STATISTICS IN MEDICINE
Statist. Med. (in press)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/sim.2352

Modelling the transmission dynamics of acute haemorrhagic conjunctivitis: Application to the 2003 outbreak in Mexico

G. Chowell[1,2,*,†], E. Shim[3,‡], F. Brauer[4,§], P. Diaz-Dueñas[5], J. M. Hyman[1,¶] and C. Castillo-Chavez[3,‖]

[1]*Theoretical Division (MS B284), Los Alamos National Laboratory, Los Alamos, NM 87545, U.S.A.*
[2]*Department of Biological Statistics and Computational Biology, Cornell University, 432 Warren Hall, Ithaca, NY 14853, U.S.A.*
[3]*Department of Mathematics and Statistics, Arizona State University, P.O. Box 871804, Tempe, AZ 85287-1804, U.S.A.*
[4]*Department of Mathematics, The University of British Columbia, Vancouver, B.C., Canada V6T 1Z2*
[5]*Instituto Mexicano del Seguro Social (IMSS), Colima, Col. 28010, Mexico*

SUMMARY

We model an outbreak of acute haemorrhagic conjunctivitis (AHC) using a simple epidemic model that includes susceptible, infectious, reported, and recovered classes. The model's framework considers the impact of underreporting and behaviour changes on the transmission rate and is applied to a recent epidemic of AHC in Mexico, using a fit to the cumulative number of cases to estimate model parameters, which agree with those derived from clinical studies. The model predicts a 'mean time from symptomatic onset to diagnosis' of 1.43 days (95 per cent CI: 1–2.5) and that the final size of the Mexican epidemic was underreported by 39 per cent. We estimate that a primary infectious case generates approximately 3 secondary cases ($\bar{R}_0^* = 2.64$, SD 0.65). We explore the impact of interventions on the final epidemic size, and estimate a 36 per cent reduction in the transmission rate due to behaviour changes. The effectiveness of the behaviour changes in slowing the epidemic is evident at 21.90 (SD 0.19) days after the first reported case. Results therefore support current public health policy including expeditious announcement of the outbreak and public health information press releases that instruct individuals on avoiding contagion and encourage them to seek diagnosis in hospital clinics. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: epidemic model; parameter estimation; parameter identifiability; underreporting; acute haemorrhagic conjunctivitis; reproductive number

*Correspondence to: G. Chowell, Theoretical Division (MS B284), Los Alamos National Laboratory, Los Alamos, NM 87545, U.S.A.
†E-mail: gchowell@t7.lanl.gov
‡E-mail: alicia@mathpost.la.asu.edu
§E-mail: brauer@math.ubc.ca
¶E-mail: mac@t7.lanl.gov
‖E-mail: chavez@math.asu.edu

Received 16 November 2004
Accepted 11 July 2005

1. INTRODUCTION

Acute haemorrhagic conjunctivitis (AHC) is a communicable disease typical of tropical, coastal cities. Outbreaks usually last 1–2 months, and secondary attack rates are greater than 50 per cent within households [1]. Susceptible individuals exposed to the virus experience a short (1–2 days) incubation period followed by the onset of painful, swollen, red (inflamed) eyes. Lacrimation, foreign-body sensation and subconjunctival haemorrhaging are common [2]. Transmission occurs primarily via person-to-person contact or contact with contaminated objects (e.g. towels). Symptoms usually persist for 3–7 days with no long-term consequences. Because of this absence of sequelae, many cases are not properly reported to public health institutions [3], and this contributes to the disease's further spread. Conversely, properly reported individuals are guided in avoiding further disease transmission. Isolation is a key prophylactic strategy, but in developing countries like Mexico, official diagnosis is required to justify worker sick leave for home isolation—which accelerates recovery and reduces frequency and duration of infectious public contact. This sick leave requirement may make it less likely that infectious individuals will isolate from coworkers, thus potentially increasing the final epidemic size.

Multiple viruses have been identified as aetiological agents of AHC, including enterovirus 70, coxsakievirus A24 variant (CA24v) and adenovirus 11. These agents have been identified by centrifugation, enhanced culture, neutralization tests, and other methods [4]. Enterovirus 70 was the first identified agent in a 1969 outbreak in western Africa. Enterovirus 70 invaded the western hemisphere in 1981, causing several outbreaks of AHC in Central America [5], South America [6], and Florida [7]. The coxsakievirus A24 variant (CA24v) was first identified in Singapore in 1970 [8], and first linked to an AHC outbreak in the western hemisphere on the islands of Trinidad, Jamaica and St. Croix, U.S. in the fall of 1986 [9]. Only months later, an outbreak of CA24v-caused AHC was observed in the Yucatan Peninsula of Mexico, where the secondary attack rate in households was 37 per cent [10]. The most recent reported epidemics of AHC have occurred in Delhi, North India in 1996 (enterovirus 70) [4]. In Saint Croix (CA24v), 1051 cases were reported in 1998 [1], and more than one million people were infected with AHC (CA24v) in South Korea [11] in 2002.

The goal here is to explore the role of non-reported AHC cases and the impacts of documented intra-outbreak behavioural changes on the dynamics of AHC contagion. A model for the 2003 outbreak of AHC in the state of Colima, Mexico is introduced. Using the model, epidemiological and control parameters and the number of secondary cases generated by a primary case in a fully susceptible population are estimated.

After presenting the model, an expression for its basic reproductive number (R_0^*) is computed. The parameter estimation procedure (in the context of available data) is described and an expression for the variance of R_0^* is derived. Subsequent to performing an analysis on parameter identifiability, a model-free estimate of the diagnostic rate is presented. Finally, results are summarized and their implications discussed.

2. MODEL

2.1. Epidemic model

The model (Figure 1) for the transmission dynamics of AHC classifies individuals as susceptible (S), exposed (E), infectious (I), diagnosed/reported (J), and recovered but not reported

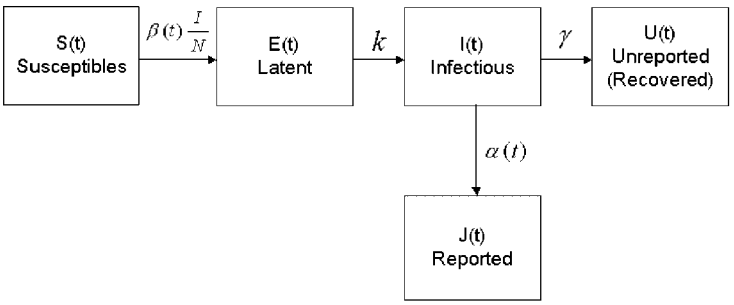

Figure 1. Schematic representation of the flow of individuals among the different classes. Susceptible individuals in contact with the virus enter the exposed class at the rate $\beta(t)I/N$ where $\beta(t)$ is the time-dependent transmission rate, I is the number of infectious individuals, and I/N is the probability that a contact is made with an infectious individual. The timescale of AHC outbreaks is typically much faster than those of demographic processes (births and deaths) (N is constant), and the final epidemic size was small compared to N. Hence, assuming a constant population size at risk is reasonable. Diagnosed/reported individuals are educated on how to avoid further contact with susceptible individuals. Hence, their contribution to further disease transmission is assumed to be negligible. Exposed individuals enter the infectious class at constant rate k (mean latent period is $1/k$). Infectious individuals are either diagnosed (reported) at the time-dependent rate, $\alpha(t)$, or recover at the constant rate, γ, without being diagnosed (underreported). Recovered individuals acquire immunity to the causative AHC virus strain for at least the duration of the outbreak, in agreement with the epidemiology of AHC [12].

(U). Susceptible individuals in contact with the virus enter the exposed class at the rate $\beta(t)I(t)/N$ where $\beta(t)$ is the time-dependent transmission rate, $I(t)$ is the number of infectious individuals at time t and $N(t)=S+E+I+J+U$ is the total population at time t. We assumed homogeneous mixing between individuals and, therefore, the fraction $I(t)/N$ is the probability that a random contact would be with an infectious individual. The timescale of AHC outbreaks is typically much faster than those of demographic processes (births and deaths), and the final epidemic size is small compared to N. Hence, we assumed that the size of the at-risk population is constant.

Since only diagnosed individuals in hospitals or clinics are granted sick days (Mexican policy) then high underreporting rates were observed during AHC outbreaks [13], limiting the impact of such a policy. These assumptions and practices add support to the homogeneous mixing modelling assumption (see, for example, References [14–18]). The incorporation of detailed population structure into the model (see, for example, References [19–21] and references therein) would not only increase the model complexity but also augment the number of parameters that must be estimated. The implementation of mathematical models can require highly detailed data that is often not available or difficult to get.

Once individuals are diagnosed/reported and, consequently, 'educated' on how to avoid further contact with susceptible individuals, we assumed that their contribution to further disease transmission is negligible. Exposed individuals enter the infectious class at constant rate k (mean latent period is $1/k$). Infectious individuals are either diagnosed (class J) at the time dependent rate $\alpha(t)$ or recover at the constant rate γ without being diagnosed (class U). Recovered individuals are assumed to acquire immunity to the causative AHC virus strain for at least the duration of the outbreak, an assumption that is in agreement with the epidemiology of AHC [12].

The transmission process (single outbreak) can be modelled using the system of non-linear differential equations:

$$\begin{aligned}
\dot{S}(t) &= -\beta(t)S(t)I(t)/N \\
\dot{E}(t) &= \beta(t)S(t)I(t)/N - kE(t) \\
\dot{I}(t) &= kE(t) - (\alpha(t)+\gamma)I(t) \\
\dot{J}(t) &= \alpha(t)I(t) \\
\dot{U}(t) &= \gamma I(t)
\end{aligned} \tag{1}$$

where the dot denotes the time derivatives. To account for behavioural changes in the population, the transmission rate $\beta(t)$ is modelled by the step function

$$\beta(t) = \begin{cases} \beta_0 & t < \tau \\ \beta_1 & t \geqslant \tau \end{cases}$$

where $\beta_1 < \beta_0$ and τ is the day when behavioural changes began to have a significant impact on transmission. The use of a time-dependent diagnostic rate $\alpha(t)$ is used to account for the low reporting rate at the beginning of the outbreak (before behavioural changes began to have an effect on the transmission dynamics of AHC) (Figure 2). For simplicity, the diagnostic

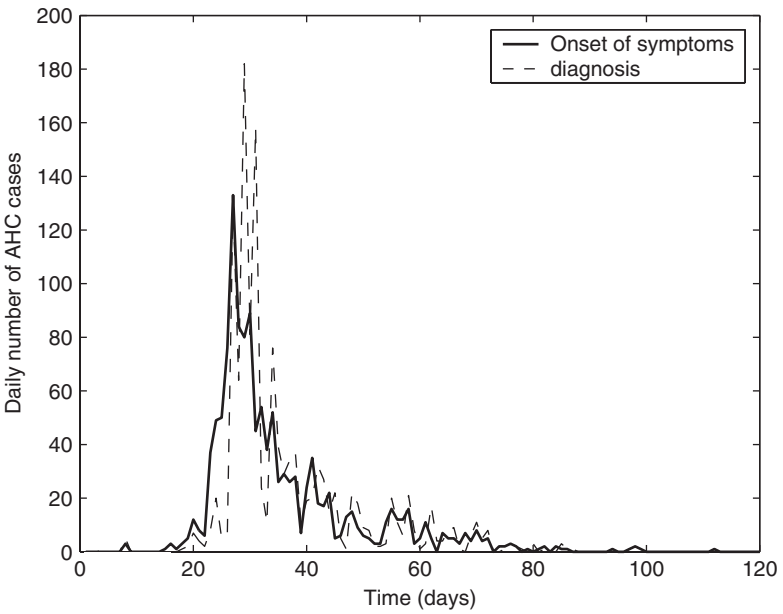

Figure 2. Daily number of AHC cases reported by date of symptom onset and date of case notification (from the 2003 outbreak in Colima, Mexico).

rate $\alpha(t)$ is also modelled by a step function:

$$\alpha(t) = \begin{cases} \alpha_0 & t < \tau \\ \alpha_1 & t \geqslant \tau \end{cases}$$

where $\alpha_0 < \alpha_1$. The model was unable to reproduce the beginning of the outbreak whenever $\alpha_0 = \alpha_1$.

2.2. *The reproductive number, R_0^**

The number of secondary cases generated by a primary infectious case or basic reproductive number (R_0) [22–24] is a measure of the power of an infectious disease to spread in a susceptible population at a demographic steady state. Once an epidemic is underway, the effective reproductive number ($R(t)$) decreases as the susceptible population is depleted, and as public health measures begin to take hold.

Public health measures may include contact tracing followed by quarantine, isolation, vaccination (if available), surveillance controls at ports of entry, and education of the population by fact-sheet dissemination. The number, type, and intensity of the control measures will depend on the disease in question and the availability of resources.

Since AHC disease symptoms vary among individuals, those with mild symptoms may never be properly diagnosed in hospital clinics, while mild cases may lead to higher incidence of infection at work or school leading to higher transmission rates in households. For benign diseases like AHC, public health recommendations include the following:

- the avoidance of direct or indirect (object sharing) contact with AHC cases;
- movement restrictions on infected individuals (sent home from school or work while symptomatic); and,
- increased hand washing [1].

Apropos the second recommendation, control of the AHC epidemic in Florida in 1981 was accelerated by closing affected schools. Likewise, during the 2002 AHC epidemic in South Korea [11], 1100 schools were closed.

The reproductive number (R_0^*) depends upon the activity of a typical infective in a large susceptible population at a demographic steady state. Typically, a minimal number of the initially infected people will be diagnosed (reported) at the rate α_0 (our estimate $\alpha_0 \approx 0.08\,\text{day}^{-1}$). This incidental diagnostic behaviour will affect the initial spread of the disease. Hence, the reproductive number $R_0^*(\alpha_0)$ is related to the basic reproductive number R_0 by the expression $R_0^*(0) = R_0$. $R_0^*(\alpha_0) = \beta_0/(\gamma + \alpha_0)$ (Appendix A), that is, the product of the initial transmission rate (β_0) and the average infectious period ($1/\gamma$) discounted by the initial diagnostic rate (α_0).

2.3. *Epidemic data*

The State of Colima is located on the Pacific coast, with a tropical climate, a mean annual temperature of 23–26°C, and an approximate population of 488 028 [25]. Data from the adenovirus-provoked AHC outbreak of September–November, 2003, were extracted from 1310 clinical records generated by the Mexican Institute of Public Health (IMSS) which provides service to 60 per cent (292 820) of the total Colima population. The relative frequency of

symptoms presented, the spatial and age-specific incidence distributions and the distribution of the time from the onset of symptoms to notification has been reported elsewhere [13]. Figure 2 shows the daily number of AHC cases by the time of onset of symptoms and by the time of notification.

3. PARAMETER ESTIMATION

The following disease-related parameters were estimated: $(\beta_0, \beta_1, k, \text{ and } \gamma)$, the time at which behavioural changes started to take place (τ), the mean time from symptom onset to diagnosis before and after interventions took effect $(1/\alpha_0$ and $1/\alpha_1)$, and the initial numbers of exposed and infectious individuals $(E(0)$ and $I(0))$ by least-squares fit of $J(t,\Theta)$ in model (1) (Θ is the vector of fitting parameters) to the cumulative number of AHC cases by date of case notification (Figure 2) under appropriate initial conditions. The best-fitting parameters were obtained by a 10-fold repeat of our fitting procedure (with different initial conditions) with parameters randomly drawn from appropriate parameter ranges $(0<\beta_0<10,\ 0<\beta_1<10,\ 0<\tau<100,\ 0<k<2,\ 0<\gamma<1,\ 0<\alpha_0<1,\ 0<\alpha_1<1)$. The resulting parameter estimates are listed in Table I, and the best model fit to the data is shown in Figure 3.

The standard deviation of the parameters was estimated by computing the asymptotic variance–covariance $AV(\hat{\Theta})$ matrix of the least-squares estimate from the expression:

$$\mathbf{AV}(\hat{\Theta}) = \sigma^2 \mathbf{B}(\Theta_0)\nabla_\Theta \mathbf{J}(\Theta_0)^{\mathrm{T}} \mathbf{G} \nabla_\Theta \mathbf{J}(\Theta_0)\mathbf{B}(\Theta_0)$$

where $\mathbf{B}(\Theta_0) = [\nabla_\Theta \mathbf{J}(\Theta_0)^{\mathrm{T}} \nabla_\Theta \mathbf{J}(\Theta_0)]^{-1}$. An estimate of this expression is the following

$$\hat{\sigma}^2 \hat{\mathbf{B}}(\hat{\Theta})\nabla_\Theta \hat{\mathbf{J}}(\hat{\Theta})^{\mathrm{T}} \mathbf{G} \nabla_\Theta \hat{\mathbf{J}}(\hat{\Theta})\hat{\mathbf{B}}(\hat{\Theta})$$

where $\hat{\mathbf{B}}(\hat{\Theta}) = [\nabla_\Theta \hat{\mathbf{J}}(\hat{\Theta})^{\mathrm{T}} \nabla_\Theta \hat{\mathbf{J}}(\hat{\Theta})]^{-1}$, $\hat{\sigma}^2 = \sum (y_i - J(t_i, \hat{\Theta}))^2/(I_{1xn}\ \mathbf{G}\ I_{nx1})$ and $\nabla_\Theta \hat{\mathbf{J}}$ are numerical derivatives of $J(\hat{\Theta})$. In order to account for the stochastic temporal dependence of the cumulative number of cases, the error structure [27] is modelled using a Brownian bridge ($\mathbf{G}$). Here, $\mathbf{G}$ is an $n \times n$ matrix such that $G_{i,j} = (1/n)\min(i,j) - (ij)/n^2$ where n is the total number of observations. That is, $\mathbf{G}$ captures the higher variability in the cumulative number of cases observed on the middle course of the epidemic and the smaller variability observed at its beginning and end.

3.1. Estimation of the variance of R_0^*

An expression for the variance of the estimated reproductive number (R_0^*) was obtained. T_1, T_2 and T_3 were designated as random variables with means $\beta_0, \alpha_0, \gamma_0$ and defined, such that $\mathbf{T} = (T_1, T_2, T_3)$ and $\Phi = (\beta_0, \alpha_0, \gamma)$. Using the Taylor series approximation of our R_0^* expression about Φ (delta method [28]) we obtain the following:

$$R_0^*(\mathbf{t}) \approx g_i(\Phi) + \sum_{i=1}^{3} g'(\Phi)(t_i - \phi_i)$$

Table I. Definitions and estimates for parameters in model (1) obtained from the best least-squares fit to the cumulative number of reported AHC cases.

Parameter	Definition	Estimate	SD
β_0	Mean transmission rate prior to behavioural changes (day^{-1} infective^{-1})	0.99	0.29
β_1	Mean transmission rate after behavioural changes (day^{-1} infective^{-1})	0.63	0.11
τ	Approximate time at which behavioural changes take place (day)	21.90	0.19
k	Rate of progression from the exposed to the infectious state (day^{-1})	0.27	0.07
γ	Recovery rate (day^{-1})	0.30	0.09
α_0	Diagnostic rate prior to behavioural changes (day^{-1})	0.08	0.01
α_1	Diagnostic rate after behavioural changes (day^{-1})	0.70	0.15

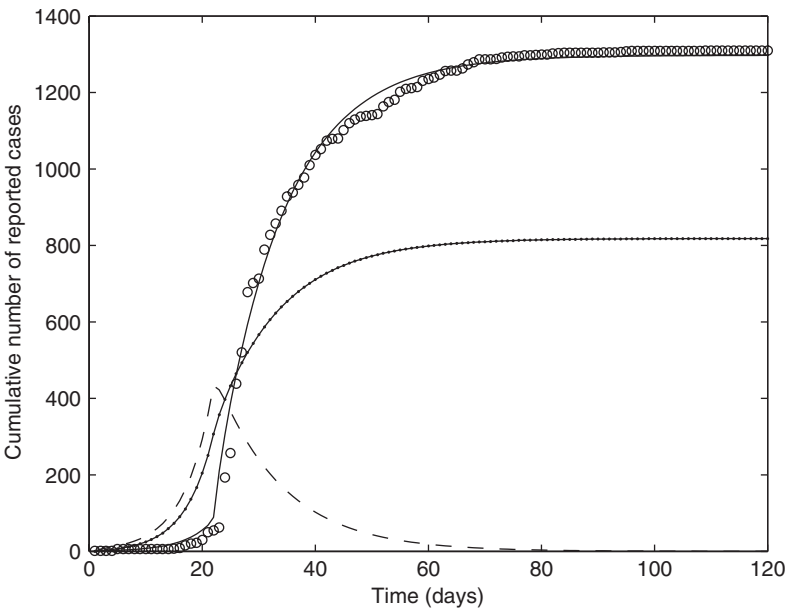

Figure 3. The best-fit solution obtained by fitting $J(t,\Theta)$ (solid line) in model (1) to the cumulative number of reported AHC cases (circles) as explained in the text (coefficient of determination is approximately 0.99 [26]). The dash–dot curve is the cumulative number of AHC unreported cases and the dash–dash curve is the number of cases incubating the disease. Model (1) predicts that behavioural changes began to have an effect on the transmission rate approximately 22 days after the first reported case.

The variance of $R_0^*(\mathbf{T})$ is approximated by

$$V(R_0^*(\mathbf{T})) \approx \sum_{i=1}^{3} [g_i'(\Phi)]^2 V(T_i) + 2\sum_{i>j} g_i'(\Phi) g_j'(\Phi) \mathrm{Cov}(T_i, T_j)$$

Hence, the variance of the estimated R_0^* is

$$V(\hat{R}_0^*) \approx \hat{R}_0^{*^2} \left\{ \frac{V(\hat{\beta}_0)}{\hat{\beta}_0^2} + \frac{V(\hat{\gamma})}{(\hat{\gamma}+\hat{\alpha}_0)^2} + \frac{V(\hat{\alpha}_0)}{(\hat{\gamma}+\hat{\alpha}_0)^2} - \left(\frac{2}{\hat{\beta}_0(\hat{\gamma}+\hat{\alpha}_0)}\right)\left(\mathrm{Cov}(\hat{\gamma},\hat{\beta}_0) - \frac{\hat{\beta}_0 \mathrm{Cov}(\hat{\alpha}_0,\hat{\gamma})}{\hat{\gamma}+\hat{\alpha}_0} + \mathrm{Cov}(\hat{\alpha}_0,\hat{\beta}_0)\right)\right\} \quad (2)$$

The corresponding values for the variance and covariance terms are obtained directly from the estimated variance–covariance matrix $\mathbf{AV}(\hat{\Theta})$.

3.2. Parameter identifiability analysis

We quantified the uncertainty in parameter estimates by systematically exploring the identifiability of the model parameters. Non-identifiability problems arise when small variations in model output correspond to large variations in some model parameters [29]. The best fit of the cumulative number of reported cases $J(t,\Theta)$ to the data was perturbed by simulating alternate realizations. To the best-fit curve $J(t,\Theta)$ was added a simulated Brownian bridge error structure computed using the increment in the 'true' $J(t,\Theta)$ from day i to day $i+1$ as the Poisson mean for the number of new cases observed in the i to $i+1$ interval. The parameter estimation procedure (described above) was then applied for each of the 1000 simulated realizations. The histograms of the parameter estimates from these simulations are shown in Figure 4, their nominal confidence intervals in close agreement with those obtained from the asymptotic variance–covariance matrix $\mathbf{AV}(\Theta)$ (Table I).

To study the applicability of the epidemic model to other outbreaks of AHC, the identifiability of model parameters was explored by performing a large number of simulations with parameter values randomly drawn from plausible ranges. Results were obtained from simulations of 20 sets of parameter values obtained by dividing the plausible range of each parameter (β_0 : 0.5–5, β_1 : 0.3–5, τ : 15–30, k : 0.1–2, γ : 0.1–1, α_0 : 0.05–0.5, α_1 : 0.1–1) into 20 equal parts which are randomly permuted to generate 20 different sets of parameter values. For each set of parameter values, 100 alternate realizations of the cumulative number of reported cases $J(t,\Theta)$ were simulated by adding the simulated Brownian bridge error structure. The nominal 95 per cent confidence intervals for the estimated parameters for each of the simulated realizations using our parameter estimation procedure (Table II) were then generated. The implications of these results are presented in Section 4.

3.3. Model-free estimate of the diagnostic rate

The mean time from symptom onset to diagnosis was estimated from clinical record data via maximum likelihood methods [30] (that is, independent of our ordinary differential equation model (1)). The date of symptomatic onset was missing or not properly recorded in 22 clinical records. Therefore, only 1288 (out of 1310) clinical records were considered. We let $t_1,\ldots,t_n$ be the 'times from symptom onset to diagnosis' (days) obtained from clinical records of each of the diagnosed (reported) individuals during the epidemic. The distribution of the 'times from onset to diagnosis' is well approximated by an exponential distribution (Figure 5). Hence,

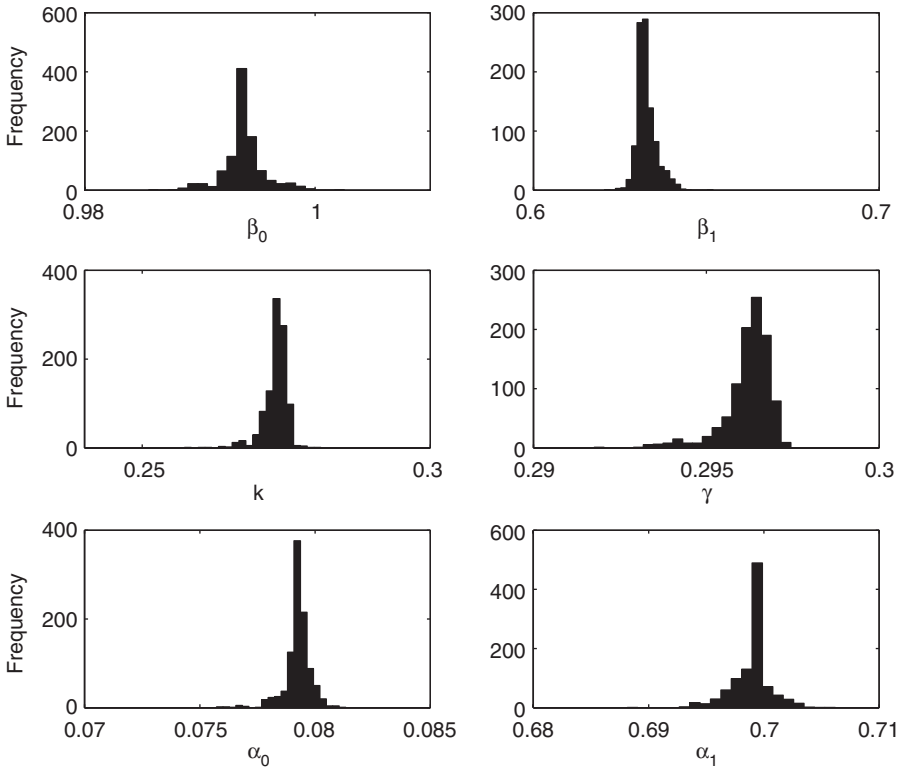

Figure 4. Distributions of the model parameter estimates obtained from simulation studies via the best fit of the cumulative number of reported cases $J(t,\Theta)$ to the data. Alternate realizations were simulated by adding a simulated Brownian bridge error structure, computed using the increment in the 'true' $J(t,\Theta)$ from day i to day $i+1$ as the Poisson mean for the number of new cases observed in the i to $i+1$ interval. One thousand realizations were simulated and our parameter estimation procedure applied to each of them. The nominal confidence intervals are in close agreement with those obtained from the asymptotic variance–covariance matrix $\mathbf{AV}(\Theta)$ (Table I).

the mean time to diagnosis (θ) is estimated by maximizing the log-likelihood equation

$$\ell(\theta|t_1,\ldots,t_n) = -\left[\sum_{i=1}^{n} t_i + n\ln(\theta)\right] \tag{3}$$

where the maximum likelihood estimator (MLE) of θ is the sample mean $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} t_i$ whose variance is $\mathrm{Var}(\hat{\theta}) = \hat{\theta}^2/n$.

4. RESULTS

Using a simple epidemiological model, the relevant parameters were estimated, agreeing well with actual data (see Figure 3), with a coefficient of determination of approximately 0.99 [26]. In turn, parameter estimates can be used to estimate the reproductive number and final

Table II. The nominal 95 per cent confidence intervals (in parenthesis) obtained from our simulation study for 20 different sets of parameter values as explained in the text.

Sim.	β_0	β_1	k	γ	α_0	α_1
1	0.5(0.495–0.506)	3.52(3.5–3.53)	0.763(0.757–0.769)	0.811(0.806–0.814)	0.0737(0.0704–0.079)	0.526(0.521–0.53)
2	0.737(0.73–0.743)	1.54(1.53–1.55)	0.432(0.405–0.451)	0.526(0.524–0.53)	0.5(0.496–0.505)	0.384(0.369–0.396)
3*	5(4.91–5.08)	3.27(0.00812–7.1)	0.811(0.797–0.827)	0.479(0.464–0.495)	0.405(0.394–0.418)	0.621(0.581–0.664)
4*	4.05(3.98–4.13)	2.28(0–6.21)	0.337(0.333–0.341)	0.147(0.144–0.151)	0.311(0.304–0.318)	0.716(0.688–0.75)
5	1.92(1.88–1.96)	3.76(3.7–3.83)	0.716(0.691–0.745)	0.668(0.66–0.678)	0.334(0.322–0.35)	1(0.977–1.03)
6	3.58(3.57–3.58)	3.02(3.01–3.03)	0.147(0.147–0.148)	0.574(0.572–0.576)	0.121(0.12–0.122)	0.763(0.758–0.767)
7*	4.29(4.2–4.39)	2.53(0.742–5.23)	0.195(0.192–0.197)	0.289(0.281–0.298)	0.0974(0.0936–0.101)	0.289(0.28–0.298)
8	2.16(2.15–2.17)	4.01(4–4.02)	0.1(0.099–0.101)	1(0.993–1.01)	0.358(0.35–0.367)	0.432(0.427–0.437)
9	0.974(0.968–0.98)	5(4.98–5.02)	1(0.996–1)	0.858(0.854–0.861)	0.453(0.451–0.455)	0.574(0.569–0.578)
10	2.39(2.36–2.42)	2.03(2.02–2.04)	0.384(0.373–0.396)	0.905(0.887–0.92)	0.476(0.466–0.489)	0.811(0.794–0.828)
11	1.45(1.44–1.45)	0.547(0.547–0.548)	0.289(0.286–0.293)	0.337(0.336–0.338)	0.192(0.191–0.194)	0.147(0.146–0.15)
12*	4.53(4.42–4.64)	4.75(2.83–7.01)	0.574(0.564–0.587)	0.763(0.742–0.793)	0.287(0.278–0.299)	0.953(0.837–1.14)
13	1.21(1.2–1.22)	4.51(4.42–4.59)	0.668(0.657–0.679)	0.195(0.193–0.197)	0.239(0.235–0.243)	0.195(0.193–0.197)
14*	3.11(3.02–3.19)	0.3(0.159–0.466)	0.242(0.236–0.249)	0.432(0.426–0.44)	0.216(0.206–0.226)	0.858(0.828–0.894)
15*	3.34(3.16–3.5)	0.795(0–6.19)	0.953(0.888–1.03)	0.384(0.376–0.392)	0.145(0.142–0.148)	0.905(0.19–1.57)
16	2.87(2.79–2.93)	2.77(2.68–2.86)	0.479(0.455–0.508)	0.953(0.946–0.958)	0.263(0.251–0.279)	0.1(0.0981–0.102)
17	2.63(2.61–2.65)	4.26(4.2–4.31)	0.621(0.611–0.632)	0.716(0.713–0.719)	0.429(0.424–0.435)	0.337(0.335–0.339)
18	1.68(1.68–1.69)	1.04(1.04–1.04)	0.526(0.519–0.532)	0.621(0.621–0.622)	0.168(0.168–0.169)	0.479(0.477–0.48)
19*	3.82(3.52–4.07)	1.29(0–9.48)	0.858(0.786–0.947)	0.242(0.231–0.252)	0.382(0.364–0.398)	0.242(−0.417–1.14)
20*	4.76(4.65–4.89)	1.78(0–10.2)	0.905(0.855–0.952)	0.1(0.0993–0.101)	0.05(0.0496–0.0503)	0.668(0.659–0.678)

For each set of parameter values, we simulated 100 alternate realizations of the cumulative number of reported cases $J(t, \Theta)$ by adding the Brownian bridge error structure and estimated parameters for each of the simulated realizations using our parameter estimation procedure. The parameter β_1 is not well identified in 8 (marked with *) out of the 20 sets of possible parameter values for other potential epidemics.

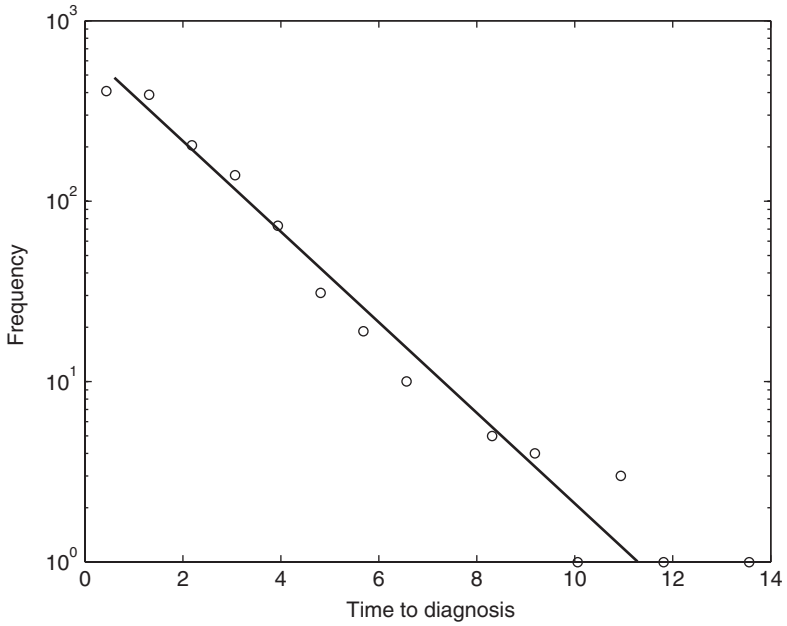

Figure 5. Semi-log plot of the distribution of the 'time from onset to diagnosis' obtained from the 1310 clinical records of AHC. Circles are the data and the solid straight line supports an exponential distribution. The maximum likelihood estimates of the mean and variance for the time of onset to diagnosis are 1.55 days (95 per cent CI: 1.46–1.63) and 2.39 day^2, respectively.

epidemic size, which was approximated as 2115 AHC cases, with an underreporting rate of approximately 38.7 per cent.

Behaviour changes began to have an impact on the transmission rate approximately 21.9 (SD 0.19) days after the first reported case. Our model predicted that the initial transmission rate $\beta_0 = 0.99$ (day^{-1} infective^{-1}) was reduced by 36.4 per cent to $\beta_1 = 0.63$ by behavioural changes and that the diagnostic rate increased from 0.08 to 0.70 (day^{-1}) (Table I). The estimates of latent period $(1/k)$ and the infectious period $(1/\gamma)$ were estimated to be approximately 3.65 and 3.37 days, respectively (see Table I). The estimated time from symptom onset to diagnosis was $1/\alpha_1 \approx 1.43$ days (95 per cent CI: 1–2.5), which compares favourably to the independent estimate of 1.55 days (95 per cent CI: 1.46–1.63) obtained using the dates of notification and dates of onset of symptoms of the 1310 individual clinical records via maximum likelihood methods (Figure 5). The two independent estimates are therefore consistent.

The simulation studies support our parameter estimates (Figure 4) for the 2003 AHC epidemic in Colima, Mexico. The parameter β_1 is not well identified in 8 of the 20 sets of possible parameter values for alternative (potential) epidemics. Clearly, β_1 should not be estimated (using this model) in non-comparable epidemic settings (Table II). Caution should be exercised when quantifying the role of interventions using the decay in the transmission rate from β_0 to β_1.

A sensitivity analysis was performed on the final epidemic size (Appendix B) to changes in the time (τ) at which interventions began to impact the transmission and diagnostic rates

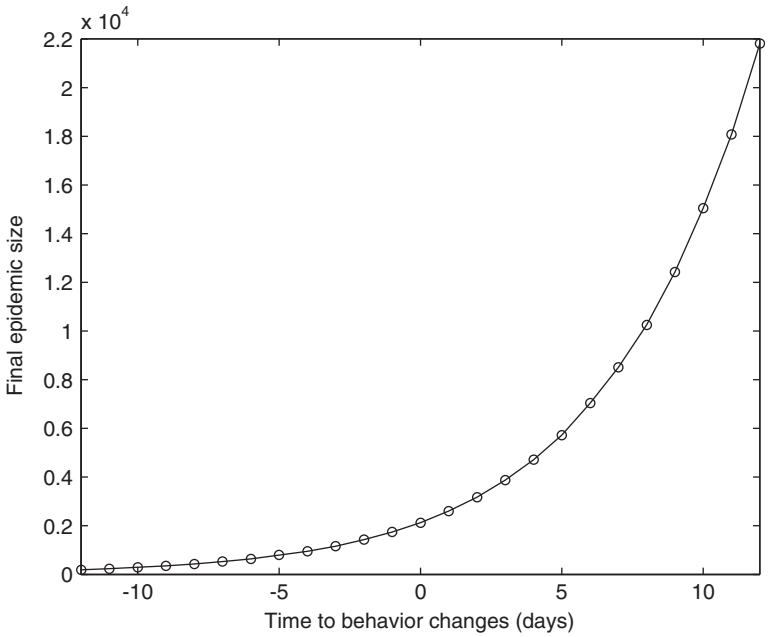

Figure 6. The sensitivity of the final epidemic size (reported and unreported cases) to the onset of behavioural changes evoked by public health measures. Negative numbers represent number of days before the actual estimated intervention time (Table I) and positive numbers represent a delay after the estimated intervention time. All other parameters have been fixed to their baseline values (Table I).

(Figure 6). Our model predicted an increase in the final epidemic size of 63 per cent (3609 cases) for a 5-day delay in the actual estimate of the effects of interventions, and a reduction of 62 per cent (1320 cases) had behavioural changes been in place 5 days sooner than the actual estimated time (Figure 6).

The expression for $R_0^* = \beta_0/(\gamma + \alpha_0)$ estimates the average number of secondary cases that a primary infectious AHC case generates in a fully susceptible population at a demographic steady state. In addition, an expression for the variance of R_0^* was derived (equation (2)). The reproductive number depends on the initial transmission rate (β_0), the infectious period ($1/\gamma$) and the initial mean time from onset of symptoms to diagnosis ($1/\alpha_0$). The reporting rate was very low at the beginning of this outbreak (Figure 2). Our estimate of the reproductive number (R_0^*) is 2.64 with standard deviation 0.65.

5. DISCUSSION

A relatively simple epidemic model (Figure 1) is used to assess underreporting, population behavioural changes, and the effects of basic public health interventions during a Mexican outbreak of acute haemorrhagic conjunctivitis (AHC). Our model was able to capture the time-course dynamics of the outbreak with sensible estimates of the relevant parameters. The force of infection (incidence) was modelled under the assumption of homogeneous mixing ($\beta SI/N$) [31] as previously demonstrated [14–18].

The introduction of models that incorporate population structure (e.g. age) is certainly possible [23], such model extensions capable of testing additional hypotheses. Unfortunately, such models often require a significant amount of data. The first mathematical model that incorporated household information was used to study an epidemic of bubonic plague [32] and age-structured models have been used to capture the seasonal dynamics of measles [22, 33]. Models considering household structure have also been used to study the distribution of the total number of secondary cases of infection in households invaded by *Variola minor* [34], influenza and the common cold [19, 35]. However, none of these models considered the role of interventions.

In this work, we estimated that 38.7 per cent of AHC cases were unreported. Undiagnosed individuals were more likely to transmit the disease (primarily due to lack of information on how to avoid further transmission). A reduction in the transmission rate as a consequence of disseminated public health information (fact sheets) was also considered. However, since there is a latency in such dissemination of public health information, the average elapsed time (τ) before interventions became effective in influencing the transmission and diagnostic rates was estimated. We estimated that the transmission rate was reduced by 36 per cent. The reporting increased due to behavioural changes in the population. Hence, mitigating the magnitude and impact of an outbreak is achievable by launching an awareness campaign as the outbreak starts; instructing individuals (via press releases) about how to avoid contagion (including indirect contact with contaminated objects such as utensils, glasses, towels, or laundry); stressing the importance of staying home from work or school until symptoms disappear; and decreasing the fraction of underreported individuals.

Parameter estimates from models should ideally be corroborated using empirical data. Unfortunately, either empirical data are not available or available empirical data can only partially validate model parameter estimates. Our estimates of the infectious period and latent periods agree with the epidemiology of AHC [36, 37]. However, the expected epidemiological variability might be due to differences between aetiological agents of AHC (i.e. Enterovirus 70, coxsakievirus A24 variant (CA24v), adenovirus 11). Our estimate from model (1) of the 'mean time to diagnosis' ($1/\alpha_1$) turned out to be in agreement with an independent estimate (obtained using maximum likelihood methods on the individual times to diagnosis of the clinical records of AHC cases generated during the epidemic).

Estimates of the final epidemic size were obtained using our analytical expression (equation 10) (Appendix B) which relied on estimates of the number of susceptible individuals around the time when the interventions were implemented (S_τ) or the outbreak concluded (S_∞). We also studied the sensitivity of model parameters to changes in the effective (assumed constant) population size (N), an uncertain quantity at the moment of modelling real disease outbreaks. Our parameter estimates are not sensitive to changes in the effective population size N, an observation that we have made elsewhere [18].

We found that a typical infectious case generates an average of approximately three secondary cases during its infectious period in a population of susceptible individuals at a demographic equilibrium. Our model predicts that behavioural changes in the population, together with a strategy of contact tracing followed by isolation of infectious cases provide an effective means of disease control. During the AHC outbreak reported here, the Mexican Institute of Public Health granted three sick days (isolation period at home) to each officially infected worker [13]. We estimated approximately 11 person-years of missed work from the 1310 reported cases (out of our estimated total of 2115 cases including unreported cases).

Additional data would be needed to assess productivity losses due to the lack of child care, the impact of unreported cases on productivity losses, medical costs, and costs incurred by public health departments, to name but a few. Even though AHC is regarded as a self-limiting disease, its control should be taken seriously, given that its economic impact can be significant.

APPENDIX A: THE REPRODUCTIVE NUMBER, R_0^*

Using the approach of van den Driessche and Watmough [24], we obtained an expression for the reproductive number R_0^* for our system with the transmission and the diagnostic rates fixed to their initial values β_0 and α_0, respectively. First, one can consider the disease transmission model consisting of initial conditions and the following system of equations:

$$\dot{x}_i = f_i(x) = \mathscr{F}_i(x) - \mathscr{V}_i(x), \quad i = 1, \ldots, 5$$

where $\dot{x}_i = f_i(x)$ represents our system (1) and

$$\mathscr{F} = \begin{pmatrix} \beta_0 SI/N \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad \mathscr{V} = \begin{pmatrix} kE \\ -kE + (\alpha_0 + \gamma)I \\ \beta_0 SI \\ -\alpha_0 I \\ -\gamma I \end{pmatrix}$$

Let x_0 denote the disease-free equilibrium of (1) and define $DF(x_0)$ and $DV(x_0)$ as follows:

$$D\mathscr{F}(x_0) = \begin{pmatrix} 0 & \beta_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$D\mathscr{V}(x_0) = \begin{pmatrix} k & 0 & 0 & 0 & 0 \\ -k & \gamma + \alpha_0 & 0 & 0 & 0 \\ 0 & \beta_0 & 0 & 0 & 0 \\ 0 & -\alpha_0 & 0 & 0 & 0 \\ 0 & -\gamma & 0 & 0 & 0 \end{pmatrix}$$

where F and V are the 2×2 matrices consisting of the first two rows and columns of $D\mathscr{F}(x_0)$ and $D\mathscr{V}(x_0)$, respectively. The basic reproductive number is given by the largest eigenvalue of FV^{-1}:

$$FV^{-1} = \begin{pmatrix} 0 & \beta_0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} k & 0 \\ -k & \gamma+\alpha_0 \end{pmatrix}^{-1}$$

$$= \frac{1}{k(\gamma+\alpha_0)} \begin{pmatrix} 0 & \beta_0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \gamma+\alpha_0 & 0 \\ k & k \end{pmatrix}$$

$$= \frac{1}{k(\gamma+\alpha_0)} \begin{pmatrix} \beta_0 k & \beta_0 k \\ 0 & 0 \end{pmatrix}$$

$$\rho(FV^{-1}) = \frac{\beta_0}{\gamma+\alpha_0} \tag{A1}$$

APPENDIX B: THE FINAL EPIDEMIC SIZE

The final size of an epidemic is tightly linked to its economic impact. Hence, it is of importance to estimate the final epidemic size. For the case of benign diseases, the number of reported cases highly underestimates the final epidemic size. Expressions for the final epidemic size have been derived for the standard SIR models and its extensions [38].

Here, an expression for the final epidemic size for our model with time-dependent transmission and diagnostic rates ($\beta(t)$ and $\alpha(t)$) is derived. From model (1), one can divide $\dot{U}(t)$ by $\dot{S}(t)$ to get

$$\frac{\mathrm{d}U(t)}{\mathrm{d}S(t)} = \frac{\gamma N I(t)}{-\beta(t)S(t)I(t)} = -\frac{\gamma}{\beta(t)}\frac{N}{S(t)} \tag{B1}$$

Using the fact that $\beta(t)$ is defined as a step function

$$U(t) - U(0) = -\gamma N\left[\int_0^\tau \frac{1}{\beta_0 S(t)}\,\mathrm{d}S + \int_\tau^t \frac{1}{\beta_1 S(t)}\,\mathrm{d}S\right]$$

$$= -\gamma N\left[\left(\frac{1}{\beta_0}-\frac{1}{\beta_1}\right)\ln(S(\tau)) - \frac{1}{\beta_0}\ln(S(0)) + \frac{1}{\beta_1}\ln(S(t))\right]$$

Since $U(0)=0$, it follows that

$$U(t) = -\gamma N\left[\left(\frac{1}{\beta_0}-\frac{1}{\beta_1}\right)\ln(S(\tau)) - \frac{1}{\beta_0}\ln(S(0)) + \frac{1}{\beta_1}\ln(S(t))\right] \tag{B2}$$

If we use the fact that $\lim_{t\to\infty} I(t)=0$ and $\lim_{t\to\infty} E(t)=0$ and let $U_\infty = \lim_{t\to\infty} U(t)$ and $S_\infty = \lim_{t\to\infty} S(t)$, then we obtain

$$U_\infty = -\gamma N\left[\left(\frac{1}{\beta_0}-\frac{1}{\beta_1}\right)\ln(S(\tau)) - \frac{1}{\beta_0}\ln(S(0)) + \frac{1}{\beta_1}\ln(S_\infty)\right] \tag{B3}$$

which gives the final number of unreported cases. We obtain similarly an expression for the total number of reported cases.

By dividing $\dot{J}(t)$ by $\dot{U}(t)$, we obtain

$$\frac{\mathrm{d}J(t)}{\mathrm{d}U(t)} = \frac{\alpha(t)I(t)}{\gamma I(t)} = \frac{\alpha(t)}{\gamma}$$

Using the fact that $\alpha(t)$ is defined as a step function

$$\begin{aligned} J(t)-J(0) &= \frac{1}{\gamma}\left[\int_0^\tau \alpha_0\,\mathrm{d}U + \int_\tau^t \alpha_1\,\mathrm{d}U\right] \\ &= \frac{1}{\gamma}[(\alpha_0-\alpha_1)U(\tau) - \alpha_0 U(0) + \alpha_1 U(t)] \end{aligned}$$

Since $U(0)=0$ and $J(0)=0$, it follows that

$$J(t) = \frac{1}{\gamma}[(\alpha_0-\alpha_1)U(\tau) + \alpha_1 U(t)] \tag{B4}$$

If we let $J_\infty = \lim_{t\to\infty} J(t)$, then we obtain

$$J_\infty = \frac{1}{\gamma}[(\alpha_0-\alpha_1)U(\tau) + \alpha_1 U_\infty] \tag{B5}$$

which gives the final number of reported cases. With equations (B3) and (B5), one can obtain an expression for the final epidemic size:

$$U_\infty + J_\infty = \frac{1}{\gamma}(\alpha_0-\alpha_1)U(\tau) + (\alpha_1+\gamma)N\left[\left(-\frac{1}{\beta_0}+\frac{1}{\beta_1}\right)\ln(S(\tau)) + \frac{1}{\beta_0}\ln(S(0)) - \frac{1}{\beta_1}\ln(S_\infty)\right] \tag{B6}$$

ACKNOWLEDGEMENTS

We are thankful to Nick Hengartner and Tom Burr for helpful conversations and two anonymous referees for useful comments. We also thank Anuj Mubayi and Mahbubur Rahman for their technical support.

REFERENCES

1. CDC. Acute haemorrhagic conjunctivitis—St. Croix, U.S. Virgin Islands, September–October 1998. *Morbidity and Mortality Weekly Report* 1998; **47**:899–901.
2. Hierholzer JC, Hatch MH. Acute haemorrhagic conjunctivitis. In *Viral Diseases of the Eye*, Darrell RW (ed.). Lea & Febiger: Malvern, PA, 1985; 165–196.
3. Dominguez A, Coll JJ, Fuentes M, Salleras L. Timeliness of notification in infectious disease cases. *Public Health Reports* 1992; **107**(4):474–476.
4. Maitreyi RS, Dar L, Muthukumar A, Vajpayee M, Xess I, Vajpayee RB, Seth P, Broor S. Acute haemorrhagic conjunctivitis due to enterovirus 70 in India. *Emerging Infectious Diseases* 1997; **8**:276–279.
5. CDC. Acute haemorrhagic conjunctivitis—Panama and Belize, 1981. *Morbidity and Mortality Weekly Report* 1981; **30**:497–500.
6. CDC. Acute haemorrhagic conjunctivitis—Latin America. *Morbidity and Mortality Weekly Report* 1981; **30**:450–451.
7. CDC. Acute haemorrhagic conjunctivitis—Key West, Florida. *Morbidity and Mortality Weekly Report* 1981; **30**:463–464.
8. Mirkovic RR, Schmidt NJ, Yin-Murphy M, Melnick JL. Enterovirus etiology of the 1970 Singapore epidemic of acute conjunctivitis. *Intervirology* 1974; **4**:119–127.
9. CDC. Epidemiologic notes and reports acute haemorrhagic conjunctivitis caused by coxsakievirus A24—Caribbean. *Morbidity and Mortality Weekly Report* 1987; **36**:245–246, 251.
10. CDC. International notes on acute haemorrhagic conjunctivitis—Mexico. *Morbidity and Mortality Weekly Report* 1989; **38**:327–329.
11. Oh M-D, Park S, Choi Y, Kim H, Lee K, Park W, Yoo Y, Kim E-C, Choe K. Acute haemorrhagic fever caused by coxsakievirus A24 variant, South Korea, 2002. *Emerging Infectious Diseases* 2003; **9**:1010–1012.
12. Gale Encyclopedia of Medicine, The Gale Group, 2002, website: (www.healthatoz.com/healthatoz/Atoz/ency/adenovirus_infections.jsp)
13. Chowell G, Castillo-Chavez C, Diaz-Duenas P. Characterization of an outbreak of acute haemorrhagic conjunctivitis in Mexico, 2003. *Digital Journal of Ophthalmology* 2005; **11**(1). http://www.djo.harvard.edu/site.php?url=/physicians/oa/712
14. Gani R, Leach S. Transmission potential of smallpox in contemporary populations. *Nature* 2001; **414**:748–751.
15. Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, Gopalakrishna G, Chew SK, Tan CC, Samore MH, Fisman D, Murray M. Transmission dynamics and control of severe acute respiratory syndrome. *Science* 2003; **300**:1966–1970.
16. Riley S, Fraser C, Donnelly CA, Ghani AC, Abu-Raddad LJ, Hedley AJ, Leung GM, Ho LM, Lam TH, Thach TQ, Chau P, Chan KP, Lo SV, Leung PY, Tsang T, Ho W, Lee KH, Lau EM, Ferguson NM, Anderson RM. Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science* 2003; **300**:1961–1966.
17. Chowell G, Fenimore PW, Castillo-Garsow MA, Castillo-Chavez C. SARS Outbreaks in Ontario, Hong Kong and Singapore: the role of diagnosis and isolation as a control mechanism. *Journal of Theoretical Biology* 2003; **24**:1–8.
18. Chowell G, Hengartner NW, Castillo-Chavez C, Fenimore PW, Hyman JM. The reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda. *Journal of Theoretical Biology* 2004; **229**:119–126.
19. Ball F, Mollison D, Scalia-Tomba G. Epidemics with two levels of mixing. *Annals of Applied Probability* 1997; **7**:46–89.
20. Becker NG, Utev S. The effect of community structure on the immunity coverage required to prevent epidemics. *Mathematical Biosciences* 1998; **147**:23–39.
21. Song B, Castillo-Chavez C, Aparicio JP. Tuberculosis models with fast and slow dynamics: the role of close and casual contacts. *Mathematical Biosciences* 2002; **180**:187–205.
22. Anderson RM, May RM. *Infectious Diseases of Humans*. Oxford University Press: Oxford, 1991.
23. Brauer F, Castillo-Chavez C. *Mathematical Models in Population Biology and Epidemiology*. Springer: New York, 2000.
24. van den Driessche P, Watmough J. Reproductive numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences* 2002; **180**:29–48.
25. INEGI (Instituto Nacional de Estadistica, Geografia e Informatica). Conteo de poblacion y vivienda, 1995.
26. Neter J, Wasserman W. *Applied Linear Statistical Models*. Richard D. Irwin, Inc.: Homewood, IL, 1974.
27. Davidian M, Giltinan DM. Nonlinear models for repeated measurement data. *Monographs on Statistics and Applied Probability*, vol. 62. Chapman & Hall/CRC Press: New York/Boca Raton, FL, 1995.
28. Casella G, Berger RL. *Statistical Inference* (2nd edn). Duxbury Press: Belmont, CA, 2002.
29. Pillonetto G, Sparacino G, Cobelli C. Numerical non-identifiability regions of the minimal model of glucose kinetics: superiority of Bayesian estimation. *Mathematical Biosciences* 2003; **184**:53–67.

30. Bickel P, Doksum KA. *Mathematical Statistics*. Holden-Day: Oakland, CA, 1977.
31. Castillo-Chavez C, Velasco-Hernandez JX, Fridman S. Modeling contact structures in biology. In *Frontiers of Theoretical Biology*, Levin SA (ed.), Lecture Notes in Biomathematics, vol. 100. Springer: Berlin, Heidelberg, New York, 1994; 454–491.
32. McKendrick AG. Applications of mathematics to medical problems. *Proceedings of the Edinburgh Mathematical Society* 1926; **44**:98–130.
33. Schenzle D. An age-structure model of pre- and post-vaccination measles transmission. *IMA Journal of Mathematics Applied in Medicine and Biology* 1984; **1**:169–191.
34. Angulo JJ, Smith TL, Tsokos JO, Tsokos CP. Population-dynamics models and a sequential test in the analysis of the influence of household setting on the spread of Variola minor. *Journal of Theoretical Biology* 1980; **82**:91–103.
35. Longini IM, Koopman JS. Household and community transmission parameters from final distributions of infections in households. *Biometrics* 1982; **38**:115–126.
36. Buehler JW, Finton RJ, Goodman RA, Choi K, Hierholzer JC, Sikes RK, Elsea WR. Epidemic keratoconjunctivitis: report of an outbreak in an ophthalmology practice and recommendations for prevention. *Infection Control* 1984; **5**:390–394.
37. Harley D, Harrower B, Lyon M, Dick A. A primary school outbreak of pharyngoconjunctival fever caused by adenovirus type 3. *Communicable Diseases Intelligence* 2001; **25**:9–12.
38. Brauer F. Some simple epidemic models. *Mathematical Biosciences and Engineering*, in press.

