Dynamically Modeling SARS and Other Newly Emerging Respiratory Illnesses: Past, Present, and Future

Author(s): Chris T. Bauch, James O. Lloyd-Smith, Megan P. Coffee and Alison P. Galvani

Source: *Epidemiology*, Vol. 16, No. 6 (Nov., 2005), pp. 791-801
Published by: Lippincott Williams & Wilkins
Stable URL: https://www.jstor.org/stable/20486145
Accessed: 27-03-2020 01:22 UTC

Dynamically Modeling SARS and Other Newly Emerging Respiratory Illnesses

Past, Present, and Future

Chris T. Bauch, James O. Lloyd-Smith,† Megan P. Coffee,‡ and Alison P. Galvani§*

Abstract: The emergence and rapid global spread of the severe acute respiratory syndrome (SARS) coronavirus in 2002–2003 prompted efforts by modelers to characterize SARS epidemiology and inform control policies. We overview and discuss models for emerging infectious diseases (EIDs), provide a critical survey of SARS modeling literature, and discuss promising future directions for research. We reconcile discrepancies between published estimates of the basic reproductive number R_0 for SARS (a crucial epidemiologic parameter), discuss insights regarding SARS control measures that have emerged uniquely from a modeling approach, and argue that high priorities for future modeling of SARS and similar respiratory EIDs should include informing quarantine policy and better understanding the impact of population heterogeneity on transmission patterns.

(*Epidemiology* 2005;16: 791–801)

The outbreak of severe acute respiratory syndrome (SARS) coronavirus provided a warning of the vulnerability of the human population to emerging diseases and was in many ways paradigmatic of emerging infectious diseases (EIDs) in the modern era. Air travel led to rapid global spread of the disease, yet electronic communication and the leadership of the World Health Organization facilitated an unprecedented international and interdisciplinary response. At the same time, action had to be taken based on sparse and evolving knowledge about a novel pathogen, and so primary methods for SARS control were the age-old measures of quarantine, case isolation, and infection control. Puzzling aspects of SARS data suggested a complex epidemiology: major outbreaks occurred in some cities but not others, and certain individuals or situations generated extraordinarily large numbers of secondary cases—the so-called superspreading events (SSEs). Questions abounded and mathematical modeling played an important role in understanding the complexities of this new disease.

In the 18 months since the SARS coronavirus was identified, more than 14 modeling studies of the 2002–2003 SARS outbreak have been published (Table 1).[1–14] The models developed during and after the SARS epidemic varied widely in terms of populations studied, motivating questions, design, and, occasionally, conclusions. While the epidemic was in progress, models were produced to predict whether a global pandemic was possible and whether control measures being applied were adequate.[1,2] Subsequent work has sought to discern general principles of EID epidemiology or to derive lessons for the next emergence of a SARS-like disease.[3,15–18]

We review this burgeoning literature, summarizing major results relevant to possible SARS reemergence or to other EIDs arising through zoonosis or bioterrorism. We emphasize dynamic (ie, mathematical or mechanistic) models that explicitly incorporate epidemiologic mechanisms (Table 2) rather than statistical models. We begin with a brief introduction to modeling techniques for emerging diseases. We then review the SARS modeling literature along 2 broad themes relevant to an unfolding epidemic: understanding the basic epidemiology of the disease and evaluating control strategies. We conclude with a general discussion on directions for further research in SARS and EID modeling.

MODELS OF EMERGING INFECTIOUS DISEASES

Modeling plays an important role in gaining insights into infectious disease epidemiology and in designing

Submitted 10 November 2004; accepted 13 April 2005.

From the *Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada; the †Department of Environmental Science, Policy and Management, University of California, Berkeley, California; ‡Harvard School of Medicine, Harvard University, Boston, Massachusetts; and the §Department of Epidemiology and Public Health, Yale School of Medicine, Yale University, New Haven, Connecticut.

Supported by CTB: a NSERC Discovery Grant (to CTB) and NSF-NIH Ecology of Infectious Disease grant DEB-0090323 and NIH-NIDA grant R01-DA10135 (to JOL-S).

Correspondence: Chris T. Bauch, Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada N1G 2W1. E-mail: cbauch@uoguelph.ca.

ISSN: 1044-3983/05/1606-0791
DOI: 10.1097/01.ede.0000181633.80269.4c

TABLE 1. Descriptions of Published SARS Models

Study	Population	Model Design	Focus	Caveats	Conclusions
Riley et al[1]	Hong Kong	Stochastic metapopulation compartmental model with hospitalized and infectious presymptom stages	Estimation of R in absence of superspreading events (SSEs); control measures; nosocomial transmission	SSEs not taken into account in estimation of R	Transmission fell during course of outbreak as a result of control measures and reduced contact; movement restrictions can be effective; hospital transmission is significant
Lipsitch et al[2]	Hong Kong, Singapore	Both deterministic and stochastic compartmental models with quarantine, isolation	Estimation of R and T; assessing control measures; probability of SARS invasion	Assumes homogeneous mixing	Need to apply multiple control measures; more variance in secondary infections, and lower R, increase extinction probability of incipient outbreaks
Lloyd-Smith et al[3]	Generic hospital/ community populations	Stochastic compartmental model with several control measures; community and hospital transmission	Emphasis on risks of nosocomial transmission and on assessing control measures	Hypothetical population; only healthcare workers (not patients or visitors) could transmit	Hospital-based precautions are most potent; combined effects of control measures can be nonlinear and unintuitive (so modeling is key); timeliness is important
Chowell et al[4]	Hong Kong, Toronto, Singapore	Deterministic compartmental model with 2 possible susceptibility levels	Population heterogeneity in susceptibility, assessing control measures	Some parameters set arbitrarily; assumes nonnegligible transmission during exposed period	Ability to control outbreak is most sensitive to effectiveness of isolation and rate of diagnosis
Hsieh et al[5]	Taiwan	Deterministic compartmental model (with discrete time steps)	Determining how best to improve control of nosocomial transmission	Simplified representation of transmission process; neglects change in transmission before and after controls	Best ways to control outbreaks are more efficient diagnosis and reclassification procedures, and reduced time to isolation
Ng et al[6]	Hong Kong, Beijing, Inner Mongolia	Deterministic compartmental model with 2 viral strains	Observed outbreak pattern in Hong Kong can only be explained by assuming 2 strains	No data confirming existence of 2 strains in humans; unclear if outbreak data can really be interpreted as 2 epidemics	A simple susceptible–infected–recovered (SIR) model does not apply to SARS; data suggest circulation of a second SARS precursor variant
Choi and Pak[7]	Toronto, Hong Kong	Discrete-time model with simple exponential growth of cases	Explanation of modeling basics for health experts	Analysis of data and model output does not exploit known and proven techniques (eg, least-squares fitting)	Simple models are didactic and can quickly provide rough estimates of important parameters

(*Continued*)

TABLE 1. (*Continued*)

Study	Population	Model Design	Focus	Caveats	Conclusions
Nishiura et al[8]	Japan	Deterministic compartmental model with quarantine, isolation	Seeks to explain why some countries did not experience a SARS outbreak	Limited sensitivity analysis; deterministic model applied to stochastic invasion	Initial number of imported cases is an important determinant of outbreak probability
Wang and Ruan[9]	Beijing	Simplified deterministic compartmental model	Parameter estimation and assessment of control measures	No sensitivity analysis; phenomenologic analysis; deterministic treatment of invasion	Applying control measures early is important in order to avoid endemic persistence
Zhou and Yan[10]	Singapore, Hong Kong, Beijing	Logistic curve fitted to data	Curve-fitting, prediction of end point of outbreak	Neglects change in transmission before and after application of controls	Logistic models can predict cumulative numbers of infected persons and end point of outbreak
Masuda et al[11]	N/A	Contact network model	SSEs, spatial effects, and social networks	Little application to real outbreaks or realistic parameterization	Social network structure impacts spread: highly connected SSEs are crucial
Meyers et al[12]	Vancouver	Contact network model	Understanding heterogeneity in SARS transmission (SSEs, geographic variation in outbreak occurrences)	Network is static and does not change during course of epidemic	Network structure, and the location of index cases within a network, can influence size of outbreaks and chances of an epidemic occurring
Wallinga and Teunis[13]	Hong Kong, Vietnam, Singapore, Canada	Likelihood-based estimation procedure for approximating chain of transmission	Estimation of the effective reproductive number R	Some limitations apply for initial cases when date of symptom onset is less than generation time	R dropped below 1 once the global SARS alert was issued, indicating effectiveness of control measures
Gumel et al[14]	Toronto, Hong Kong, Singapore, Beijing	Deterministic compartmental model with quarantine, isolation	Assessing the efficacy of control measures	Assumes homogeneous mixing; numeric estimates subject to strong assumptions	A perfect isolation policy alone is sufficient to control SARS, with or without quarantine; resources should be devoted disproportionately to isolation programs

NA indicates not applicable.

control strategies, whether for novel infectious agents such as SARS or known pathogens that are reemerging under new circumstances. There are many instances in which EID epidemiology or health policy has been informed by models. For example, models aided in controlling the foot-and-mouth outbreak in Britain in 2001,[19–21] understanding the hazard posed by monkeypox after smallpox eradication,[22,23] forecasting HIV trends,[24] analyzing patterns of reemergence of diseases such as dengue,[25] and assessing threats of bioterrorism from infectious agents such as anthrax or smallpox.[26–28]

Modeling allows for estimation of epidemiologic parameters from data, identification of likely mechanisms underlying observed patterns, assessment of the relative merits of alternative control strategies, and prediction of epidemiologic or evolutionary dynamics. The first step in designing a modeling study requires identification of the questions being asked followed by development of a model appropriate to

TABLE 2. Glossary of Terms Used in Dynamic Modeling

Dynamic (mechanistic/mathematical) model	A model that explicitly incorporates known or hypothesized epidemiologic mechanisms, allowing for prediction of future temporal development under various scenarios (eg, individual-based simulation model)
Phenomenologic model	A model designed to reproduce trends observed in data without specifying the underlying mechanisms (eg, fitting a curve to data)
Compartmental model	An epidemic model that divides the population into mutually exclusive categories (compartments) based on disease status, age, risk factor, and so on; often abbreviated by letters denoting the compartments included in the model (SIR, SEIR, and so on; see Table 1), eg, susceptible (S), infected but not infectious (E), infectious (I), recovered (R)
Deterministic model	A model in which the future can be unambiguously determined based on precise knowledge of the present state of the system, eg, ordinary differential equation models
Stochastic model	A model in which some or all events are governed by chance (and hence not deterministic); multiple outcomes are possible for a given starting state of the system
Parameter	An independent quantity that determines some aspect of the system's behavior (eg, duration of infectiousness)
Variable	A dependent quantity that evolves over time in response to the initial state of the system, values of model parameters, and stochastic effects, if any
Basic reproductive number (R_0)	The average number of secondary infections produced by a typical infected individual in an otherwise susceptible population; R_0 is often computed as a threshold quantity derived from a dynamic model; in homogeneously mixing populations, this threshold quantity agrees exactly with the above definition of R_0; however, in heterogeneously mixing populations, they may differ[59]
Effective reproductive number (R)	Average number of secondary infections produced by a typical infected individual
Control reproductive number (R_c)	The effective reproductive number in the presence of control measures
Generation time/serial interval (T)	The average time between the onset of symptoms in a given infected individual and the onset of symptoms in individuals that person has infected
Theta (θ)	Proportion of transmission occurring before an individual exhibits symptoms

those questions. A broad range of modeling approaches may be adopted (Tables 1 and 2).

When developing a model, one must decide how much detail to incorporate. A highly detailed, realistic model makes all assumptions explicit and affords greater opportunities to investigate governing mechanisms and assess control strategies. For example, to evaluate restrictions on travel, it is necessary to incorporate spatial structure into the model.[1] However, greater detail is not an unqualified good. Model results are often sensitive to parameters in ways that may be difficult to foresee or understand; therefore, setting parameters arbitrarily, without reference to data, is hazardous.[29] This danger only increases with the number of parameters to be estimated. The problem can be ameliorated by analyzing the sensitivity of model predictions to parameters whose values are not known precisely.[30] However, this process of sensitivity analysis is often neglected in practice. An even greater challenge is to evaluate sensitivity to model structure, ie, to different possible ways of distilling a complex reality into an idealized modeling framework. Simpler models, by contrast, require fewer estimated parameters and are therefore easier to analyze. However, the resulting phenomenologic description of the system imparts less ability to understand governing mechanisms and assess control measures. The implicit simplifying assumptions involved can also introduce inaccuracies (although this is not necessarily the case). Selection of the most appropriate level of detail depends on the questions being asked. Comparing results from independent modeling studies with differing levels of detail can be fruitful—in a sense, the ultimate structural sensitivity analysis—and we carry out such a comparison in this review.

Another important distinction for modeling studies of EIDs is between deterministic and stochastic models (Table 2). Both disease transmission between individuals and disease progression within individuals are probabilistic by nature. In the beginning stages of an outbreak, the number of infected individuals is very small, and so chance events play an important role in transmission dynamics. As a result of chance, diseases with epidemic potential can die out, or they can generate more infections than expected. Deterministic models cannot capture this variation, and in the early stages of an outbreak, such models underestimate both the probability of extinction and the probability of an unusually rapid outbreak. In some cases, chance events only serve to produce

variance around a population average that is accurately predicted by a deterministic model, whereas in other cases, the average behavior of a stochastic model differs from the deterministic prediction. Deterministic models are not necessarily inappropriate for studying invasion, but they should be applied cautiously and with an appreciation for stochastic effects.

In tackling SARS, researchers have taken a variety of approaches, ranging from the simplest possible epidemic models to sophisticated models incorporating various population heterogeneities (Table 1). In the following sections, we review the studies in detail.

CHARACTERIZING EPIDEMIOLOGIC PROPERTIES OF EMERGING INFECTIOUS DISEASES

Forecasting the course of an epidemic and the best strategies for its containment can often be based on a handful of epidemiologic quantities: the basic reproductive number (R_0),[31] the degree of individual heterogeneity in infectiousness, the generation time (T), and the proportion of transmission occurring before the symptomatic period (θ).[32] Estimating these parameters is one of the first challenges when a new infectious disease emerges.

The Basic Reproductive Number (R_0)

The *basic reproductive number* is the average number of secondary infections produced by a typical infected individual in a wholly susceptible population (Table 2).[31,33] If R_0 is less than one, each infection will not replace itself on average, and the outbreak will not be sustained. If R_0 exceeds one, then the number of infected individuals multiplies and an epidemic may ensue. As the epidemic spreads, the proportion of the population that is susceptible declines. Thus, in a homogeneously mixed population, infectious individuals begin to "waste" contacts on individuals that are already infected or immune, and so the number of new infections produced per infected individual also declines. Here, it becomes useful to define the *effective reproductive number* (R) as the average number of secondary infections produced by a typical infected individual (in a population with some proportion of susceptible individuals). In a homogeneously mixed population, $R = fR_0$, in which f is the proportion of susceptible individuals at a given point in time. (With heterogeneous mixing, R declines even faster.[34]) Initially, $R = R_0$, but R declines over time as susceptibles are depleted. Around the epidemic peak, $R = 1$, and thereafter, the epidemic starts to burn out. It is also useful to distinguish the *control reproductive number* (R_c) as the value of R in the presence of control measures. [14] If R_c can be sustained at values below one, then the disease will eventually be eradicated.

R_0 is a population-average quantity that neglects variation in the number of secondary cases produced per infected individual. However, such variation was often evident in the initial chain of SARS transmission before control measures were in place. Most infected individuals did not transmit the infection further, but in a few cases, known as *superspreading events* (SSEs), a small number of infected individuals produced a disproportionately large number of secondary cases. Lipsitch et al[2] have shown that, for a given R, a greater variance in secondary cases per index case reduces the probability of an epidemic. Lloyd-Smith et al[34a] have shown that individual variation is significant for at least 10 important EIDs. Using a stochastic modeling framework, they have demonstrated that greater variation leads to less frequent but more explosive outbreaks. Thus, in addition to knowing R_0, some measure of variance in secondary transmission is an important measure in classifying the tenacity of a nascent epidemic.

Computation of R_0 for SARS

Model-free estimates of R_0 can be determined from mapping out the chain of transmission early in the epidemic. However, these data are often not comprehensive, and so R_0 is usually estimated by fitting incidence time series to an epidemic model. Almost all studies reviewed here used this latter method. These studies produced considerable variation in estimates of R_0 (or R) for SARS, even for the same populations (Table 1, Fig. 1). Understanding the source of these discrepancies is clearly important, and a critical comparison of modeling studies can reconcile such differences.

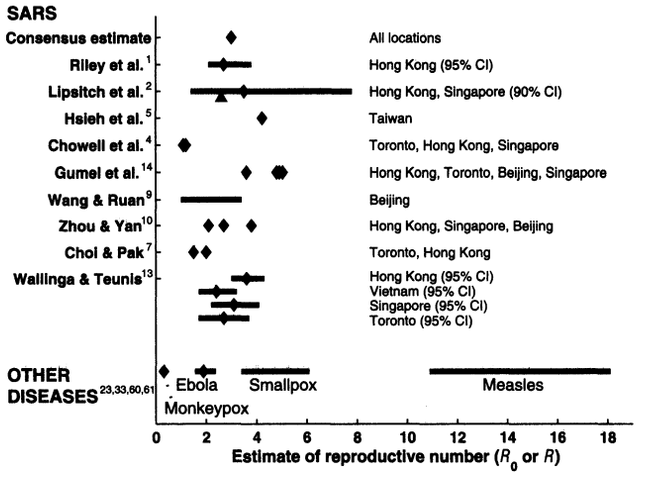

FIGURE 1. R_0 estimates from various studies for SARS and other diseases.[1–14,23,33,60,61] "Consensus estimate" refers to the SARS R_0 estimate (approximately 3) arrived at in this article by critically comparing independent SARS studies. For Lipsitch et al,[1] the triangle denotes their best deterministic estimate, whereas the diamond denotes the mean of their Bayesian estimate. The interval for Wang and Ruan[9] indicates a range of possible values depending on assumptions made about the generation time (T).

There are at least 3 possible sources for such discrepancies: different model structure, different input parameter values, and different estimation procedures (eg, curve-fitting vs Bayesian statistics). For the case of SARS, studies with radically different model structures often came up with comparable estimates for R_0. For example, Zhou and Yan[10] fitted the time series of cumulative probable cases to a simple logistic curve, estimating $R_0 = 2.1$ for Hong Kong. (However, because this approach integrates over both precontrol and postcontrol phases, their estimate is an average of a larger "true" R_0 before control measures and a smaller R_c after control measures.) By comparison, Riley et al[1] developed a detailed stochastic spatial simulation of SARS in Hong Kong and estimated $R_0 = 2.7$ during the precontrol phase (and in the absence of SSEs; including the SSEs would have increased this estimate, perhaps by 20%). Taking a very different approach by applying maximum-likelihood procedures to reconstruct the chain of transmission, Wallinga and Teunis[13] estimated $R = 3.6$ for Hong Kong.

In other cases, different types of models produced very different results. For example, Chowell et al[4] assumed a structured population consisting of 2 groups with different susceptibility, estimating $R_0 = 1.2$ for Hong Kong. Because these researchers derived a mathematical expression for R_0, the source of this discrepancy is immediately evident: if model parameters are chosen such that the population is uniformly susceptible, the result is $R_0 = 2.6$, roughly in line with other estimates. (We note in passing that this ease of comparison demonstrates 1 of the advantages of rigorous mathematical analysis over complex individual-based simulations that must be reimplemented by other investigators for results to be confirmed.)

Other discrepancies can likewise be explained. Wang and Ruan[9] used a simplified compartmental model to estimate that R_0 lies between 1.1 and 3.3 for Beijing. However, the lower bound is highly unrealistic because it was derived under the incorrect assumption that transmission ceased entirely after admission to the hospital.[35] Choi and Pak[7] fit cumulative case data in the early phases of the outbreaks to a geometric growth curve and estimated $R_0 = 1.5$ for Toronto and $R_0 = 2.0$ for Hong Kong. These estimates are also lower than they should be because the authors incorrectly used the incubation period (5 days) in their calculations instead of the generation time, or serial interval, of the disease. (See the definitions in Table 2.) Using the generation time instead (eg, 8.4 days in Singapore[2]) yields $R_0 = 2.5$ for Toronto and $R_0 = 3.4$ for Hong Kong.

The R_0 estimates of Gumel et al[14] are uniformly higher than other estimates (Fig. 1). The authors note that their possible overestimation of the efficacy of isolation would lead to overestimates of the transmission rate; this in turn will cause overestimation of R_0. The nosocomial transmission model of Hsieh et al[5] incorporates details of the hospital admission and reclassification procedures. By fitting the model to time series of cumulative cases and deaths, they estimated $R_0 = 4.2$ for Taiwan. However, they cautioned that this value may be an overestimate as a result of uncertainties regarding the temporal variation in infectiousness and viral load. They also noted that the value may reflect peculiarities of the Taiwan outbreak such as the relatively high rate (73%) of nosocomial transmission.

This critical comparison suggests that major discrepancies among published studies can largely be accounted for and that R_0 for SARS is approximately 3 in populations where an outbreak occurred. The fact that independent studies using different models can be reconciled is very encouraging. However, important questions remain, particularly with respect to heterogeneity. The dependence of R_0 on heterogeneity in host susceptibility is apparent in the results of Chowell et al,[4] as pointed out previously. Other heterogeneities such as spatial structure[1,36] and pathogen strain structure[6] can also alter predictions. Most of the published models did not incorporate such heterogeneities, although these are important for SARS. For example, there was considerable age-related variation in susceptibility and transmissibility,[35,37–40] an increased risk of infection for healthcare workers who inadvertently acted to bridge within-hospital and community-wide epidemics,[3,41] a clustering together of individuals with above-average susceptibility in healthcare settings (eg, the elderly),[35,42,43] and, possibly, several different modes of transmission.[6] Clusters of infections among particular groups occurred often, and models have begun to examine the importance of social networks in the spread of SARS.[11,12] Considerable spatial heterogeneity also existed for SARS on a global scale, with some countries experiencing relatively large outbreaks and others having little or no activity (although these differences are probably partly the result of chance).[44] Individual heterogeneity in infectiousness can be quantified from detailed contact tracing data, but the associated model-free estimation of R_0 (as the mean number of cases caused in the chain of transmission) may be biased low as a result of missing data.[34a]

The Generation Time (T)

To determine the speed with which an epidemic spreads, one must know both R_0 and the generation time (T) of the infection. This information is more important for epidemic EIDs than endemic diseases: when T is larger, more time is available to apply control measures, conduct gene sequencing, characterize pathogen epidemiology, educate the public, set up isolation facilities, and conduct contact tracing. Conversely, when T is smaller, even a disease with a modest R_0 can spread out of control as health authorities lose the "race to trace."[45] Generation time is implicitly incorporated into most epidemic models in the form of assumptions about incubation periods and infectious periods. It should be em-

phasized, however, that the simplest types of compartmental models (ordinary differential equations) misrepresent the population distributions of these periods, such that too many individuals progress too rapidly to symptoms and then recovery. More realistic (empiric) distributions of incubation and infectious periods may be preferable.[1,3] This point has been made before,[46,47] but the distinction may be particularly important for SARS because evidence shows that viral load (and perhaps infectiousness) increases for the first 10 to 15 days after symptoms appear.[48]

The Proportion of Transmission Occurring Before Symptom Onset (θ)

The parameter θ is the proportion of transmission that occurs before symptoms appear.[32] θ determines the potential efficacy of various control measures and suggests whether models should include nosocomial transmission. For diseases with large θ, transmission in the general population will be more widespread, and presymptomatic controls such as quarantine should be more effective. Conversely, for small θ, we may deduce that quarantine is effective only insofar as it reduces the time to isolation, and nosocomial transmission should be more important because individuals become infectious after exhibiting symptoms and being admitted to the hospital.

SARS had a relatively low R_0 (approximately 3), a small value of θ (<0.11),[32] and T of approximately 1 week. Smallpox is the most similar disease with respect to these 3 parameters, and it is interesting to note that smallpox and SARS are to date the only 2 infectious diseases that have been eradicated worldwide (although smallpox was endemic and different control strategies were applied). In the next section, we review control strategies for SARS and discuss how epidemic models were used to assess their efficacy.

ASSESSING CONTROL STRATEGIES

As is typical of many EIDs, there were no vaccines or drugs against SARS. Thus, the only ways of limiting transmission were the time-honored methods of quarantine, case isolation, travel restrictions, and barrier precautions (eg, masks, gloves, handwashing). Such control strategies have been applied for centuries and will doubtless be the first line of defense against future EIDs as well, yet they had received relatively limited attention from modelers before the emergence of SARS. The qualitative impact of control measures is known without the aid of models: they tend to decrease transmission. However, models offer a rigorous analytical framework that provides a nuanced, quantitative understanding of the impact of control measures. The efficacy of multiple control measures may depend on control and epidemiologic parameters in a complex, nonlinear way rather than in a simple additive fashion. For instance, there may exist a threshold in contact-tracing delays beyond which an epidemic "blows up." In such cases, models can characterize this nonlinear dependence and quantify such thresholds. Also, when control measures are excessively costly, inconvenient, or burdensome, rigorous modeling can help determine whether and when partial relaxation of certain control measures is desirable. In the following discussion of SARS control strategies, we emphasize conclusions that can be derived only by a modeling approach.

Expressed in its most basic form (and ignoring population heterogeneity and temporal development), R_0 is the product of the duration of infectiousness (D), the number of contacts of an infected individual with susceptible individuals per unit time (c), and the probability of transmission per contact between an infected and a susceptible individual (β). Hence, $R_0 = Dc\beta$. Reducing R_0 below 1 depends on efforts to reduce D (eg, through antimicrobial therapy or reducing time from infection to isolation), c (eg, through quarantine, case isolation, movement restrictions), and β (eg, through masks, gloves, handwashing). We use this convenient (if idealized) categorization in the following discussion of control measures.

Quarantine and Isolation

Quarantine, which reduces the contact rate (c) of possibly infected individuals before symptoms appear, controls an outbreak through 2 routes. First, the quarantine of incubating individuals prevents them from spreading the infection once they become infectious. Second, the monitoring of quarantined individuals for evidence of symptoms reduces the time from onset of symptoms to hospital admission. If θ is large, then quarantine acts through both routes described previously, whereas if θ is small, the first route is much less important. For the case of SARS, with a relatively small value of θ, quarantine should work mostly by reducing the onset-to-admission time; this intuition is borne out by modeling results.[3]

The few SARS models that studied quarantine have generated some surprising predictions. Lipsitch et al[1] developed a deterministic compartmental susceptible–exposed–infected–removed (SEIR) model for transmission in the general population with a parameter controlling the proportion of contacts of probable cases that are quarantined. They found that, beyond a certain threshold, quarantining a larger fraction of an infected person's contacts actually decreases the average number of days spent in quarantine per person, because a smaller fraction of the population must ultimately be quarantined if the epidemic is controlled in its early stages. Lloyd-Smith et al[3] developed a detailed mechanistic model that accounted for delays in contact tracing and imperfect quarantine. They concluded that the impact of quarantine actually increases as R_0 increases. In fact, for sufficiently low values of R_0, quarantine has little discernible effect. These models assume that the rate of quarantine of undiagnosed cases

grows linearly with the number of undiagnosed cases; in fact, the process is nonlinear, because the availability of contacts to trace depends on diagnosis of other cases. A more sophisticated mathematical analysis is required to take this nonlinearity into account.[32,49]

On a larger spatial scale, restricting movement between subpopulations is equivalent to reducing the contact rate (c). Modeling movement restriction requires a spatially structured model such as the metapopulation model based on Hong Kong's districts developed by Riley et al.[1] Their model predicted that imposing movement restrictions between the districts of Hong Kong would have reduced transmission substantially. Although their depiction of both natural movement patterns and movement controls was highly idealized, their conclusion raised the important point that countries lacking sufficient resources for less disruptive measures could effectively fight SARS through movement restrictions. The contact rate may also have dropped because of behavioral changes in the general public such as decreased attendance at social functions and decreased use of public transport. Riley et al[1] attributed part of the drop in transmission in Hong Kong to this effect.

Barrier Precautions

The probability of transmission per contact (β) can be reduced through handwashing and through barrier precautions such as masks and gloves.[3–5] In their model incorporating separate community and hospital-based populations, Lloyd-Smith et al[3] found that contact precautions applied to all individuals in the hospital are always more effective than precautions directed only at identified patients with SARS (arising from the possibility of infected but undiagnosed healthcare workers). During the SARS outbreaks, precautions were typically restricted to known SARS cases, and it is clear, with hindsight, that undiagnosed or misdiagnosed patients were a critically important source of SARS transmission worldwide.[35,50] Hence, the benefit of strict hospital-wide contact precautions is an important lesson to apply to future difficult-to-diagnose SARS-like illnesses.

Duration of Infectiousness

A third way to bring R_0 below 1 is by reducing D either by decreasing the natural duration of infectiousness in the host through antimicrobial therapies, or by decreasing the effective duration of infectiousness through faster hospitalization, classification, and isolation. For most novel pathogens (including SARS), only the latter approach is possible. For SARS, many models indicated that decreasing the time between symptom onset and case isolation has a disproportionate effect in controlling the outbreak.[1,2,4] This effect arises because little or no transmission occurs before symptoms are apparent, and indeed the predicted effect would have been greater still had models incorporated the suspected increase in SARS infectiousness in the first 10 days after infection.[48] Hsieh et al[5] modeled the Taiwan outbreak and found that most nosocomial transmission occurred in the time between admission of suspected SARS cases and reclassification as probable SARS or non-SARS cases. Hence, the authors recommended that efficient diagnosis should be a top priority in control of nosocomial outbreaks.

The Importance of Timing

A recurring theme in many SARS models is the importance of timely application of control measures.[2,3,8,9,13,14] These models show that quarantine and isolation have a disproportionate impact on epidemic control if applied early in the outbreak. Conversely, delays in imposing control can lead to large case burdens or even failure of potentially successful containment measures. The dependence of outbreak size on time of application is nonlinear: there are crucial periods early in the outbreak beyond which the effectiveness of control measures is severely degraded. This qualitative truth is familiar to epidemiologists; the role of models is to quantitatively identify critical periods for policy implementation and to characterize the nonlinear dependence of outbreak size on the efficacy and timeliness of control measures. A less scientific but no less significant role for models is to demonstrate the disastrous consequences of delays or the implications if certain pathogens were to be introduced into virgin populations, helping to overcome political inertia and spur decisive measures.

Effects of Combined Control Measures

Authors have been in general agreement that the application of multiple control measures was crucial in controlling the SARS outbreak.[1–3] However, according to the models used in these studies, some control measures were much more effective than others. For example, Lipsitch et al[2] predicted that reductions in time to isolation should have a disproportionate impact on transmission, whereas Gumel et al[14] found that isolation is much more effective than quarantine and suggested that more resources should be devoted to isolation programs than quarantine programs. Moreover, models can play an important role in highlighting nonintuitive interactions between simultaneous control measures. Lloyd-Smith et al[3] identified major tradeoffs between control measures, showing that it is fruitless to upgrade isolation practices if patients are not hospitalized quickly enough, or, conversely, that more rapid hospitalization does not help if transmission by identified patients is not sufficiently reduced. In general, although multiple control measures can lead to valuable synergy, blind improvement of any control measure may be a less wise investment than targeted upgrades of identified weaknesses.

A final contribution of modeling is in estimating the efficacy of control measures through interpretation of inci-

dence data, particularly by comparing fitted values of parameters before and after imposition of controls (Lloyd-Smith et al., unpublished data).[1,4,13] Using this method, Riley et al[1] illustrated a decrease in both the transmission rate and onset-to-admission time as the Hong Kong SARS epidemic was brought under control. They attributed the decline in transmission to changes in the public contact pattern, although it has been noted elsewhere that this decline may also have been the result of spatial structure.[36] Chowell et al[4] studied the 2 successive Toronto outbreaks in this way, and demonstrated an increase in the effectiveness of isolation and a decrease in the time from infection to diagnosis during the second outbreak compared with the first outbreak. Finally, Lloyd-Smith et al (unpublished data) showed that SARS data from Beijing and Singapore are more consistent with control measures that entirely block transmission by some individuals than with measures that partially reduce transmission by all individuals.

DISCUSSION

This review suggests promising directions for future research on SARS and SARS-like EIDs. Most modeling studies assumed a relatively homogeneous population, although SARS exhibited considerable heterogeneity in space, transmissibility, and susceptibility. Such heterogeneities can significantly alter parameter estimates and model predictions.[1,4,51] Heterogeneous mixing patterns inherent in human societies tend to impede transmission and facilitate contact tracing and quarantine efforts relative to random mixing assumed in simpler models. Heterogeneity is expected to affect transmission more for SARS (which is primarily spread through droplets and close contact) than many other diseases such as measles (which is easily spread through an airborne route). Despite reports of age-related variation in infectiousness,[35] no age-structured SARS models have been published to date, and only a few models have incorporated limited treatments of spatial structure,[1] social structure,[3,11,12] or variation in susceptibility.[4] Furthermore, despite the paramount role played by air travel in spreading EIDs, modelers have only recently begun developing frameworks for assessing the efficacy of air travel restrictions (Earn et al., unpublished data). Advances in geographic information systems offer additional possibilities for analyzing geographic differences in R_0, hospitalization trends, death rates, and contact patterns.[52–56]

Few models have addressed quarantine realistically by taking factors such as quarantine failure into account.[3,57] Unlike many measures that cause negligible inconvenience to the general population, quarantine can impose a large burden: in Taiwan, 130,000 individuals were quarantined because of SARS.[5] Therefore, "fine-tuning" of quarantine policy as guided by epidemic models might be appropriate, particularly if the burden on the population can be reduced without much increase in cases. Models can help determine what classes of contacts deserve priority, how resources should be allocated to the various stages of the contact tracing process, and when quarantine (or isolation) can be ended. Another way to make models more useful to policymakers is by integrating them into cost-effectiveness analyses.[58] Such analysis can provide guidance on choice of control strategies in situations in which resources are limited, as is inevitably the case.

There is also a continuing role for models in understanding the basic epidemiology of SARS and other emerging infectious diseases. Superspreading events have a powerful impact on invasion dynamics for SARS and many other EIDs (Lloyd-Smith et al., unpublished data),[1,35] yet little is known about the causes of SSEs. Contributing factors may include misdiagnosis, age of patients, misguided medical procedures, heterogeneous contact rates or viral shedding, and coinfections with other respiratory pathogens (Lloyd-Smith et al., unpublished data).[35] Models could aid in integrating these complex factors, establishing their relative importance, and exploring how best to control SSEs. Another critical direction for future modeling work includes the implications for control of time-varying SARS viral load.[48]

SARS models have in many ways only "scratched the surface." New modeling studies devoted to addressing these issues can also take advantage of recent advances in the understanding of SARS epidemiology. Perhaps the most important benefit of further SARS modeling efforts is not preparation for the possible reemergence of SARS, but rather preparation for future respiratory EIDs such as pandemic influenza.

This review has emphasized the effect of modeling on the study of SARS. However, SARS has also affected modeling approaches. Discussions of control have traditionally focused on reducing R_0. This is suitable for the control and elimination of endemic diseases in which prevalence is substantial and widespread. However, epidemic EIDs are crucially affected by chance events in the early phases of their emergence, and control efforts are shaped by reliance on quarantine and isolation. In the wake of SARS, other parameters have been introduced or reemphasized as being complementary to R_0 for assessing the efficacy of control measures. These parameters included the proportion of transmission occurring before onset of symptoms,[32] variance in the number of secondary cases produced per infected individual (Lloyd-Smith et al., unpublished data),[2] and heterogeneity in number of contacts.[12]

SARS had pandemic potential, but a prompt and effective public health response curtailed this threat. The success of future efforts to curtail emerging infectious diseases will depend on worldwide collaboration among public health officials, medical doctors, epidemiologists, and epidemic modelers. Some of the benefits of this integration of expertise have been illustrated by this review. A critical comparison of

independent studies, running the gamut of possible models, can help clarify aspects of SARS epidemiology, inform control policy, and solidify confidence in the parameter estimates and prescriptions that emerge from modeling. However, for this synergy to be fully exploited, there must be more widespread and timely access to data by independent research groups. Data collected with public funds should be made widely accessible if public benefit is to be maximized.

ACKNOWLEDGMENTS

CTB was supported by an NSERC Discovery Grant. JOL-S was supported by NSF-NIH Ecology of Infectious Disease Grant DEB-0090323 and by NIH-NIDA Grant R01-DA10135.

REFERENCES

1. Riley S, Fraser C, Donnelly CA, et al. Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science*. 2003;300:1961–1966.
2. Lipsitch M, Cohen T, Cooper B, et al. Transmission dynamics and control of severe acute respiratory syndrome. *Science*. 2003;300:1966–1970.
3. Lloyd-Smith JO, Galvani AP, Getz WM. Curtailing transmission of severe acute respiratory syndrome within a community and its hospital. *Proc R Soc Lond B*. 2003;270:1979–1989.
4. Chowell G, Fenimore PW, Castillo-Garsow MA, et al. SARS outbreaks in Ontario, Hong Kong and Singapore: the role of diagnosis and isolation as a control mechanism. *J Theor Biol*. 2003;224:1–8.
5. Hsieh YH, Chen CW, Hsu SB. SARS outbreak, Taiwan, 2003. *Emerg Infect Dis*. 2003;10:201–206.
6. Ng TW, Turinici G, Danchin A. A double epidemic model for the SARS propagation. *BMC Infect Dis*. 2003;3:19–35.
7. Choi BC, Pak AW. A simple approximate mathematical model to predict the number of severe acute respiratory syndrome cases and deaths. *J Epidemiol Community Health*. 2003;57:831–835.
8. Nishiura H, Patanarapelert K, Sriprom M, et al. Modelling potential response to severe acute respiratory syndrome in Japan: the role of initial attack size, precaution, and quarantine. *J Epidemiol Community Health*. 2004;58:186–191.
9. Wang W, Ruan S. Simulating the SARS outbreak in Beijing with limited data. *J Theor Biol*. 2004;227:369–379.
10. Zhou G, Yan G. Severe acute respiratory syndrome epidemic in Asia. *Emerg Infect Dis*. 2003;9:1608–1610.
11. Masuda N, Konno N, Aihara K. Transmission of severe acute respiratory syndrome in dynamical small-world networks. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2004;69:031917.
12. Meyers LA, Pourbohloul B, Newman MEJ, et al. Network theory and SARS: predicting outbreak diversity. *J Theor Biol*. 2005;232:71–81.
13. Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am J Epidemiol*. 2004;160:509–516.
14. Gumel AB, Ruan S, Troy D, et al. Modeling strategies for controlling SARS outbreaks. *Proc R Soc Lond B*. 2004;271:2223–2232.
15. Anderson RM, Fraser C, Ghani AC, et al. Epidemiology, transmission dynamics and control of SARS: the 2002–2003 epidemic. *Phil Trans R Soc Lond B*. 2004;359:1091–1105.
16. Galvani AP. Emerging infections: what have we learned from SARS? *Emerg Infect Dis*. 2004;10:1351–1352.
17. Donnelly CA, Fisher MC, Fraser C, et al. Epidemiological and genetic analysis of severe acute respiratory syndrome. *Lancet Infect Dis*. 2004; 4:672–683.
18. Chowell G, Castillo-Chavez C, Fenimore PW, et al. Model parameters and outbreak control for SARS. *Emerg Infect Dis*. 2004;10:1258–1263.
19. Ferguson NM, Donnelly CA, Anderson RM. Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. *Nature*. 2001;413:542–548.
20. Ferguson NM, Donnelly CS, Anderson RM. The foot-and-mouth epidemic in Great Britain: pattern of spread and impact of interventions. *Science*. 2001;292:1155–1160.
21. Keeling MJ, et al. Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science*. 2001;294: 813–817.
22. Jezek Z, Grab B, Dixon H. Stochastic model for interhuman spread of monkeypox. *Am J Epidemiol*. 1987;126:1082–1092.
23. Fine PEM, Jezek Z, Grab B, Dixon H. The transmission potential of monkeypox in human populations. *Int J Epidemiol*. 1988;17:643–650.
24. May RM, Anderson RM. Transmission dynamics of HIV infection. *Nature*. 1987;326:137–142.
25. Cummings DA, Irizarry RA, Huang NE, et al. Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. *Nature*. 2004; 427:344–347.
26. Webb GF, Blaser MJ. Dynamics of bacterial phenotype selection in a colonized host. *Proc Natl Acad Sci U S A*. 2002;99:3135–3140.
27. Wein LM, Craft DL, Kaplan EH. Emergency response to an anthrax attack. *Proc Natl Acad Sci U S A*. 2003;100:4346–4351.
28. Bauch CT, Galvani AP, Earn DJD. Group interest versus self-interest in smallpox vaccination policy. *Proc Natl Acad Sci U S A*. 2003;100: 10564–10567.
29. May RM. Uses and abuses of mathematics in biology. *Science*. 2004; 303:790–793.
30. Blower SM, Dowlatabadi H. Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. *Int Stat Rev*. 1994;62:229–243.
31. Diekmann O, Heesterbeek JAP. *Mathematical Epidemiology of Infectious Diseases*. New York: John Wiley & Sons; 2000.
32. Fraser C, Riley S, Anderson RM, et al. *Proc Natl Acad Sci U S A*. 2004;101:6146–6151.
33. Anderson RM, May RM. *Infectious Diseases of Humans: Dynamics and Control*. Oxford: Oxford University Press; 1991.
34. Keeling MJ. The effects of local spatial structure on epidemiological invasions. *Proc Roy Soc Lond B*. 1999;266:859–869.
34a. Lloyd-Smith JO, Schreiber SJ, Kopp PE, et al. Superspreading and the impact of individual variation on disease emergence. *Nature*. In press.
35. Shen Z, Ning F, Zhou W, et al. Superspreading SARS events, Beijing. *Emerg Infect Dis*. 2004;10:256–260.
36. Dye C, Gay N. Modeling the SARS epidemic. *Science*. 2003;300:1884–1885.
37. Donnelly CA, Ghani AC, Leung GM, et al. Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. *Lancet*. 2003;361:1761–1766.
38. Pan X, Zhu Z, Xu F, et al. Evaluation of control measures implemented in the severe acute respiratory syndrome outbreak in Beijing, 2003. *JAMA*. 2003;290:3215–3221.
39. Li G, Zhao Z, Chen L, et al. Mild severe acute respiratory syndrome. *Emerg Infect Dis*. 2003;9:1182–1183.
40. Hon KL, Leung CW, Cheng A, et al. Clinical presentations and outcome of severe acute respiratory syndrome in children. *Lancet*. 2003;361: 1701–1703.
41. Marra MA, Jones SJ, Astell CR, et al. The genome sequence of the SARS-associated coronavirus. *Science*. 2003;300:1399–1404.
42. Wong PN, Mak SK, Lo KY, et al. Clinical presentation and outcome of severe acute respiratory syndrome in dialysis patients. *Am J Kidney Dis*. 2003;42:1075–1081.
43. Kwan BC, Leung CB, Szeto CC, Wang AY, Li PK. Severe acute respiratory syndrome in a hemodialysis patient. *Am J Kidney Dis*. 2003;42:1069–1074.
44. Galvani AP, Lei X, Jewell NP. Severe acute respiratory syndrome: temporal stability and geographic variation in death rates and doubling times. *Emerg Infect Dis*. 2003;9:991–994.
45. Kaplan EH, Craft DL, Wein LM. Emergency response to a smallpox attack: the case for mass vaccination. *Proc Natl Acad Sci U S A*. 2002;99:10935–10940.
46. Keeling MJ, Grenfell BT. Disease extinction and community size: modeling the persistence of measles. *Science*. 1997;275:65–67.
47. Lloyd AL. Realistic distributions of infectious periods in epidemic models: changing patterns of persistence and dynamics. *Theor Popul Biol*. 2001;60:5–71.

48. Peiris JS, Lai ST, Poon LL, et al. Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. *Lancet*. 2003;361:1319–1325.
49. Eames KTD, Keeling MJ. Contact tracing and disease control. *Proc R Soc Lond B*. 2003;270:2565–2571.
50. Leo YS, et al. Severe acute respiratory syndrome—Singapore, 2003. *MMWR Morb Mortal Wkly Rep*. 2003;52:405–411.
51. Durrett R, Levin S. The importance of being discrete (and spatial). *Theor Popul Biol*. 1994;46:363–394.
52. Rogers DJ, Randolph SE. Studying the global distribution of infectious diseases using GIS and RS. *Nat Rev Microbiol*. 2003;1:231–237.
53. Munch Z, Van Lill SW, Booysen CN, et al. Tuberculosis transmission patterns in a high-incidence area: a spatial analysis. *Int J Tuberc Lung Dis*. 2003;7:271–277.
54. Brooker S, Beasley M, Ndinaromtan M, et al. Use of remote sensing and a geographical information system in a national helminth control programme in Chad. *Bull World Health Organ*. 2002;80:783–789.
55. Tanser F, Lesueur D, Solarsh G, et al. The application of geographical information systems to important public health problems in Africa. *Trop Med Int Health*. 2000;5:40–46.
56. Zenilman JM, Glass G, Shields T, et al. Geographic epidemiology of gonorrhea and chlamydia on a large military installation-application of a GIS system. *Sex Transm Infect*. 2002;78:40–44.
57. Day T. Predicting quarantine failure rates. *Emerg Infect Dis*. 2004;10:487–488.
58. Edmunds WJ, Medley GF, Nokes DJ. Evaluating the cost effectiveness of vaccination programs: a dynamic perspective. *Stat Med*. 1999;18: 3263–3282.
59. Keeling MJ, Grenfell BT. Individual-based perspectives on R_0. *J Theor Biol*. 2000;203:51–61.
60. Gani R, Leach S. Transmission potential of smallpox in contemporary populations. *Nature*. 2002;414:748–51.
61. Chowell G, Hengartner NW, Castillo-Chavez C, et al. The basic reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda. *J Theor Biol*. 2004;229:119–126.