm{L}\left(\hat{\theta} \mid data\right)\right)$ is the log-likelihood maximized over the unknown parameters (θ), given the model and the data, and K is the number of parameters estimated in the model. Because some of our datasets are small, we used the modified criterion $\mathrm{AIC_c}$, which reduces to the conventional expression as sample size N becomes larger (Anderson et al. 2000):

$$\mathrm{AIC_c} = -2\ \ln\left(\mathrm{L}\left(\hat{\theta} \mid data\right)\right) + 2K + \frac{2K(K+1)}{N-K-1}$$

We rescaled the AIC_c by subtracting the minimum score for each dataset, and present the resulting values ΔAIC_c. We then calculated Akaike weights w_i for each of the three candidate models:

$$w_i = \frac{\exp\left(-\frac{1}{2}\Delta AIC_{c,i}\right)}{\sum_{j=1}^{3}\exp\left(-\frac{1}{2}\Delta AIC_{c,j}\right)}.$$

The Akaike weight w_i can be interpreted as the approximate probability that model *i* is the best model of the set of candidate models considered, in the sense of combining accurate representation of the information in the data with a parsimonious number of parameters (Anderson et al. 2000).

Parameter estimation and model selection from mean and proportion of zeros

For surveillance datasets, frequently only limited information was available. If this included the total number of disease introductions and the number of these that led to no secondary cases then $\hat{p}_0$, the proportion of primary cases for whom Z=0, could be estimated. If the total number of second-generation cases is reported (Heiner et al. 1971), then it was divided by the number of introductions to estimate $\hat{R}$. In some instances the number of cases in later generations was also available, but this information was not used because we could not attribute these cases to specific sources of infection. In the studies on measles in the United States and Canada, data were not available to estimate $\hat{R}$ ourselves so separate estimates of $\hat{R}$ from the original reports were used (Gay et al. 2004; King et al. 2004).

Given estimates of the mean ($\hat{R}$) and proportion of zeros ($\hat{p}_0$) of a negative binomial distribution, the shape parameter *k* can be estimated by numerically solving the equation $\hat{p}_0 = \left(1 + \hat{R}/k\right)^{-k}$. We denoted the resulting estimates $\hat{k}_{pz}$. To ascertain the accuracy of this method of estimation, we compared $\hat{k}_{pz}$ and $\hat{k}_{mle}$ for several outbreaks for which we had full information on *Z* (Figure S1). The proportion of zeros estimate is quite accurate, particularly for $\hat{k}$ <1, but is usually slightly higher than $\hat{k}_{mle}$ and has a broader confidence interval.

To compare the Poisson, geometric and negative binomial models, likelihood scores were determined from proportion-of-zeros data as products of p_0 and $(1 - p_0)$ factors. For a dataset with *N* individuals, of whom $\hat{p}_0 N$ caused no transmission, the likelihood is:

$$L\left(\hat{\theta}, \text{model } j \mid data\right) = \left(p_0^{(j)}\right)^{\hat{p}_0 N} \left(1 - p_0^{(j)}\right)^{(1-\hat{p}_0)N}$$

where $p_0^{(1)} = \exp(-R_0)$ for the Poisson model, $p_0^{(2)} = (1 + R_0)^{-1}$ for the geometric model, and $p_0^{(3)} = (1 + R_0/k)^{-k}$ for the negative binomial model. AIC_c scores and Akaike weights were then generated as described in the previous section.

Bootstrap methodology

Bootstrap confidence intervals were generated for estimated parameters $\hat{R}_0$ and $\hat{k}$ by resampling from outbreak datasets of *Z* values. We employed the bias-corrected percentile method (Efron and Tibshirani 1993; Manly 1998), because both parameters are restricted to positive real values and tended to have skewed bootstrap distributions

for which the median of bootstrap estimates did not equal the parameter estimate from the original dataset. Bootstrap confidence intervals of asymmetric distributions are still prone to errors in coverage (Manly 1998), so the displayed intervals are intended as approximate ranges of uncertainty. We report 90% confidence intervals to avoid known inaccuracies in identifying extreme percentile values of asymmetric bootstrap distributions. We employed 10,000 resamples with replacement to generate our simulated bootstrap distributions.

For estimates of $\hat{k}$ based on the proportion of zeros in a sample, we generated bootstrap confidence intervals directly from the binomial distribution, i.e. without Monte Carlo simulation of the bootstrap distribution. For a dataset with *N* cases of which a proportion $\hat{p}_0$ are zeros, the bootstrap distribution of the proportion of zeros is precisely binomial($N, \hat{p}_0$). To generate the 90% confidence interval, we therefore calculated the 5th and 95th percentiles of the binomial distribution directly (using the Matlab function **binoinv**), and converted the resulting values of $\hat{p}_0$ into estimates of $\hat{k}$ using the method described above. Because we had no information on the distribution of *Z* values in the original dataset (beyond the proportion of zeros), bootstrap estimation of a confidence interval for $\hat{R}$ was not possible. In several instances, confidence intervals for $\hat{R}$ were available from the original published source and reported accordingly.

In certain cases when the full distribution of *Z* values was known, but only one or two *Z* values were greater than zero, the binomial bootstrap was used to obtain the exact bootstrap confidence interval. For example, in the H5N1 avian influenza dataset there is one non-zero value of Z=2 and 32 values of Z=0. The bootstrap distribution of

Z=2 cases is then equal to binomial(N=33,p=1/33), from which exact percentile values could be obtained as above. Since all other Z values in the bootstrap dataset will equal zero, these are easily converted to $\hat{R}$ and $\hat{k}$ using maximum-likelihood estimation.

For all methods of bootstrapping, datasets with very few non-zero values of Z generated significant proportions of bootstrapped datasets with all zeros. The negative binomial shape parameter $\hat{k}$ cannot be calculated from $\hat{p}_0 = \left(1 + \hat{R}/k\right)^{-k}$ in this case. Because all-zero datasets are completely homogeneous and have variance equal to their mean value, we assumed $\hat{k}_{pz} \to \infty$ for these datasets. When 5% or more of bootstrapped datasets had all zero values, then one limit of the 90% CI was determined by this assumption. We marked such instances with a superscript u in Tables 1 and 2.

Proportion of transmission due to top 20%

The proportion of transmission due to a given proportion of the population, plotted in Figure 1B, was calculated as follows. First we estimated R_0 and k, which specify the pdf $f_\nu(x)$ and cdf $F_\nu(x)$ of the gamma-distribution describing the individual reproductive number ν for a given disease and population. We then calculated the cumulative distribution function for transmission of the disease:

$$F_{\text{trans}}(x) = \frac{1}{R_0}\int_0^x u\, f_\nu(u)\, du$$

such that $F_{\text{trans}}(x)$ is the proportion of all transmission due to infectious individuals with $\nu < x$. The proportion of transmission due to individuals with $\nu > x$ is thus $1-F_{\text{trans}}(x)$, while the proportion of individuals with $\nu > x$ is $1-F_\nu(x)$. These quantities were plotted parametrically as a function of x to make Figure 1B. Similarly, the proportion of

transmission due to the most infectious 20% of cases, t_{20}, was calculated by finding x_{20} such that $1-F_i(x_{20})=0.20$, then $t_{20}=1-F_{trans}(x_{20})$.

Superspreading events (SSEs)

An extensive review of superspreading events involving diseases of casual contact, including the data plotted in Figure 2A, is shown in Table S1. As stated in the main text, certain themes emerge repeatedly in case reports of these SSEs, from which we can perhaps learn how better to prevent them. Unrecognized and misdiagnosed cases were a common source of SSEs, as were atypical presentations (or novel diseases) that drew doctors and students to the patient. (On this point, we re-emphasize a conclusion from earlier SARS modeling, which emphasized that strict infection control should be imposed for all individuals in an affected hospital, rather than merely for identified patients (Lloyd-Smith et al. 2003).) Severe coughing, due either to pulmonary involvement of the disease in question or to coinfections with other respiratory pathogens, is often linked to SSEs with suspected airborne transmission. Also common in SSE accounts are circumstances in which the index patient had close or prolonged contact with a large number of others—these events are often quite commonplace (e.g. students in high school hallways, people at parties, dances, or rock concerts, or individuals who continue to work despite symptoms), but sometimes are extraordinary (e.g. several hours aboard a grounded airliner with its ventilation system disabled). Funerals are implicated in SSEs of diseases sometimes thought too lethal for extensive transmission (e.g. Ebola hemorrhagic fever, pneumonic plague, smallpox).

The percentile intervals in Figure 2A were generated directly from the Poisson distribution, with reproductive numbers drawn from specific studies of the relevant diseases where possible (Jezek and Fenner 1988; Gani and Leach 2001; Lipsitch et al. 2003; Chowell et al. 2004), or otherwise from compiled estimates (Anderson and May 1991). These latter estimates of R_0 are intended to be indicative only, since they do not necessarily describe the same population setting or disease strain as the SSEs in question.

The proportion of cases causing SSEs was estimated directly from outbreak datasets for Figure 2B, using the 99th-percentile definition of an SSE proposed in the main text. To compare these observed proportions to the Poisson and negative binomial models for Z, we established 95% confidence intervals for the SSE proportion under each of those models (using mle parameters). For instance, if Z~Poisson(R) the expected proportion of cases causing SSEs is $\Psi_{\text{Poisson}}=1-F_{\text{Poisson}(R)}(Z^{(99)})$, where $F_{\text{Poisson}(R)}(x)$ is the cdf of the Poisson(R) distribution, and $Z^{(99)}$ is the smallest integer satisfying $F_{\text{Poisson}(R)}(Z^{(99)})\geq 0.99$. For an outbreak with Z~Poisson(R) and N cases, the number of SSEs will be binomially distributed, N_{SSE}~binomial(N,Ψ_{Poisson}). To calculate the 95% confidence interval for the observed proportion of SSEs, we find the 2.5 and 97.5 percentiles of N_{SSE}~binomial(N,Ψ_{Poisson}) and divide by N. Then 95% of outbreaks with N cases and Z~Poisson(R) will have an observed proportion of SSEs (Ψ_{obs}) in this range. 95% confidence intervals for Z~NegB(R,k) were calculated in a precisely analogous manner.

Note that the proportion of 99th-percentile SSEs, Ψ_{Poisson}, is often less than 1%, because Poisson(R) is a discrete distribution and for arbitrary R there is unlikely to be

an integer $Z^{(99)}$ such that $F_{\text{Poisson}(R)}(Z^{(99)})$ equals 0.99 exactly. As a result, the proportion of cases causing SSEs under the negative binomial model, $\Psi_{R,k}$, may approach some value less than 0.01 as $k\to\infty$. In plotting Figure 2C, we chose values of R such that $\Psi_{\text{Poisson}}=\Psi_{R,\infty}=0.01$ and all plotted lines approached the same asymptotic value. These values were computed simply by examining Poisson cdf's for different R. Precise values of R in Figure 2C are 0.148, 0.436, 1.279, 2.330, 3.507, 6.099, 10.345, and 20.323. Note that this effect of the discreteness of the Poisson distribution, while a nuisance in making plots, has little practical impact because most diseases have $k<5$ (Table 1).

Branching process model and analysis

We studied the properties of stochastic disease invasions using a single-type branching process model, which allowed us to incorporate individual heterogeneity in infectiousness by varying the offspring distribution. This model of invasion assumes that the supply of susceptible individuals is not limiting for the outbreak, and that the numbers of secondary cases ("offspring") caused by each infectious individual are independent and identically distributed. Branching process models are summarized in depth elsewhere (Harris 1989), as are their particular applications to modeling disease invasion (Diekmann and Heesterbeek 2000).

The heart of a branching process model is the offspring distribution, which describes the probability distribution of the number of new cases Z caused by each infectious individual, i.e. it sets $p_k=\Pr(Z=k)$ for $k=0,1,2,3,\ldots$. Analysis of branching

process models centers on the probability generating function (pgf) of the offspring distribution, $g(s)$:

$$g(s) = \sum_{k=0}^{\infty} p_k s^k, \ |s| \leq 1$$

Two important properties of the epidemic process follow directly from $g(s)$. The basic reproductive number, R_0, is by definition the mean value of Z, and is equal to $g'(1)$. The probability that an infectious individual will cause no secondary infections, $p_0=\Pr(Z=0)$, is $g(0)$. Thus a great deal can be learned about an outbreak from the y-intercept of the pgf and its slope at $s=1$.

The n^{th} iterate of the pgf, $g_n(s)$, is the pgf of Z_n, the number of cases in the n^{th} generation, and is defined as follows: $g_0(s)=s$, $g_1(s)=g(s)$, and $g_{n+1}(s)=g(g_n(s))$ for $n=1,2,3,\ldots$ (Harris 1989). The probability that the epidemic has gone extinct by the n^{th} generation is thus $g_n(0)$. We denote the probability of extinction as $n\to\infty$ by q, then q is a solution to the equation $q=g(q)$, which from monotonicity and convexity of $g(s)$ has at most one solution on the interval (0,1) (Harris 1989). When $R_0\leq 1$, the only solution to $q=g(q)$ is $q=1$ and disease extinction is certain; when $R_0>1$, there is a unique positive solution less than one (Harris 1989).

Finally, the pgf for the total number of individuals infected in all generations of a minor outbreak (i.e. one that goes extinct) is defined implicitly as $G(s)=sg(G(s))$ (Harris 1989). The expected size of a minor outbreak is then $G'(1)$, and can be calculated numerically for a given $g(s)$.

For our treatment of the transmission process, we assume that each infectious case has an individual reproductive number ν, drawn from some distribution with pdf

$f_\nu(u)$. Demographic stochasticity in transmission is then represented by a Poisson process, as is standard in branching process treatments of epidemics (Diekmann and Heesterbeek 2000). This yields the following pgf for a Poisson distribution with mean ν distributed as $f_\nu(u)$:

$$g(s) = \int_0^\infty e^{-u(1-s)} f_\nu(u)\, du$$

If ν is a constant, R_0, then the pgf is:

$$g(s) = e^{-R_0(1-s)}$$

If ν is exponentially distributed with mean R_0, the resulting offspring distribution is geometric with mean R_0 (Taylor and Karlin 1998) and pgf:

$$g(s) = \left(1 + R_0(1-s)\right)^{-1}$$

If ν is gamma distributed, with mean R_0 and shape parameter k, the resulting offspring distribution is negative binomial, also with mean R_0 and shape parameter k (Taylor and Karlin 1998), with pgf:

$$g(s) = \left(1 + \frac{R_0}{k}(1-s)\right)^{-k}$$

This expression was applied in all of the general branching process results shown above to derive our results. The expression $q=g(q)$ was solved numerically to generate Figures 3B and S2B, showing the dependence of the extinction probability on R_0 and k. The negative binomial pgf itself is plotted in Figure S2A, showing how the probability of infecting zero others (p_0) increases sharply with k for a given R_0. The probability of extinction in the n^{th} generation (Figure S2C) was calculated using $g_n(0)-g_{n-1}(0)$. These numerical solutions match the averaged output of many

simulations precisely, for R_0 above and below zero, and for $k \to 0$ and $k \to \infty$. The expected size of minor outbreaks (Figure S2D) was plotted by solving $G'(1)$ numerically for a range of values of R_0 and k.

Branching process simulations

Major outbreaks (Figures 3C, S2E, S2F): A branching process epidemic was implemented by simulation, beginning with a single infectious individual. For each infectious individual, the individual reproductive number ν was drawn from a gamma distribution with chosen values of R_0 and k, using the **gamrnd** function in Matlab (v6.1 R13, MathWorks, Cambridge MA) adapted to allow non-integer k. The number of secondary cases Z caused by that individual was then determined by drawing a Poisson random variable with mean ν, using the Matlab function **poissrnd**. Each individual was infectious for only one generation, and the total number of infected individuals in each generation was summed. The first generation to reach 100 cases was used as an arbitrary benchmark of epidemic growth rate.

Control policies—theoretical development

We consider an epidemic that has a natural (i.e. uncontrolled) offspring distribution $Z \sim \text{NegB}(R_0,k)$, from which we know the probability of infecting zero others is $p_0=(1+R_0/k)^{-k}$. Under the homogeneous partial (HP) control policy, every individual's infectiousness is reduced by a factor c, so the expected number of secondary cases is reduced from ν to $\nu_c^{\text{HP}}=(1-c)\nu$. We model transmission as a Poisson process, $Z \sim \text{Poisson}(\nu)$; imposing HP control, in which each transmission event from the

original process is eliminated with probability c (regardless of the ν value of the source case) changes this to a marked Poisson process so that $Z_c^{\mathrm{HP}} \sim \mathrm{Poisson}((1-c)\nu)$ (Taylor and Karlin 1998). If uncontrolled individual reproductive numbers are gamma-distributed, $\nu \sim \mathrm{gamma}(R_0,k)$, then only the scale parameter of the resulting negative binomial distribution is affected by HP control and $Z_c^{\mathrm{HP}} \sim \mathrm{NegB}((1-c)R_0,k)$. The variance-to-mean ratio of Z_c^{HP} is $1+(1-c)R/k$, and decreases monotonically as control effort increases (Figure S3C).

Under random absolute (RA) control, each infected individual is controlled perfectly (such that they cause zero secondary infections) with probability c. Imposition of RA control influences transmission only for the fraction $1-p_0$ of individuals whose natural Z value is greater than zero—of these a fraction c have $Z_c^{\mathrm{RA}}=0$, while the remaining fraction $1-c$ are unaffected and have $Z_c^{\mathrm{RA}}=Z$. Under an RA control policy, therefore, the proportion of cases causing zero infections is $p_0^{\mathrm{RA}} = p_0+c(1-p_0)$ and the population mean $R_c^{\mathrm{RA}} = \frac{1}{N}\sum_{i=1}^{N} Z_i \Pr(\text{case } i \text{ not controlled}) = (1-c)\frac{1}{N}\sum_{i=1}^{N} Z_i = (1-c)R_0$. The exact distribution of Z_c^{RA} is defined by $\Pr(Z_c^{\mathrm{RA}}=0)= p_0^{\mathrm{RA}}$ and $\Pr(Z_c^{\mathrm{RA}}=j)=(1-c)\Pr(Z=j)$ for all $j>0$, i.e. the distribution of Z_c^{RA} has an expanded zero class relative to Z, while for non-zero values its density is simply reduced by a factor $(1-c)$ from $Z \sim \mathrm{NegB}(R_0,k)$. Hence, the offspring distribution under RA control has pgf:

$$g_{\mathrm{RA}}(s) = c + (1-c)\left(1+\frac{R_0}{k}(1-s)\right)^{-k}$$

The variance-to-mean ratio of Z_c^{RA} can be calculated from $\left(g''_{\mathrm{RA}}(1)+g'_{\mathrm{RA}}(1)-\left(g'_{\mathrm{RA}}(1)\right)^2\right)/g'_{\mathrm{RA}}(1)$ (Harris 1989) and shown to equal $1+R_0/k+cR_0$, which increases monotonically as c increases. For direct comparison with other offspring distributions, this composite distribution under RA control can be approximated by a new negative binomial distribution, $Z_c^{\mathrm{RA}} \sim \mathrm{NegB}(R_c^{\mathrm{RA}}, k_c^{\mathrm{RA}})$ where R_c^{RA} is given above and k_c^{RA} is estimated using the proportion of zeros method as the solution to $p_0^{\mathrm{RA}} = p_0 + c(1-p_0) = \left(1+R_c^{\mathrm{RA}}/k_c^{\mathrm{RA}}\right)^{-k_c^{\mathrm{RA}}}$. This approximation yields better than 95% overlap with the exact distribution for k≤1, and better than 85% overlap for almost all of parameter space (Figure S3A). (The proportion of overlap is calculated as $1-\left(\sum_{i=0}^{\infty}\left|Z_{i,\mathrm{exact}}-Z_{i,\mathrm{approx}}\right|\right)\Big/2$, which scales from 0 to 1 as the two distributions go from completely non-overlapping to identical.) The approximation approaches exactness for $c\to 0$ and $c\to 1$, and is least accurate for large values of k because it is unable to mimic the bimodal distribution of Z_c^{RA} (Figure S3B). The approximated shape parameter k_c^{RA} decreases monotonically as control effort c increases (Figure S3C).

Relative efficacy of control policies

For HP control, with all individuals' transmission reduced by a factor c, the offspring distribution is $Z_c^{\mathrm{HP}} \sim \mathrm{NegB}((1-c)R_0,k)$ and has pgf:

$$g_{\mathrm{HP}}(s) = \left(1+(1-c)\frac{R_0}{k}(1-s)\right)^{-k}.$$

For RA control, with a random proportion c of individuals controlled absolutely, the pgf is as given above:

$$g_{\mathrm{RA}}(s) = c + (1-c)\left(1 + \frac{R_0}{k}(1-s)\right)^{-k}.$$

Claim: For all $c \in (0,1-1/R_0)$, the probability of extinction is always greater under RA control than under HP control.

Proof of claim: Define $G(x) = \left(1 + x\frac{R_0}{k}(1-s)\right)^{-k}$ where X is a Bernoulli random variable with a probability $1-c$ of success. Since G is a convex function, Jensen's inequality implies that

$$g_{\mathrm{HP}}(s) = G\big(E(X)\big) < E\big(G(X)\big) = g_{\mathrm{RA}}(s) \qquad (*)$$

whenever $c \in (0,1)$ and $s \in [0,1)$. Furthermore, for the n^{th} iterates of the pgf we have from (*) that

$$g_{\mathrm{HP},n}(0) < g_{\mathrm{RA},n}(0)$$

so the probability of disease extinction by the n^{th} generation is always greater under RA control. Thus if $c \in (0,1\text{-}1/R_0)$, the probability of ultimate extinction under RA control is greater than that under HP control, i.e. $q^{\mathrm{RA}} > q^{\mathrm{HP}}$. If $c > 1-1/R_0$, then $R_c^{\mathrm{HP}} = R_c^{\mathrm{RA}} < 1$ so that $q^{\mathrm{RA}} = q^{\mathrm{HP}} = 1$.

To consider the efficacy of control policies targeting the more infectious individuals in a population, we consider a general branching process whose pgf is given by

$$g(s) = \int_0^\infty e^{-u(1-s)} f_\nu(u)\,du$$

where $f_\nu(u)$ is the pdf of the individual reproductive number ν for the outbreak in question.

For a control strategy $C : [0,\infty) \to [0,1]$ in which the probability of absolutely controlling a case with individual reproductive number ν is $C(\nu)$, the pgf of the branching process becomes

$$g_C(s) = c + \int_0^\infty e^{-u(1-s)} (1 - C(u)) f_\nu(u)\,du$$

where

$$c = \int_0^\infty C(u) f_\nu(u)\,du$$

is the fraction of individuals controlled on average. For example, randomized absolute (RA) control corresponds to choosing $C(\nu)=c$ for all ν. Maximally-targeted control, in which the top $c \times 100\%$ of infectious individuals are controlled absolutely, corresponds to choosing

$$C(\nu) = \begin{cases} 0 & \text{if } \nu < \nu_c \\ 1 & \text{if } \nu \geq \nu_c \end{cases}$$

where ν_c satisfies $\int_{\nu_c}^\infty f_\nu(u)\,du = c$.

Note that when ν is gamma-distributed with mean R_0 and shape parameter k, the pgf under maximally-targeted control is

$$g_{\max}(s) = c + \left(1 + \frac{R_0}{k}(1-s)\right)^{-k} \left(1 - \frac{\Gamma(k, \nu_c(k/R_0 + 1 - s))}{\Gamma(k)}\right)$$

where $\Gamma(k,b)=\int_b^\infty t^{k-1}e^{-t}dt$ and $\Gamma(k)=\Gamma(k,0)$.

For any distribution of ν represented by $f_\nu(u)$, we can make the following claim.

Claim: Let C_1 and C_2 be two control strategies that satisfy $\int_0^\infty C_i(u)f_\nu(u)du = c$ and

$$\int_x^\infty C_1(u)f_\nu(u)du > \int_x^\infty C_2(u)f_\nu(u)du \qquad (**)$$

for all $x > 0$, so that C_1 targets higher-ν individuals to a greater degree. Then the reproductive number under strategy 1 ($R_c^{C_1}$) is less than that under strategy 2 ($R_c^{C_2}$). Moreover, if $R_c^{C_2}>1$, then the probability of extinction is greater under strategy 1.

Proof of Claim: The claim $R_c^{C_1} < R_c^{C_2}$ is equivalent to $g'_{C_1}(1) < g'_{C_2}(1)$. Recall that if X and Y are positive random variables such that P($X > x$) > P($Y > x$) for all $x > 0$, then E(X)>E(Y) (*1*). Define X_i to be the positive random variable with the pdf

$$\frac{1}{1-c}(1-C_i(u))f_\nu(u)$$

for i=1,2. By (**) we have

$$P(X_2 > x) = \int_x^\infty \frac{1}{1-c}(1-C_2(u))f_\nu(u)du > \int_x^\infty \frac{1}{1-c}(1-C_1(u))f_\nu(u)du = P(X_1 > x)$$

for all $x > 0$. Hence

$$g'_{C_2}(1) = (1-c)E(X_2) > (1-c)E(X_1) = g'_{C_1}(1).$$

The second assertion of the claim is equivalent to the statement that $g_{C_1}(s) > g_{C_2}(s)$ for all $s \in [0,1)$. To prove this, define Y_i=exp($-X_i(1-s)$). Since exp($-x(1-s)$) is a

decreasing function of x for $s \in [0,1)$ and $P(X_2>x)>P(X_1>x)$ for all $x > 0$, we have $P(Y_1>x)>P(Y_2>x)$ for all $x>0$. Hence, $g_{C_1}(s) = c + (1-c)E(Y_1) > c + (1-c)E(Y_2) = g_{C_2}(s)$, and as argued above we have $g_{C_1,n}(0) > g_{C_2,n}(0)$ for all generations n and therefore $q^{C_1} > q^{C_2}$.

To see the utility of this claim, let us consider two control strategies C_1 and C_2 that control two portions of the population in different ways. Suppose strategy C_i controls the less-infectious portion of the population (i.e. $\nu < \nu^*$) with probability a_i and controls the more-infectious portion of the population (i.e. $\nu \geq \nu^*$) with probability b_i. In other words,

$$C_i(\nu) = \begin{cases} a_i & \text{if } \nu < \nu^* \\ b_i & \text{if } \nu \geq \nu^* \end{cases}$$

Moreover, let us assume that both strategies control the same fraction of individuals, i.e. $\int_0^\infty C_i(u) f_\nu(u) du = c$ for i=1,2. Suppose that strategy 1 targets more-infectious individuals to a greater degree than strategy 2, i.e. $b_1 > b_2$ and thus $a_1 < a_2$. This is a generalized formulation of the targeted control scenario discussed in the main text (Figure 4D), for which strategy 1 defines ν^* as the solution to $\int_0^{\nu^*} f_\nu(u) du = 0.80$ and takes $b_2 = 4 \times a_2$, whereas strategy 2 is non-targeted RA control with $a_2 = b_2 = c$. For $\nu \geq \nu^*$:

$$\int_\nu^\infty C_1(u) f_\nu(u) du = b_1 \int_\nu^\infty f_\nu(u) du$$

$$> b_2 \int_\nu^\infty f_\nu(u) du = \int_\nu^\infty C_2(u) f_\nu(u) du$$

and for $\nu < \nu^*$:

$$\int_{\nu}^{\infty} C_1(u) f_\nu(u)\,du = c - a_1 \int_0^{\nu} f_\nu(u)\,du$$

$$> c - a_2 \int_0^{\nu} f_\nu(u)\,du = \int_{\nu}^{\infty} C_2(u) f_\nu(u)\,du.$$

Condition (**) is fulfilled, so $R_c^{C_1} < R_c^{C_2}$ and $q^{C_1} > q^{C_2}$, corroborating the simulation results for targeted control in the main text (Figure 4D).

In general, the more a control policy targets the more-infectious individuals, the higher the probability of disease extinction and the slower the growth rate of an outbreak in the event of non-extinction. If we define strategy TA as any absolute control policy targeting infectious individuals more than random, then for a given control effort $c \in (0,1)$ we have

$$g_{\mathrm{TA}}(s) > g_{\mathrm{RA}}(s) > g_{\mathrm{HP}}(s)$$

for all $s \in [0,1)$, so targeted absolute control is always more effective than random absolute control, which in turn is always better than homogeneous partial control.

Control policies—simulations

Control simulations (Figures 4D, S3D): The branching process simulation from Figure 3C was used. For HP control, every infected case's individual reproductive number was reduced to $(1-c)\nu$ before a Poisson random variate was drawn to determine the number of infections caused. For RA control, every infected case had probability c of having ν reduced to zero before the Poisson random variate was drawn. For targeted absolute control, the total proportion of the population subject to control was c, but the

probability of control for a top-20% individual was four times greater than that for a bottom-80% individual, e.g. Pr(control, top-20%)=1/4 and Pr(control, bottom-80%)=1/16, yielding Pr(control, overall)=1/10. Under four-fold targeting, equal effort (in terms of absolute numbers controlled) is expended on top-20% and bottom-80% individuals.

Targeted control was simulated as follows. For each combination of R_0 and k, the cutoff value of ν dividing top-20% from bottom-80% infectiousness was established from the cdf of ν. During the simulation, after a value of ν was drawn from the gamma(R_0,k) distribution for each infected individual, they were assigned to the top-20% or bottom-80% categories. For individuals in either category, a uniform random variate on (Curtis et al. 1999) was drawn, and if it was less than the probability of control for that category then that individual's value of ν was reset to zero. The realized number of secondary infections Z_c was then generated by drawing a Poisson random variate with mean ν.

For the simulations shown in Figure 4D, control was initiated in the second generation (i.e. the index case was not subject to control), representing a delay in recognition of the outbreak. Containment of an outbreak was defined as preventing it from growing to the point of a generation with 100 cases. Since a branching process that escapes control will grow without bound, results were not sensitive to this arbitrary threshold. The relative effect of targeted control (the quantity plotted in the bottom of Figure 4D) was computed as follows. The uncontrolled probability of a major outbreak for the given R_0 and k was computed as 1−Pr(containment|0% control). The contribution of control efforts to containment was then calculated as:

Contrib(control policy) = Pr(containment|control policy)–Pr(containment|0% control).

The relative effect of targeted control, plotted in Figure 4D, was then:

Relative effect = Contrib(targeted control)/Contrib(random control).

This quantity equaled 1 for $k \to \infty$, since targeting has no effect on a homogeneous population, but is greater than 1 for all finite values of k.

Figure captions – appendix A

Figure S1

Comparison of maximum likelihood and proportion of zeros estimates and 90% confidence intervals for the negative binomial shape parameter k. Each point corresponds to an outbreak for which we have full information on Z, so we are able to estimate $\hat{k}_{mle}$ and the corresponding bias-corrected bootstrap confidence interval. For the same dataset, we then discarded all information except the mean and proportion of zeros and estimated $\hat{k}_{pz}$ and the binomial-bootstrap confidence interval.

Figure S2

Branching process results for Z~NegB(R_0,k). (A) The probability generating function of the negative binomial distribution, plotted for R_0=3 and different shape parameters k. The y-intercept of the pgf equals p_0, the probability that an infected individual will infect nobody, and is a major factor in the rising probability of extinction as k decreases. The extinction probability q is determined by the point of intersection of the pgf with a line of slope 1 (dashed) through the origin. (B) The probability of stochastic extinction given introduction of a single infected individual, q, rises to 1 as $k \to 0$ for any value of R_0. (C) The probability of stochastic extinction by the n^{th} generation of transmission, q_n, for R_0=3 and a range of k. Interestingly, for the third and subsequent generations, the k=1 case has the highest continuing probability of extinction. (D) Expected size of a minor outbreak (i.e. an outbreak that dies out spontaneously) versus R_0. Curves for all k values are identical for R_0<1. (E) Growth rate of simulated outbreaks with R_0=1.1 and one initial case, conditional on non-extinction. Boxes show median and interquartile

range of the first disease generation with 100 cases; whiskers show 95-percentile range, and crosses show outliers. Percentages show the proportion of 10,000 simulated outbreaks that reached the 100-case threshold (i.e. roughly 1-q). (F) Growth rate of simulated outbreaks with R_0=3. Both (E) and (F) are exactly analogous to Figure 3C except for different values of R_0.

Figure S3

Impact of control measures. (A) Accuracy of the approximation whereby the offspring distribution under random absolute (RA) control is represented by a negative binomial distribution, $Z_c^{\mathrm{RA}} \sim \mathrm{NegB}(R_c^{\mathrm{RA}}, k_c^{\mathrm{RA}})$. Calculation of the proportion of overlap is described in the SOM text. (B) Precise and approximated NB offspring distributions under RA control for R_0=3. From bottom to top, five curves for both the precise (red solid lines and circles) and approximate (black dotted lines and squares) distributions show k=0.1, 0.5, 1, 3, and 10. (C) The approximated shape parameter k_c^{RA} decreases monotonically as control effort c increases. Curves depict uncontrolled k=1000 (blue), k=1 (green), and k=0.1 (red), for R_0=1 (solid), R_0=3 (dotted), and R_0=10 (dashed). (D) Effect of control measures targeting the most infectious individuals. The plot is exactly analogous to Figure 3D, except that in the targeted control scenario individuals in the top 20% of infectiousness are ten-fold more likely to be controlled than those in the bottom 80%.

Figure S4

Contact rate data from Zaire, 1980-84, during intensive surveillance for monkeypox emergence. Contacts were defined as any person who was close enough to touch a case during the rash period, including anyone living, eating, or sleeping in the same household as a monkeypox patient during the infectious period. Data from (Jezek et al. 1987; Fine et al. 1988).

Figure S1

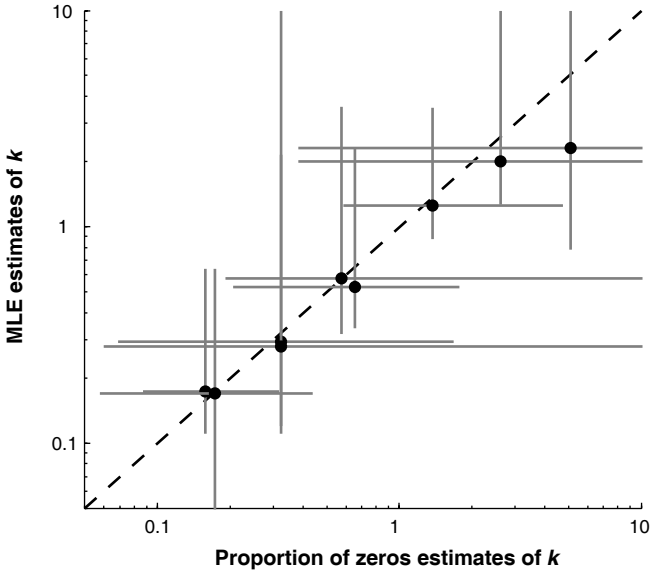

10
1
0.1
MLE estimates of k
0.1
1
10
Proportion of zeros estimates of k

Figure S2

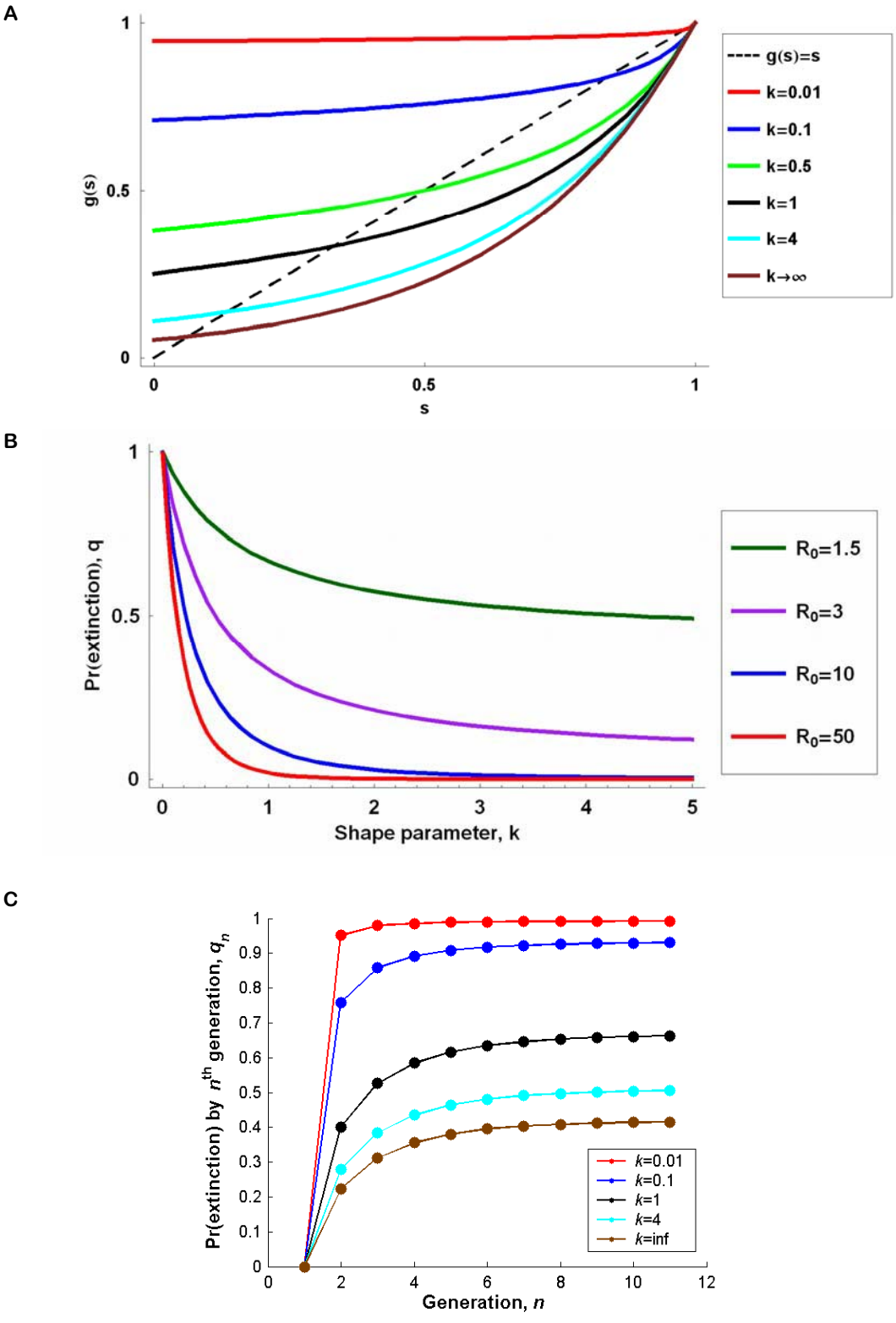
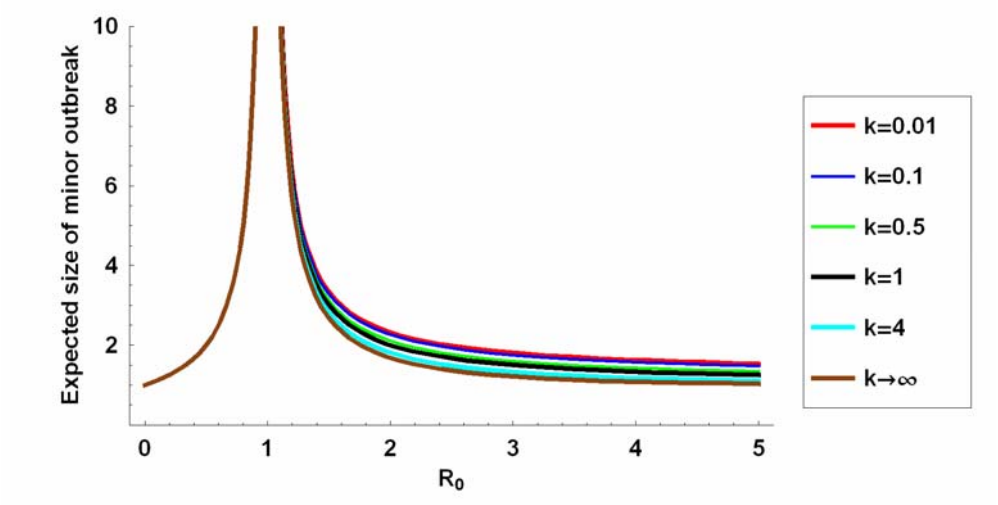
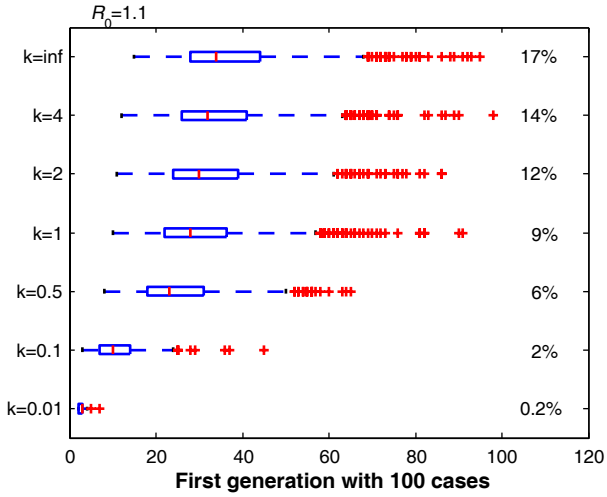

A
g(s)
s
g(s)=s
k=0.01
k=0.1
k=0.5
k=1
k=4
k→∞
B
Pr(extinction), q
Shape parameter, k
R0=1.5
R0=3
R0=10
R0=50
C
Pr(extinction) by nth generation, qn
Generation, n
k=0.01
k=0.1
k=1
k=4
k=inf

Figure S2 (cont)

D



E



F

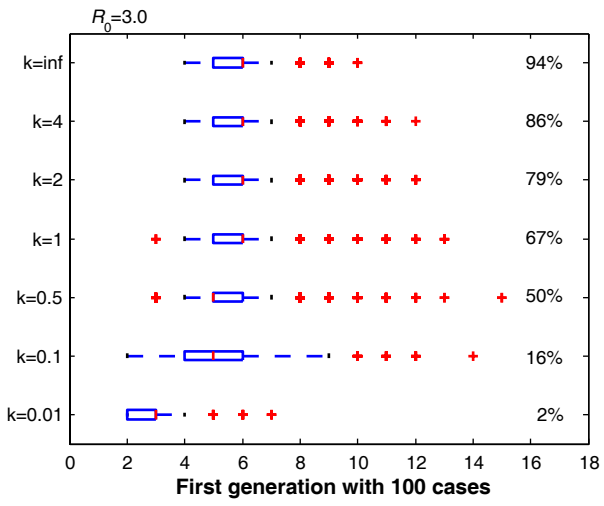


Figure S3

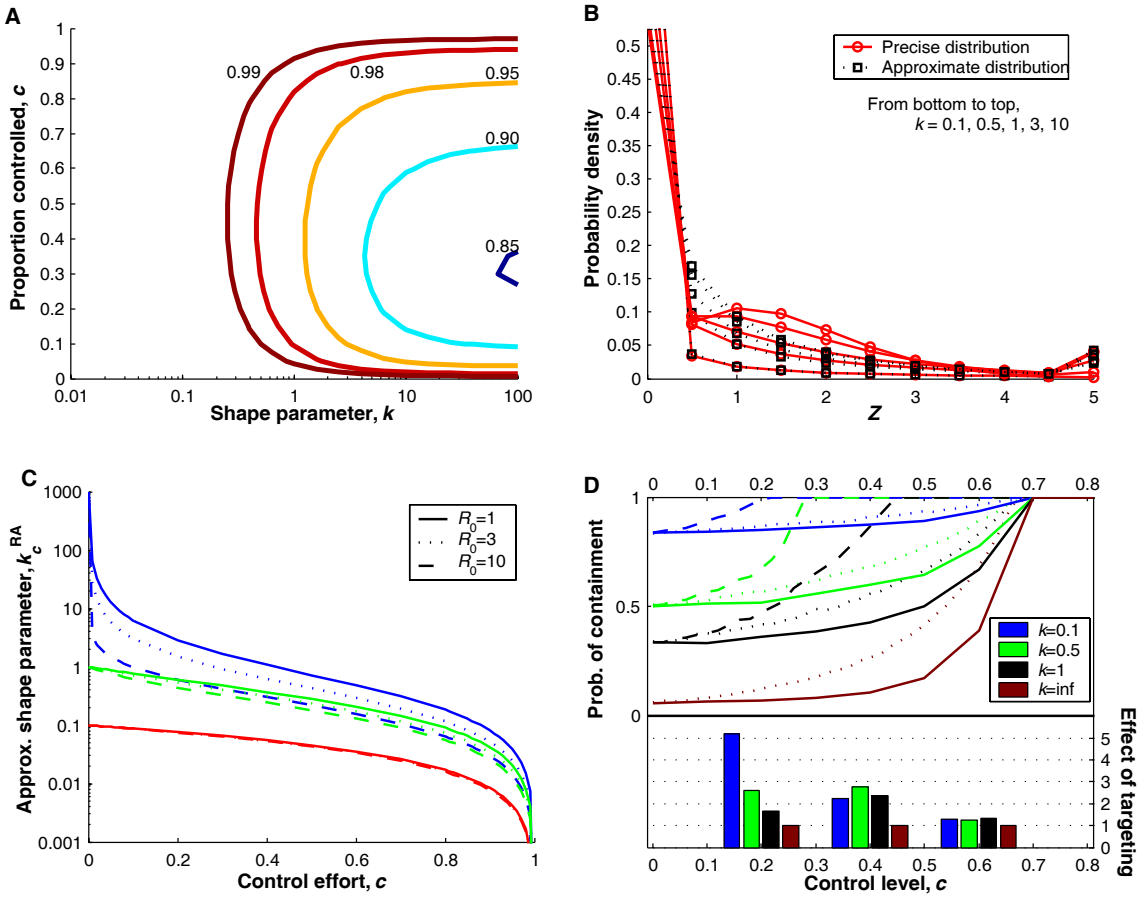
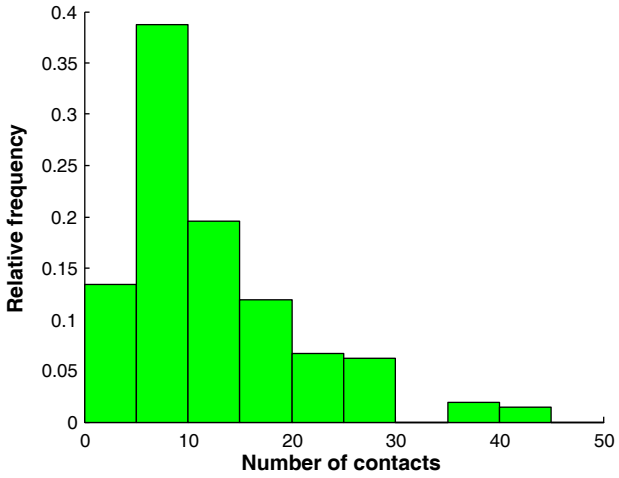

A
Proportion controlled, c
Shape parameter, k
0.99
0.98
0.95
0.90
0.85
B
Probability density
Z
Precise distribution
Approximate distribution
From bottom to top,
k = 0.1, 0.5, 1, 3, 10
C
Approx. shape parameter, k_c^RA
Control effort, c
R0=1
R0=3
R0=10
D
Prob. of containment
Effect of targeting
Control level, c
k=0.1
k=0.5
k=1
k=inf

Figure S4


Relative frequency
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
10
20
30
40
50
Number of contacts

Appendix B – OUTBREAK DATASETS AND SUPERSPREADING EVENTS

SARS, Singapore 2003 (Leo et al. 2003)

This dataset describes the progression of SARS in Singapore, beginning with the index case who imported the infection from Hong Kong. The first case had onset of symptoms on Feb 25, 2003. The government was notified of an unusual cluster of pneumonia cases on March 6, and again on March 14 for a cluster of six persons, including two HCWs, with atypical pneumonia. The first control measures were imposed on March 22, 2003. A case in the third generation had onset of symptoms on March 12, fully ten days before control measures were instituted. Given the mean serial interval of 8.4 days reported for Singapore (Lipsitch et al. 2003), and the inevitable delays to full implementation of control measures, we considered transmission by the first 3 generations to have been uncontrolled, and combine it into one dataset with N=57 in our analysis. Transmission data from the fourth through seventh generations were pooled to create the dataset under control measures, with N=114. Control measures imposed during this period included use of isolation and full contact precautions with all identified SARS patients, twice-daily screening of HCWs for fever, limitation of hospital visitors, and later the shutdown of a vegetable market where a SSE that occurred after control had been initiated. The dataset excludes 22 patients whose direct contacts were poorly defined, or who were cases translocated to Singapore (and the seven contacts of one of those cases). Note that the maximum-likelihood estimate of R_0 for Singapore (1.63; 90% CI (0.54,2.65)) is somewhat lower than estimated elsewhere (3.1; 95% CI (2.3,4.0)) (Wallinga and Teunis 2004). This may be because of excluded data, or because our dataset includes the period between the WHO's global alert on

SARS (March 12) and the imposition of formal control measures (March 22), during which time transmission may have been reduced by informal changes of behavior. Analysis of a dataset including only the first 2 generations of transmission in Singapore (N=22) yields $\hat{R}_{0,mle}$=2.55 (90% CI (0.50,4.50)) and $\hat{k}_{mle}$=0.21 (90% CI (0.15,∞)).

SARS, Beijing 2003 (Shen et al. 2004)

This dataset describes a hospital outbreak of SARS in the period before SARS was recognized in Beijing. The index case was an elderly woman hospitalized for diabetes, who caught SARS while a patient in the hospital. She infected 33 others, including many fellow patients. Second-generation cases included patients and visitors, and transmission by the second generation occurred in the hospital (to patients and visitors), in homes, and in a workplace. The hospital had not implemented isolation or quarantine procedures during the second generation's infectious period. Later in the outbreak administrative controls reduced contact rates, but infection control measures (masks, gloves, etc.) and respiratory isolation were never in place. We regard the first and second generations of spread as a natural experiment in SARS nosocomial transmission. To avoid concern of selection bias (i.e. that this outbreak was traced and reported because it began with an extraordinary superspreading event), we have removed the index case (Z=33) from our main analysis, and used only the Z values from the second generation cases (N=33) to calculate the values in Table 1. Analysis including the index case yields a more highly overdispersed distribution for ν ($\hat{R}_{0,mle}$=1.88, 90%CI (0.41,3.32); $\hat{k}_{mle}$=0.12, 90%CI (0.078,0.42)), as expected given the addition of an extreme SSE. The dataset under control was comprised of data from the

third and fourth generations of cases (*N*=43), after the hospital's imposition of limits on visitors and social contacts.

Smallpox (Variola major), Europe 1958-1973 (Fenner et al. 1988, p. 1077)

This dataset is a summary of smallpox importations into Europe from 1958-1973, and thus combines data collected over a long time period in many countries, probably with varying degrees of smallpox vaccination. Two outbreaks were excluded from the analysis, because one of them had three primary cases and the other had no primary case (infection was apparently transmitted on a carpet). The remaining outbreaks each had a single index case, and the number of infections in the first indigenous generation (i.e. cases within Europe) was taken as the *Z* value for each index case. Information on later generations is tabulated in the source material, but was excluded from this analysis because it was unclear if and when control was imposed in each outbreak, and there is no way to divide the total number of cases in the second indigenous generation among the possible source cases in the first indigenous generation.

Smallpox (Variola major), Benin 1967 (Henderson and Yekpe 1969)

A village-based outbreak occurred in Benin (formerly Dahomey) in 1967. The existence of the outbreak was concealed from authorities for three months, after which a vaccination team arrived but is suspected not to have affected the natural die-out of the outbreak. Contact tracing was by recollection of the villagers and some links are uncertain. Vaccination scar rates were <20% among children, and >70% among adults.

Transmission was predominantly by intimate contacts within households, rather than via frequent casual contacts among villagers. Limited control measures were imposed by the villagers, but were judged by the authors of the report to have had little effect on transmission so we have not divided the dataset.

Smallpox (Variola major), West Pakistan 1968-1970 (Heiner et al. 1971)

This is surveillance data from 47 outbreaks in rural West Pakistan, focusing on transmission within compounds inhabited by extended families. Of 47 outbreaks, 26 led to secondary transmission, with a total of 70 second-generation cases. Since all compound residents were in reasonably close contact, generations of cases were assigned based on the interval between exposure to the index case and onset of illness; for second generation cases this interval was 9-21 days. The population is reported to be relatively homogeneous. There was no isolation of contacts from cases, and vaccination is reported to have "played a minor role", though it was also observed that previously-vaccinated index cases tended to be less infectious. Severe illness was associated with higher infectiousness in this study. A similar study in East Pakistan in 1967 reported 30 smallpox outbreaks, with R~2.2 (stated verbally in the paper) and p_0=13/30, yielding an estimate of $\hat{k}_{pz}$=0.49 (Thomas et al. 1971).

Smallpox (Variola major), Kuwait 1967 (Arita et al. 1970)

In this outbreak, smallpox was suspected relatively quickly and control measures were imposed rapidly in the affected hospital. One unrecognized case had been transferred to another hospital, however, and initiated further spread there before

the disease was recognized and control was imposed. The outbreak was stopped by this expanded control effort. The background level of vaccination is not reported, but Kuwait had been free of endemic smallpox for a decade at the time of the outbreak. Control measures included intensive surveillance of hospitals suspected to be infected, with vaccination of all patients. Household contacts of infected individuals were vaccinated and placed under surveillance, and a mass vaccination campaign was initiated that covered 80% of the total population of Kuwait by the midway point of the outbreak (i.e. the date by which symptoms had appeared for roughly half of all cases).

Smallpox (Variola minor), England 1966 (Shooter 1980)

This outbreak of Variola minor, the less common and less severe form of smallpox, was initiated by a laboratory release in Birmingham, England. Because smallpox had been eliminated from England for decades, the outbreak went unsuspected until a case in the fourth generation of transmission was diagnosed and control efforts were initiated. Thorough investigations were conducted by British and US experts, but the results seem to have been published only as an appendix to a parliamentary inquiry into a 1978 release of smallpox from the same laboratory in Birmingham (Shooter 1980). The contact tracing dataset is quite complete, though there were several cases for whom a source of infection was not established. We have excluded the latter from our analysis. Vaccination levels in the general population were roughly 60% (Fenner et al. 1988, p. 1071).

Monkeypox, Zaire 1980-1984 (Jezek et al. 1987; Fine et al. 1988)

From 1980-1984, intensive surveillance and epidemiologic investigations were carried out in Zaire to monitor the risk of monkeypox emergence into the niche left empty by the recent eradication of smallpox. 147 monkeypox cases were judged to be primary cases infected by an animal source. These data are tabulated in several publications, with the greatest detail shown in Jezek et al (Jezek et al. 1987), who break down each outbreak by number of secondary cases per index case (Z) for each generation. In our analysis, we used the data for the first generation of human-to-human transmission only, to minimize the influence of control measures. Scars from smallpox vaccination (which is cross-protective for monkeypox) were seen on 68% of investigated contacts (Jezek and Fenner 1988, p. 99), but concern was expressed that vaccine protection may have been waning. Occasional instances of subclinical infection were reported, raising the possibility that these transmission figures are an underestimate (Jezek and Fenner 1988).

Pneumonic plague, 6 outbreaks 1907-1993 (Gani and Leach 2004)

Datasets from six outbreaks of pneumonic plague (*Yersinia pestis*) were compiled by Gani & Leach for their excellent recent analysis of the transmission and control of plague outbreaks. They employ an approach similar to ours, comparing Poisson and geometric models for the offspring distribution with aggregated data on Z (for all six datasets, before control measures), and conclude that the geometric distribution provides a superior fit. (Note that our analysis, while including the more flexible negative binomial distribution as a candidate model, also selected the geometric

model as the best combination of accuracy and parsimony in fitting the aggregated data (Table 1).) Because several of the source reports were published in inaccessible or foreign-language publications, we contacted Dr. Raymond Gani directly and he kindly provided the raw data from their analyses. We based our analysis of pneumonic plague on these data, with reference to the source report for the Mukden outbreak which we analyzed more closely in Table 2 (Tieh et al. 1948). Mukden is a city in Manchuria, China, which experienced a pneumonic plague outbreak in 1946 with 12 cases before control measures and 27 cases after the advent of control. Control measures included isolation and quarantine (in a suburban area) of all patients and contacts, disinfection and locking of infected houses, and wearing of masks required for all contacts and advised for the general population.

Avian H5N1 Influenza, Southeast Asia 2004 (Ferguson et al. 2004)

These data come from surveillance of the avian influenza epidemic in Southeast Asia, as summarized in April 2004. Of 33 confirmed H5N1 cases in humans, 31 were conclusively attributed to avian-to-human transmission, while 2 cases were possibly due to human-to-human transmission (from the same source case). While this human-to-human transmission was not confirmed, the pattern is consistent with recent reports of limited transmission of H5N1 avian influenza by prolonged intimate contact within families in Thailand ProMED-mail. Avian influenza, human - East Asia (48): Thailand. ProMED-mail 2004; 28 Sep: 20040928.2680. <http://www.promedmail.org>. Accessed 28 September 2004..

Rubella, Hawaii 1970 (Hattis et al. 1973)

In this outbreak, an army recruit returned to Hawaii from the US mainland for the Christmas holidays. He imported rubella, and proceeded to infect "every identified susceptible contact he had during the 72-hour period of his prodromal illness" (Hattis et al. 1973). His extreme infectiousness may have been linked to a persistent nonproductive cough linked to an earlier (separate) respiratory illness. The great majority of secondary cases did not cause further transmission; there was only one other infection event reported in the outbreak. Several cases were not epidemiologically linked to any source of transmission, and were omitted from the analysis. While this outbreak is probably exceptional in the extreme infectiousness of the index case, the authors of the original report conclude that highly heterogeneous infectiousness is necessary to explain observed patterns of rubella epidemiology in Hawaii. In particular, they posit that "During an uncomplicated rubella infection the average individual may have minimal contagious potential", while "Other persons may have a substantially greater potential for spread". Proposed factors influencing the potential for spread by individuals were age, sex, and coexisting or previous respiratory infections (the latter factor supported by unpublished evidence from military camps). "Spreader to spreader" contact is proposed to be necessary for sustained rubella transmission in a population, explaining why extended rubella outbreaks are most often observed in large, crowded population groups. The authors conclude that the proposed individual variation in infectiousness, combined with the sparse population distribution of Hawaii in the 1960s, could explain "why the highly susceptible population of Hawaii can encounter dozens and perhaps hundreds of rubella introductions each year without resulting in a full-scale

epidemic". This qualitative hypothesis is highly similar to the model-based conclusions reached in our study.

Hantavirus (Andes virus), Argentina 1996 (Wells et al. 1997)

This outbreak is the first reported instance of human-to-human transmission of a hantavirus, and is perhaps representative of a zoonotic pathogen beginning to adapt to a human host. It is definitely an anomalous pattern for hantavirus, as human-to-human transmission has not been reported elsewhere despite intensive surveillance. Contact tracing for this outbreak was imprecise, in part because several of the infected individuals had contact with more than one earlier case. The dataset of Z values analyzed was drawn from a diagrammed transmission chain and text descriptions in the outbreak report. In instances where the source of transmission was vague (i.e. transmission lines to two source cases in the published transmission chain), we adopted the conservative policy of dividing the secondary cases evenly between the possible sources in making our estimates of $\hat{R}_{0,mle}$ and $\hat{k}_{mle}$. The confidence intervals reported in Table 1 include the upper and lower bounds of 90% confidence intervals computed for all alternative assumptions regarding these vaguely attributed cases. There is no mention of control measures in the outbreak report, possibly because human-to-human transmission was not thought to be a threat.

Ebola Hemorrhagic Fever, Uganda 2000 (Francesconi et al. 2003)

These data come from a traced portion of a large outbreak (425 presumptive cases) from Aug 2000 to Jan 2001. The study methodology was retrospective contact

tracing, with the stated goal of determining the original "primary" cases of the outbreak (i.e. those who had acquired infection directly from the zoonotic reservoir). Cases (or their next of kin) were asked to identify persons from whom they had probably acquired the disease, who were in turn asked to identify who had infected them. Primary cases were defined as those whose sources of infection could not be identified. Prospective contact tracing was conducted to the extent that lists of contacts of identified cases (information that was "routinely collected") were matched with a list of reported cases. This data collection technique may bias the dataset toward surviving chains of transmission, since these are the ones that lead to the later-generation cases from which contact tracing began. The effort at prospective contact tracing would have mitigated this to some extent, but the level of tracing effort was certainly lower than for the retrospective work. The resulting dataset is conspicuously low in $Z=0$ entries, just as we would expect for a methodology that is biased against detecting chains that have died out.

Measles, US 1997-1999 (Gay et al. 2004)

In this summary of measles elimination efforts in the United States, 165 separate chains of measles transmission were identified (of which 107 were classified as importations). 122 outbreaks consisted of a single case with no secondary transmission (yielding an estimate of $p_0=122/165$). Insufficient data were reported to estimate R directly, but estimation of R was a major goal of the source paper so we used their estimate and 95% confidence interval. These estimates of R were derived from three approaches, all based on the assumption that $Z\sim\text{Poisson}(R)$. Our analysis shows that the

negative binomial offspring distribution is strongly favored by AIC_c model selection, but it is not clear what impact this would have on estimation of R using the methods described. We used the broadest confidence interval reported to account for this uncertainty. Vaccination levels in the US are reported to be above 90% in school-aged children (Papania and Wharton 2002), but are possibly lower in other populations.

Measles, Canada 1998-2001 (King et al. 2004)

As for the US measles dataset, this is routine surveillance data tracking progress on elimination of measles from Canada. 49 outbreaks were reported, of which 35 had only one case. Again we were unable to estimate R directly, and took estimates and confidence intervals (based on Z~Poisson(R)) from the source paper. The vaccination level in the general population is reported to be 95-100%. The authors raise the interesting point that long chains of transmission have occurred exclusively in religious communities that actively resist immunization, suggesting that an important determinant of the individual reproductive number ν in this context is the susceptibility of one's contacts.

Table S1. Superspreading events in the published literature

Disease	*Z*	Setting	Patient	Circumstances	Ref.
Ebola HF	46	Community	?M	Active social life, including workplace contacts; possibility of spread by injection (re-used needles).	(Smith et al. 1978)
Ebola HF	28-38+	Hospital	29M	"Popular" doctor, with many visitors during hospitalization before death.	(Khan et al. 1999)
Ebola HF	21+	Funeral	45F	Misdiagnosed, leading to traditional funeral with washing and handling of cadaver.	(Khan et al. 1999)

Influenza	38	Airplane	21F	All infections occurred aboard grounded airplane with ventilation system turned off for three hours; severe cough.	(Moser et al. 1979)
Lassa fever	16	Hospital	25F	Misdiagnosed; atypical presentation with severe cough. Possible airborne spread via air currents from bed to rest of ward.	(Carey et al. 1972)
Measles	69	High school	16F	Hacking cough; high school setting	(Chen et al. 1989)
Measles	84	High school	16M	Hacking cough; high school setting	(Chen et al. 1989)

Measles	250	Dance party	?M	First arrival of measles in Greenland—true virgin population. Index case attended crowded "dancing-lik" party.	(Christensen et al. 1953)
Mycoplasma pneumonia	26	Fraternity banquet	Unk.*	"Gross bacchanal" fraternity banquet: inebriation, cigar smoke membrane irritation, vomiting, shouting; participants "drenched with food missiles, drinks and gastric contents".	(Evatt et al. 1971)
Pneumonic plague	32	Funeral	?W	Funeral attenders and visitors of an unrecognized case.	(Hopkins et al. 1971)

Rubella	18	Home and parties	20M	Previous (ongoing) respiratory illness with cough.	(Hattis et al. 1973)
Rubella	37+	Discotheque	?M	Crowded discotheque; possible airborne spread via air flow from index case to crowd. Singing thought to aid aerosol-ization.	(Marks et al. 1981)
SARS	13	Hotel and hospital	64M	Undiagnosed: SARS not yet recognized.	(Tsang et al. 2003)
SARS	20	Hospital	47M	Undiagnosed: SARS not yet recognized.	(Kamps and Hoffmann 2004)

SARS	187+	Apartment block	26M	Amoy Gardens outbreak. Hypothesis: unsealed plumbing and bathroom fans led to aerosolized virus, infecting many in apartment complex.	(Yu et al. 2004)
SARS	21	Hospital	22?	Undiagnosed: SARS not yet recognized.	(Leo et al. 2003)
SARS	23	Hospital	27?	Undiagnosed: SARS not yet recognized. Patient was HCW infected nosocomially.	(Leo et al. 2003)
SARS	23	Hospital	53?	Patient infected nosocomially, co-morbidities.	(Leo et al. 2003)

SARS	40+	Hospital	60?	Misdiagnosed. Patient infected nosocomially, co-morbidities.	(Leo et al. 2003)
SARS	12	Vegetable market, hospital	64?	Misdiagnosed, with co-morbidities. Patient transmitted with minimal contact (e.g. twice to taxi drivers).	(Leo et al. 2003)
SARS	44	?	?	Co-morbidities.	(Kamps and Hoffmann 2004)
SARS	137	Hospital worker	43M	Co-morbidities; 'popular hospital laundry worker', continued work despite symptoms	(Kamps and Hoffmann 2004)

SARS	33	Hospital	62W	Undiagnosed: SARS not yet recognized. Patient infected nosocomially, with co-morbidities. High contact rate (many visitors) and no precautions in hospital.	(Shen et al. 2004)
SARS	10	Hospital	70W	Undiagnosed: SARS not yet recognized. Patient infected nosocomially, no precautions in hospital.	(Shen et al. 2004)

SARS	8	Hospital	69W	Undiagnosed: SARS not yet recognized. Patient infected nosocomially, no precautions in hospital.	(Shen et al. 2004)
SARS	12	Construction site	23M	High number of contacts at home and worksite.	(Shen et al. 2004)
SARS	19	Home, hospital	?M	Misdiagnosed due to unknown contact history, co-morbidities.	(Varia et al. 2003)

SARS	24/2*	Home, emergency room, ICU, hospital	?M	Unprotected exposure to index patient and wife of emergency personnel in ambulance, and of patients and staff in emergency room. Intubation procedure infected HCWs despite protective equipment.	(Varia et al. 2003)
Smallpox	19	?	?	No details available.	(Fenner et al. 1988, p.1077)
Smallpox	11	Social contacts	38M	Undiagnosed: smallpox not suspected. Visited with family and friends following travel abroad.	(Fenner et al. 1988, p.1092)

Smallpox	38	Hospital spread to HCWs and patients	30M	Undiagnosed: smallpox not suspected. Noted as interesting case and shown to students and staff in hospital.	(Fenner et al. 1988, p.1092)
Smallpox	16			Undiagnosed: mild ambulant case, not recognized as smallpox.	(Fenner et al. 1988, p.1908)
Smallpox	17	Hospital		Airborne spread despite "rigorous isolation"; aided by severe bronchitis, low humidity, and strong air currents	(Fenner et al. 1988, p.193)
Streptococcus group A (type 46)	10	Army barrack	?M	Asymptomatic case, with strongly positive nose and throat cultures.	(Hamburger et al. 1945)

Streptococcus group A (type 1)	100+	Hospital cafeteria	?M	Food handler with strongly positive nose culture and very high hand cultures; directly handled each piece of apple pie (popular item in cafeteria).	(Hamburger et al. 1945)
Tuberculosis	40/2*	Rock concert	?	2 index cases in rock band, infected “hundreds, if not thousands” of fans, at least 40 active cases. Airborne spread aided by singing.	(Houk 1980)
Tuberculosis	56		9M	Undiagnosed case, children not usually infectious with TB	(Curtis et al. 1999)

Table notes

Fractional entries in *Z* column denote more than one possible index case.

Patient column shows age and sex of index case, when known. (? indicates factor not reported in source literature.)

* index case not identified.

HCW: healthcare worker

References

Abbot, P. and L. M. Dill (2001). Sexually transmitted parasites and sexual selection in the milkweed leaf beetle, *Labidomera clivicollis*. Oikos **92**: 91-100.

Able, D. J. (1996). The contagion indicator hypothesis for parasite-mediated sexual selection. Proceedings of the National Academy of Sciences of the United States of America **93**: 2229-2233.

Anderson, R. M., S. P. Blythe, S. Gupta and E. Konings (1989). The transmission dynamics of the human immunodeficiency virus type-1 in the male homosexual community in the United Kingdom - the influence of changes in sexual behavior. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **325**: 45-98.

Anderson, D. R., K. P. Burnham and W. L. Thompson (2000). Null hypothesis testing: Problems, prevalence, and an alternative. Journal of Wildlife Management **64**: 912-923.

Anderson, R. M. and G. P. Garnett (2000). Mathematical models of the transmission and control of sexually transmitted diseases. Sexually Transmitted Diseases **27**: 636-643.

Anderson, R. M. and R. M. May (1991). *Infectious diseases of humans: dynamics and control*, Oxford University Press.

Antonovics, J., Y. Iwasa and M. P. Hassell (1995). A generalized model of parasitoid, venereal and vector-based transmission processes. American Naturalist **145**: 661-675.

Arita, I., E. Shafa and M. A. Kader (1970). Role of Hospital in Smallpox Outbreak in Kuwait. American Journal of Public Health **60**: 1960-1966.

Bailey, N. T. J. (1975). *The mathematical theory of infectious diseases*. London, Griffin.

Ball, F., D. Mollison and G. Scalia-Tomba (1997). Epidemics with two levels of mixing. The Annals of Applied Probability **7**: 46-89.

Bauch, C. and D. A. Rand (2000). A moment closure model for sexually transmitted disease transmission through a concurrent partnership network. Proceedings of the Royal Society of London Series B-Biological Sciences **267**: 2019-2027.

Beckage, N. E. (1997). *Parasites and pathogens: effects on host hormones and behavior*. New York, Chapman & Hall.

Becker, N. (1989). *Analysis of infectious disease data*. London, Chapman & Hall.

Becker, N. and I. Marschner (1990). The effect of heterogeneity on the spread of disease. in *Stochastic processes in epidemic theory: proceedings of a conference held in Luminy, France, October 23-29, 1988*. P. Picard, J. P. Gabriel and C. Lefevre, Eds. New York, Springer-Verlag. **86:** 90-103.

Begon, M., M. Bennett, R. G. Bowers, N. P. French, S. M. Hazel and J. Turner (2002). A clarification of transmission terms in host-microparasite models: numbers, densities and areas. Epidemiology and Infection **129**: 147-153.

Bjornstad, O. N., B. F. Finkenstadt and B. T. Grenfell (2002). Dynamics of measles epidemics: Estimating scaling of transmission rates using a Time series SIR model. Ecological Monographs **72**: 169-184.

Blanchard, J. F. (2002). Populations, pathogens, and epidemic phases: closing the gap between theory and practice in the prevention of sexually transmitted diseases. Sexually Transmitted Infections **78**: I183-I188.

Blower, S. M., H. B. Gershengorn and R. M. Grant (2000). A tale of two futures: HIV and antiretroviral therapy in San Francisco. Science **287**: 650-654.

Blower, S. M., T. C. Porco and G. Darby (1998). Predicting and preventing the emergence of antiviral drug resistance in HSV-2. Nature Medicine **4**: 673-678.

Booth, C. M., L. M. Matukas, G. A. Tomlinson, A. R. Rachlis, D. B. Rose, H. A. Dwosh, S. L. Walmsley, T. Mazzulli, M. Avendano, P. Derkach, I. E. Ephtimios, I. Kitai, B. D. Mederski, S. B. Shadowitz, W. L. Gold, L. A. Hawryluck, E. Rea, J. S. Chenkin, D. W. Cescon, S. M. Poutanen and A. S. Detsky (2003) Clinical features and short-term outcomes of 144 patients with SARS in the greater Toronto area. Journal of the American Medical Association **289**: 1-9.

Boots, M. and R. J. Knell (2002). The evolution of risky behaviour in the presence of a sexually transmitted disease. Proceedings of the Royal Society of London Series B-Biological Sciences **269**: 585-589.

Bowden, F. J. and G. P. Garnett (2000). *Trichomonas vaginalis* epidemiology: parameterising and analysing a model of treatment interventions. Sexually Transmitted Infections **76**: 248-256.

Carey, D. E., J. Casals, R. F. Addy, H. A. White, G. Stroh, A. Fom, L. Pinneo, Henderso.Be and G. E. Kemp (1972). Lassa Fever - Epidemiological Aspects of 1970 Epidemic, Jos, Nigeria. Transactions of the Royal Society of Tropical Medicine and Hygiene **66**: 402-408.

Castillo-Chavez, C. (1989). *Mathematical and statistical approaches to AIDS epidemiology*. Berlin, Springer-Verlag.

Caswell, H. (2001) Matrix Population Models, 2nd ed. Sinauer, Sunderland MA.

CDC (2002). Sexually Transmitted Disease Surveillance, 2001. Atlanta, GA, U.S Department of Health and Human Services.

Chen, R. T., G. M. Goldbaum, S. G. F. Wassilak, L. E. Markowitz and W. A. Orenstein (1989). An Explosive Point-Source Measles Outbreak in a Highly Vaccinated Population - Modes of Transmission and Risk-Factors for Disease. American Journal of Epidemiology **129**: 173-182.

Chowell, G., C. Castillo-Chavez, P. W. Fenimore, C. M. Kribs-Zaleta, L. Arriola and J. M. Hyman (2004). Model parameters and outbreak control for SARS. Emerging Infectious Diseases **10**: 1258-1263.

Chowell, G., N. W. Hengartner, C. Castillo-Chavez, P. W. Fenimore and J. M. Hyman (2004). The basic reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda. Journal of Theoretical Biology **229**: 119-126.

Christensen, P. E., H. Schmidt, H. O. Bang, V. Andersen, B. Jordal and O. Jensen (1953). An Epidemic of Measles in Southern Greenland, 1951 - Measles in Virgin Soil .2. The Epidemic Proper. Acta Medica Scandinavica **144**: 430-449.

Curtis, A. B., R. Ridzon, R. Vogel, S. McDonough, J. Hargreaves, J. Ferry, S. Valway and I. M. Onorato (1999). Extensive transmission of Mycobacterium tuberculosis from a child. New England Journal of Medicine **341**: 1491-1495.

Cyranoski, D. & A. Abbott (2003) Apartment complex holds clues to pandemic potential of SARS. Nature **423**: 3-4.

De Jong, M. C. M., O. Diekmann and J. A. P. Heesterbeek (1995). How does transmission of infection depend on population size? in *Epidemic models: their structure and relation to data*. D. Mollison, Ed., Cambridge University Press**:** 84-94.

Dieckmann, U., J. A. J. Metz, M. W. Sabelis and K. Sigmund, Eds. (2002). *Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management*. Cambridge Series in Adaptive Dynamics. Cambridge, Cambridge University Press.

Diekmann, O., K. Dietz and J. A. P. Heesterbeek (1991). The basic reproduction ratio for sexually-transmitted diseases 1. Theoretical considerations. Mathematical Biosciences **107**: 325-339.

Diekmann, O. and J. A. P. Heesterbeek (2000). *Mathematical epidemiology of infectious diseases: model building, analysis, and interpretation*. Chichester, Wiley.

Diekmann, O., J. A. P. Heesterbeek and J. A. J. Metz (1990). On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. Journal of Mathematical Biology **28**: 365-382.

Dietz, K. and K. P. Hadeler (1988). Epidemiological models for sexually-transmitted diseases. Journal of Mathematical Biology **26**: 1-25.

Donnelly, C.A., A.C. Ghani, G. M. Leung, A. J. Hedley, C. Fraser, S. Riley, L. J. Abu-Raddad, L. M. Ho, T. Q. Thach, P. Chau, K. P. Chan, T. H. Lam, L. Y. Tse, T. Tsang, S. H. Liu, J. H. B. Kong, E. M. C.Lau, N. M. Ferguson and R. M. Anderson (2003) Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. Lancet **361**: 1761-1766.

Donovan, B. (2000). The repertoire of human efforts to avoid sexually transmissible diseases: past and present - Part 1: Strategies used before or instead of sex. Sexually Transmitted Infections **76**: 7-12.

Drazen, J.M. (2003) Case clusters of the severe acute respiratory syndrome. New England Journal of Medicine **348**: e6-7.

Drosten C., S. Günther, W. Preiser, S. van der Werf, H. R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R. A. M. Fouchier, A. Berger, A. M. Burguiere, J. Cinatl, M. Eickmann, N. Escriou, K. Grywna, S. Kramme, J. C. Manuguerra, S. Muller, V. Rickerts, M. Sturmer, S. Vieth, H. D. Klenk, A. D. M. E. Osterhaus, H. Schmitz and H. W. Doerr (2003) Identification of a novel coronavirus in patients with severe acute respiratory syndrome. New England Journal of Medicine **348**:1967-1976.

Dwosh, H.A., H. H. L. Hong, D. Austgarden, S. Herman, R. Schabas (2003) Identification and containment of an outbreak of SARS in a community hospital. Canadian Medical Association Journal **168**: 1415-1420.

Dye, C. and N. Gay (2003). Modeling the SARS epidemic. Science **300**: 1884-1885.

Eames, K. T. D. and M. J. Keeling (2002). Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. Proceedings of the National Academy of Sciences of the United States of America **99**: 13330-13335.

Efron, B. and R. J. Tibshirani (1993). *An introduction to the bootstrap*. London, Chapman & Hall.

Eubank, S., H. Guclu, V. S. A. Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai and N. Wang (2004). Modelling disease outbreaks in realistic urban social networks. Nature **429**: 180-184.

Evatt, B. L., W. R. Dowdle, M. Johnson and C. W. Heath (1971). Epidemic Mycoplasma Pneumonia. New England Journal of Medicine **285**: 374-&.

Farrington, C. P., M. N. Kanaan and N. J. Gay (2003). Branching process models for surveillance of infectious diseases controlled by mass vaccination. Biostatistics **4**: 279-295.

Fenner, F., D. A. Henderson, I. Arita, Z. Jezek and I. D. Ladnyi (1988). *Smallpox and its eradication*. Geneva, World Health Organization.

Fenton, A., J. P. Fairbairn, R. Norman and P. J. Hudson (2002). Parasite transmission: reconciling theory and reality. Journal of Animal Ecology **71**: 893-905.

Ferguson, N. M., C. Fraser, C. A. Donnelly, A. C. Ghani and R. M. Anderson (2004). Public health risk from the avian H5N1 influenza epidemic. Science **304**: 968-969.

Ferguson, N.M., Galvani, A.P., Bush, R.M. (2003) Ecological and immunological determinants of influenza evolution. Nature **422**: 428-433.

Ferguson, N. M. and G. P. Garnett (2000). More realistic models of sexually transmitted disease transmission dynamics - Sexual partnership networks, pair models, and moment closure. Sexually Transmitted Diseases **27**: 600-609.

Fine, P. E. M., Z. Jezek, B. Grab and H. Dixon (1988). The Transmission Potential of Monkeypox Virus in Human-Populations. International Journal of Epidemiology **17**: 643-650.

Francesconi, P., Z. Yoti, S. Declich, P. A. Onek, M. Fabiani, J. Olango, R. Andraghetti, P. E. Rollin, C. Opira, D. Greco and S. Salmaso (2003). Ebola hemorrhagic fever transmission and risk factors of contacts, Uganda. Emerging Infectious Diseases **9**: 1430-1437.

Galvani, A.P., Lei, X., Jewell, N.P. (2003) Severe acute respiratory syndrome: temporal stability and geographic variation in death rates and doubling times. Emerging Infectious Diseases **9**: 991-994.

Gani, R. and S. Leach (2001). Transmission potential of smallpox in contemporary populations. Nature **414**: 748-751.

Gani, R. and S. Leach (2004). Epidemiologic determinants for modeling pneumonic plague outbreaks. Emerging Infectious Diseases **10**: 608-614.

Garnett, G. P. (2002). An introduction to mathematical models in sexually transmitted disease epidemiology. Sexually Transmitted Infections **78**: 7-12.

Garnett, G. P. (2002). The geographical and temporal evolution of sexually transmitted disease epidemics. Sexually Transmitted Infections **78**: I14-I19.

Garnett, G. P., K. J. Mertz, L. Finelli, W. C. Levine and M. E. St Louis (1999). The transmission dynamics of gonorrhoea: modelling the reported behaviour of infected patients from Newark, New Jersey. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **354**: 787-797.

Gay, N. J., G. De Serres, C. P. Farrington, S. B. Redd and M. J. Papania (2004). Assessment of the status of measles elimination from reported outbreaks: United States, 1997-1999. Journal of Infectious Diseases **189**: S36-S42.

Getz, W. M. and J. Pickering (1983). Epidemic models - thresholds and population regulation. American Naturalist **121**: 892-898.

Gold, R. S. and M. J. Skinner (1996). Judging a book by its cover: Gay men's use of perceptible characteristics to infer antibody status. International Journal of Std & Aids **7**: 39-43.

Grenfell, B. T., O. N. Bjornstad and J. Kappey (2001). Travelling waves and spatial hierarchies in measles epidemics. Nature **414**: 716-723.

Ha, L. D., S. A. Bloom, N. Q. Hien, S. A. Maloney, L. Q. Mai, K. C. Leitmeyer, B. H. Anh, I. G. Reynolds, J. M. Montgomery, J. A. Comer, P. W. Horby and A. J. Plant (2004). Lack of SARS transmission among public hospital workers, Vietnam. Emerging Infectious Diseases **10**: 265-268.

Hadeler, K. P. and C. Castillo-Chavez (1995). A core group model for disease transmission. Mathematical Biosciences **128**: 41-55.

Hamburger, M., M. J. Green and V. G. Hamburger (1945). The Problem of the Dangerous Carrier of Hemolytic Streptococci .2. Spread of Infection by Individuals with Strongly Positive Nose Cultures Who Expelled Large Numbers of Hemolytic Streptococci. Journal of Infectious Diseases **77**: 96-108.

Hamilton, W. D. and M. Zuk (1982). Heritable true fitness and bright birds - a role for parasites. Science **218**: 384-387.

Harris, T. E. (1989). *The theory of branching processes*. New York, Dover.

Hattis, R. P., S. B. Halstead, K. L. Herrmann and J. J. Witte (1973). Rubella in an Immunized Island Population. Journal of the American Medical Association **223**: 1019-1021.

Haydon, D. T., M. Chase-Topping, D. J. Shaw, L. Matthews, J. K. Friar, J. Wilesmith and M. E. J. Woolhouse (2003). The construction and analysis of epidemic trees with reference to the 2001 UK foot-and-mouth outbreak. Proceedings of the Royal Society of London Series B-Biological Sciences **270**: 121-127.

Health Canada (2003) Summary of Severe Acute Respiratory Syndrome (SARS) Cases: Canada and International: April 29, 2003. Accessed online at http://www.hc-sc.gc.ca/pphb-dgspsp/sars-sras/eu-ae/sars20030429_e.html

Heesterbeek, J. A. P. and J. A. J. Metz (1993). The saturating contact rate in marriage- and epidemic models. Journal of Mathematical Biology **31**: 529-539.

Heiner, G. G., N. Fatima and F. R. McCrumb (1971). Study of Intrafamilial Transmission of Smallpox. American Journal of Epidemiology **94**: 316-326.

Henderson, R. and M. Yekpe (1969). Smallpox Transmission in Southern Dahomey - a Study of a Village Outbreak. American Journal of Epidemiology **90**: 423-428.

Hethcote, H. W. (2000). The mathematics of infectious diseases. Siam Review **42**: 599-653.

Hethcote, H. W. and J. W. Van Ark (1987). Epidemiological models for heterogeneous populations: Proportionate mixing, parameter estimation, and immunization programs. Mathematical Biosciences **84**: 85-118.

Hethcote, H. W. and J. A. Yorke (1984). *Gonorrhea transmission dynamics and control*. Berlin, Springer-Verlag.

Hong Kong Department of Health (2003) Situation report on severe acute respiratory syndrome, June 9 2003. Accessed online at http://www.info.gov.hk/dh/new/2003/03-06-09e.htm

Hopkins, D. R., J. M. Lane, E. C. Cummings and J. D. Millar (1971). 2 Funeral-Associated Smallpox Outbreaks in Sierra-Leone. American Journal of Epidemiology **94**: 341-&.

Houk, V. N. (1980). Spread of tuberculosis via recirculated air in a naval vessel: the Byrd study. Annals of the New York Academy of Sciences **353**: 10-24.

Hsieh, Y. H. and S. P. Sheu (2001). The effect of density-dependent treatment and behavior change on the dynamics of HIV transmission. Journal of Mathematical Biology **43**: 69-80.

Hyman, J. M. and J. Li (1997). Behavior changes in SIS STD models with selective mixing. Siam Journal on Applied Mathematics **57**: 1082-1094.

Jansen, V. A. A., N. Stollenwerk, H. J. Jensen, M. E. Ramsay, W. J. Edmunds and C. J. Rhodes (2003). Measles outbreaks in a population with declining vaccine uptake. Science **301**: 804-804.

Jezek, Z. and F. Fenner (1988). *Human Monkeypox*. Basel, Karger.

Jezek, Z., B. Grab and H. Dixon (1987). Stochastic Model for Interhuman Spread of Monkeypox. American Journal of Epidemiology **126**: 1082-1092.

Kamps, B. S. and C. Hoffmann (2003). *SARS Reference*, 3rd ed. Flying Publisher. Accessed online at SARSReference.com.

Kavaliers, M., D. D. Colwell and E. Choleris (1999). Parasites and behavior: an ethopharmacological analysis and biomedical implications. Neuroscience and Biobehavioral Reviews **23**: 1037-1045.

Keeling, M. J. (1999). The effects of local spatial structure on epidemiological invasions. Proceedings of the Royal Society Biological Sciences Series B **266**: 859-867.

Keeling, M. J. and B. T. Grenfell (1998). Effect of variability in infection period on the persistence and spatial spread of infectious diseases. Mathematical Biosciences **147**: 207-226.

Keeling, M. J., M. E. J. Woolhouse, R. M. May, G. Davies and B. T. Grenfell (2003). Modelling vaccination strategies against foot-and-mouth disease. Nature **421**: 136-142.

Keeling, M. J., M. E. J. Woolhouse, D. J. Shaw, L. Matthews, M. Chase-Topping, D. T. Haydon, S. J. Cornell, J. Kappey, J. Wilesmith and B. T. Grenfell (2001). Dynamics of the 2001 UK foot and mouth epidemic: Stochastic dispersal in a heterogeneous landscape. Science **294**: 813-817.

Kennedy, C. E. J., J. A. Endler, S. L. Poynton and H. McMinn (1987). Parasite load predicts mate choice in guppies. Behavioral Ecology and Sociobiology **21**: 291-295.

Khan, A. S., F. K. Tshioko, D. L. Heymann, B. Le Guenno, P. Nabeth, B. Kerstiens, Y. Fleerackers, P. H. Kilmarx, G. R. Rodier, O. Nkuku, P. E. Rollin, A. Sanchez, S. R. Zaki, R. Swanepoel, O. Tomori, S. T. Nichol, C. J. Peters, J. J. Muyembe-Tamfum and T. G. Ksiazek (1999). The reemergence of Ebola hemorrhagic fever, Democratic Republic of the Congo, 1995. Journal of Infectious Diseases **179**: S76-S86.

Kiesecker, J. M., D. K. Skelly, K. H. Beard and E. Preisser (1999). Behavioral reduction of infection risk. Proceedings of the National Academy of Sciences of the United States of America **96**: 9165-9168.

King, A., P. Varughese, G. De Serres, G. A. Tipples and J. Waters (2004). Measles elimination in Canada. Journal of Infectious Diseases **189**: S236-S242.

Koopman, J. S. (2002). Modeling infection transmission - The pursuit of complexities that matter. Epidemiology **13**: 622-624.

Kretzschmar, M. (2000). Sexual network structure and sexually transmitted disease prevention - A modeling perspective. Sexually Transmitted Diseases **27**: 627-635.

Kretzschmar, M. and K. Dietz (1998). The effect of pair formation and variable infectivity on the spread of an infection without recovery. Mathematical Biosciences **148**: 83-113.

Kretzschmar, M., J. C. Jager, D. P. Reinking, G. Vanzessen and H. Brouwers (1994). The basic reproduction ratio R(0) for a sexually-transmitted disease in a pair formation model with 2 types of pairs. Mathematical Biosciences **124**: 181-205.

Kretzschmar, M. and M. Morris (1996). Measures of concurrency in networks and the spread of infectious disease. Mathematical Biosciences **133**: 165-195.

Kretzschmar, M., Y. van Duynhoven and A. J. Severijnen (1996). Modeling prevention strategies for gonorrhea and chlamydia using stochastic network simulations. American Journal of Epidemiology **144**: 306-317.

Langmuir, A. D. (1980). Changing concepts of airborne infection of acute contagious diseases: a reconsideration of classic epidemiological theories. Annals of the New York Academy of Sciences **353**: 35-44.

Lee, N., D. Hui, A. Wu, P. Chan, P. Cameron, G. M. Joynt, A. Ahuja, M. Y. Yung, C. B. Leung, K. F. To, S. F. Lui, C. C. Szeto, S. Chung and J. J. Y. Sung (2003). A major outbreak of severe acute respiratory syndrome in Hong Kong. New England Journal of Medicine **348**: 1986-1994.

Leo, Y. S., M. Chen, B. H. Heng, C. C. Lee, N. Paton, B. Ang, P. Choo, S. W. Lim, A. E. Ling, M. L. Ling, B. K. Tay, P. A. Tambyah, Y. T. Lim, G. Gopalakrishna, S. Ma, L. James, P. L. Ooi, S. Lim, K. T. Goh, S. K. Chew and C. C. Tan (2003). Severe acute respiratory syndrome - Singapore, 2003. Morbidity and Mortality Weekly Report **52**: 405-411.

Levin, S. A., B. Grenfell, A. Hastings and A. S. Perelson (1997). Mathematical and computational challenges in population biology and ecosystems science. Science **275**: 334-343.

Lipsitch, M., T. Cohen, B. Cooper, J. M. Robins, S. Ma, L. James, G. Gopalakrishna, S. K. Chew, C. C. Tan, M. H. Samore, D. Fisman and M. Murray (2003). Transmission dynamics and control of severe acute respiratory syndrome. Science **300**: 1966-1970.

Lloyd, A. L. (2001). Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods. Proceedings of the Royal Society of London Series B-Biological Sciences **268**: 985-993.

Lloyd-Smith, J. O., A. P. Galvani and W. M. Getz (2003). Curtailing transmission of severe acute respiratory syndrome within a community and its hospital. Proceedings of the Royal Society of London Series B-Biological Sciences **270**: 1979-1989.

Lloyd-Smith, W. M. Getz, H. V. Westerhoff (2004). Frequency-dependent incidence in models of sexually transmitted diseases: portrayal of pair-based transmission and effects of illness on contact behaviour. Proceedings of the Royal Society of London Series B-Biological Sciences **271**: 625-634.

Lockhart, A. B., P. H. Thrall and J. Antonovics (1996). Sexually transmitted diseases in animals: Ecological and evolutionary implications. Biological Reviews of the Cambridge Philosophical Society **71**: 415-471.

Loehle, C. (1995). Social barriers to pathogen transmission in wild animal populations. Ecology **76**: 326-335.

Longini, I. M., J. S. Koopman, A. S. Monto and J. P. Fox (1982). Estimating Household and Community Transmission Parameters for Influenza. American Journal of Epidemiology **115**: 736-751.

Mandavilli, A. (2003) SARS epidemic unmasks age-old quarantine conundrum. Nature Medicine **9**: 487.

Manly, B. F. J. (1998). *Randomization, bootstrap and Monte Carlo methods in biology*. London, Chapman & Hall.

Marks, J. S., M. K. Serdula, N. A. Halsey, M. V. H. Gunaratne, R. B. Craven, K. A. Murphy, G. Y. Kobayashi and N. H. Wiebenga (1981). Saturday Night Fever - a Common-Source Outbreak of Rubella among Adults in Hawaii. American Journal of Epidemiology **114**: 574-583.

Masuda, N., N. Konno and K. Aihara (2004). Transmission of severe acute respiratory syndrome in dynamical small-world networks. Physical Review E **69**: Art. No. 031917.

Maunder, R., J. Hunter, L. Vincent, J. Bennett, N. Peladeau, M. Leszcz, J. Sadavoy, L. M. Verhaeghe, R. Steinberg and T. Mazzulli (2003). The immediate psychological and occupational impact of the 2003 SARS outbreak in a teaching hospital. Canadian Medical Association Journal **168**: 1245-1251.

May, R. M. and R. M. Anderson (1987). Transmission dynamics of HIV infection. Nature **326**: 137-142.

McCallum, H., N. Barlow and J. Hone (2001). How should pathogen transmission be modelled? Trends in Ecology & Evolution **16**: 295-300.

McCallum, H., N. Barlow and J. Hone (2002). Modelling transmission: mass action and beyond - Response from McCallum, Barlow and Hone. Trends in Ecology & Evolution **17**: 64-65.

Mertz, G. J., J. Benedetti, R. Ashley, S. A. Selke and L. Corey (1992). Risk factors for the sexual transmission of genital herpes. Annals of Internal Medicine **116**: 197-202.

Moser, M. R., T. R. Bender, H. S. Margolis, G. R. Noble, A. P. Kendal and D. G. Ritter (1979). Outbreak of Influenza Aboard a Commercial Airliner. American Journal of Epidemiology **110**: 1-6.

Newshan, G., B. Taylor and R. Gold (1998). Sexual functioning in ambulatory men with HIV/AIDS. International Journal of STD & AIDS **9**: 672-676.

Papania, M. J. and M. Wharton (2002). Measles. VPD Surveillance Manual, Centers for Disease Control. Accessed online at http://www.cdc.gov/nip/publications/surv-manual/chpt06_measles.pdf.

Park, B. J., A. J. Peck, M. J. Kuehnert, C. Newbern, C. Smelser, J. A. Comer, D. Jernigan and L. C. McDonald (2004). Lack of SARS transmission among healthcare workers, United States. Emerging Infectious Diseases **10**: 244-248.

Peiris, J. S. M., C. M. Chu, V. C. C. Cheng, K. S. Chan, I. F. N. Hung, L. L. M. Poon, K. I. Law, B. S. F. Tang, T. Y. W. Hon, C. S. Chan, K. H. Chan, J. S. C. Ng, B. J. Zheng, W. L. Ng, R. W. M. Lai, Y. Guan and K. Y. Yuen (2003). Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. Lancet **361**: 1767-1772.

Penn, D. and W. K. Potts (1998). Chemical signals and parasite-mediated sexual selection. Trends in Ecology & Evolution **13**: 391-396.

Piot, P., M. Bartos, P. D. Ghys, N. Walker and B. Schwartlander (2001). The global impact of HIV/AIDS. Nature **410**: 968-973.

Porco, T. C., K. A. Holbrook, S. E. Fernyak, D. L. Portnoy, R. Reiter and T. J. Aragon (2004). Logistics of community smallpox control through contact tracing and ring vaccination: a stochastic network model. BMC Public Health **4**.

Poutanen, S. M., D. E. Low, B. Henry, S. Finkelstein, D. Rose, K. Green, R. Tellier, R. Draker, D. Adachi, M. Ayers, A. K. Chan, D. M. Skowronski, I. Salit, A. E. Simor, A. S. Slutsky, P. W. Doyle, M. Krajden, M. Petric, R. C. Brunham and A. J. McGeer (2003). Identification of severe acute respiratory syndrome in Canada. New England Journal of Medicine **348**: 1995-2005.

Rao, A. R., E. S. Jacob, Kamalaks.S, Appaswam.S and Bradbury (1968). Epidemiological Studies in Smallpox . A Study of Intrafamilial Transmission in a Series of 254 Infected Families. Indian Journal of Medical Research **56**: 1826-1854.

Reilley, B., M. Van Herp, D. Sermand, N. Dentico (2003) SARS and Carlo Urbani. New England Journal of Medicine **348**: 1951-1952.

Rice, J. A. (1995). *Mathematical statistics and data analysis*. Belmont CA, Duxbury Press.

Riley, S., C. Fraser, C. A. Donnelly, A. C. Ghani, L. J. Abu-Raddad, A. J. Hedley, G. M. Leung, L. M. Ho, T. H. Lam, T. Q. Thach, P. Chau, K. P. Chan, P. Y. Leung, T. Tsang, W. Ho, K. H. Lee, E. M. C. Lau, N. M. Ferguson and R. M. Anderson (2003). Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions. Science **300**: 1961-1966.

Schiltz, M. A. and T. G. M. Sandfort (2000). HIV-positive people, risk and sexual behaviour. Social Science & Medicine **50**: 1571-1588.

Seto, W. H., D. Tsang, R. W. H. Yung, T. Y. Ching, T. K. Ng, M. Ho, L. M. Ho and J. S. M. Peiris (2003). Effectiveness of precautions against droplets and contact in prevention of nosocomial transmission of severe acute respiratory syndrome (SARS). Lancet **361**: 1519-1520.

Shen, Z., F. Ning, W. G. Zhou, X. He, C. Y. Lin, D. P. Chin, Z. H. Zhu and A. Schuchat (2004). Superspreading SARS events, Beijing, 2003. Emerging Infectious Diseases **10**: 256-260.

Shooter, R. A. (1980). Report of the investigation into the cause of the 1978 Birmingham smallpox occurrence. London, H.M. Stationery Office**:** 108-134.

Smith, D. H., D. Francis, D. I. H. Simpson and R. B. Highton (1978). The Nzara outbreak of viral haemorrhagic fever. in *Ebola Virus Haemorrhagic Fever*. S. B. Pattyn, Ed., Elsevier**:** 137-141.

Smith, D. L., B. Lucey, L. A. Waller, J. E. Childs and L. A. Real (2002). Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes. Proceedings of the National Academy of Sciences of the United States of America **99**: 3668-3672.

Starks, P., C. Blackie and W. M. Getz (2000). The effect of HIV infections on human sexual behavior. *XIII International AIDS Conference*, Durban, South Africa.

Sullivan, A. D., J. Wigginton and D. Kirschner (2001). The coreceptor mutation CCR5 Delta 32 influences the dynamics of HIV epidemics and is selected for by HIV. Proceedings of the National Academy of Sciences of the United States of America **98**: 10214-10219.

Taylor, H. M. and S. Karlin (1998). *An introduction to stochastic modeling*. San Diego, Academic Press.

Thomas, D. B., I. Arita, McCormac.Wm, M. M. Khan, S. Islam and T. M. Mack (1971). Endemic Smallpox in Rural East Pakistan II. Intravillage Transmission and Infectiousness. American Journal of Epidemiology **93**: 373-383.

Thrall, P. H. and J. Antonovics (1997). Polymorphism in sexual versus non-sexual disease transmission. Proceedings of the Royal Society of London Series B-Biological Sciences **264**: 581-587.

Tieh, T. H., E. Landauer, F. Miyagawa, G. Kobayashi and G. Okayasu (1948). Primary Pneumonic Plague in Mukden, 1946, and Report of 39 Cases with 3 Recoveries. Journal of Infectious Diseases **82**: 52-58.

Tsang, K. W., P. L. Ho, G. C. Ooi, W. K. Yee, T. Wang, M. Chan-Yeung, W. K. Lam, W. H. Seto, L. Y. Yam, T. M. Cheung, P. C. Wong, B. Lam, M. S. Ip, J. Chan, K. Y. Yuen and K. N. Lai (2003). A cluster of cases of severe acute respiratory syndrome in Hong Kong. New England Journal of Medicine **348**: 1977-1985.

Tsang, T., T. Lai-Yin, L. Pak-Yin, M. Lee, J. S. Wu, Y. C. Wu, I. H. Chiang, K. T. Chen, K. H. Hsu, T. J. Chen, L. T. Lee, S. J. Twu, S. Chunsuttiwat, P. Sawanpanyalert, K. Ungchusak, A. Cahaovavanich and S. L. Roy (2003). Update: Outbreak of severe acute respiratory syndrome --- Worldwide, 2003. Morbidity and Mortality Weekly Report **52**: 241-248.

Twu, S. J., T. J. Chen, C. J. Chen, S. J. Olsen, L. T. Lee, T. Fisk, K. H. Hsu, S. C. Chang, K. T. Chen, I. H. Chiang, Y. C. Wu, J. S. Wu and S. F. Dowell (2003). Control measures for severe acute respiratory syndrome (SARS) in Taiwan. Emerging Infectious Diseases **9**: 718-720.

Varia, M., S. Wilson, S. Sarwal, A. McGeer, E. Gournis, E. Galanis and B. Henry (2003). Investigation of a nosocomial outbreak of severe acute respiratory syndrome (SARS) in Toronto, Canada. Canadian Medical Association Journal **169**: 285-292.

Waldstatter, R. (1989). Pair formation in sexually transmitted diseases. in *Mathematical and statistical approaches to AIDS epidemiology*. C. Castillo-Chavez, Ed. Berlin, Springer-Verlag**:** 260-274.

Wallinga, J., W. J. Edmunds and M. Kretzschmar (1999). Perspective: Human contact patterns and the spread of airborne infectious diseases. Trends in Microbiology **7**: 372-377.

Wallinga, J. and P. Teunis (2004). Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. American Journal of Epidemiology **160**: 509-516.

Webster, J. P., J. I. Hoffman and M. Berdoy (2003). Parasite infection, host resistance and mate choice: battle of the genders in a simultaneous hermaphrodite. Proceedings of the Royal Society of London Series B-Biological Sciences **270**: 1481-1485.

Wells, R. M., S. S. Estani, Z. E. Yadon, D. Enria, P. Padula, N. Pini, J. N. Mills, C. J. Peters, E. L. Segura, N. Guthmann, E. Arguelo, F. Klein, R. Levy, C. Nagel, R. Calfin, F. deRosas, M. Lazaro, H. Rosales and P. Sandoval (1997). An unusual hantavirus outbreak in southern Argentina: Person-to-person transmission? Emerging Infectious Diseases **3**: 171-174.

Wenzel, R.P. & Edmond, M.B. (2003) Managing SARS amidst uncertainty. New England Journal of Medicine **348**:1947-1948.

Woolhouse, M. E. J., C. Dye, J. F. Etard, T. Smith, J. D. Charlwood, G. P. Garnett, P. Hagan, J. L. K. Hii, P. D. Ndhlovu, R. J. Quinnell, C. H. Watts, S. K. Chandiwana and R. M. Anderson (1997). Heterogeneities in the transmission of infectious agents: Implications for the design of control programs. Proceedings of the National Academy of Sciences of the United States of America **94**: 338-342.

World Health Organization (2003a) China daily report of SARS cases. Accessed online at http://www.who.int/csr/sars/china2003_06_11.pdf

World Health Organization (2003b) Update 58 - First global consultation on SARS epidemiology, travel recommendations for Hebei Province (China), situation in Singapore. Accessed online at http://www.who.int/csr/sars/archive/2003_05_17/en/

Yu, I. T. S., Y. G. Li, T. W. Wong, W. Tam, A. T. Chan, J. H. W. Lee, D. Y. C. Leung and T. Ho (2004). Evidence of airborne transmission of the severe acute respiratory syndrome virus. New England Journal of Medicine **350**: 1731-1739.