LETTERS

Inapparent infections and cholera dynamics

Aaron A. King[1,2], Edward L. Ionides[3], Mercedes Pascual[1,4] & Menno J. Bouma[5]

In many infectious diseases, an unknown fraction of infections produce symptoms mild enough to go unrecorded, a fact that can seriously compromise the interpretation of epidemiological records. This is true for cholera, a pandemic bacterial disease, where estimates of the ratio of asymptomatic to symptomatic infections have ranged from 3 to 100 (refs 1–5). In the absence of direct evidence, understanding of fundamental aspects of cholera transmission, immunology and control has been based on assumptions about this ratio and about the immunological consequences of inapparent infections. Here we show that a model incorporating high asymptomatic ratio and rapidly waning immunity, with infection both from human and environmental sources, explains 50 yr of mortality data from 26 districts of Bengal, the pathogen's endemic home. We find that the asymptomatic ratio in cholera is far higher than had been previously supposed and that the immunity derived from mild infections wanes much more rapidly than earlier analyses have indicated. We find, too, that the environmental reservoir[5,6] (free-living pathogen) is directly responsible for relatively few infections but that it may be critical to the disease's endemicity. Our results demonstrate that inapparent infections can hold the key to interpreting the patterns of disease outbreaks. New statistical methods[7], which allow rigorous maximum likelihood inference based on dynamical models incorporating multiple sources and outcomes of infection, seasonality, process noise, hidden variables and measurement error, make it possible to test more precise hypotheses and obtain unexpected results. Our experience suggests that the confrontation of time-series data with mechanistic models is likely to revise our understanding of the ecology of many infectious diseases.

Cholera is a diarrhoeal disease caused by enteric infection with the bacterium *Vibrio cholerae*. Six of the seven cholera pandemics that have swept the globe since 1817 originated in the low-lying, densely populated regions north of the Bay of Bengal, where the disease is endemic. Although much attention has been focused on cholera[1,8], unsolved puzzles remain about its mode of transmission and the role of host immunity in its dynamics. This is largely because, in regions where cholera is endemic, most cholera cases are mild or asymptomatic but the true extent of asymptomatic infection has been difficult to assess. Estimates of the ratio of asymptomatic to symptomatic cases vary greatly, and the importance of inapparent infections in the dynamics of cholera outbreaks is unknown. To determine what role is played by inapparent infections, we used an approach that allows indirect inference about unobserved variables.

A remarkably rich data set on the pattern of cholera epidemics exists in the form of mortality records kept by the sanitary commissioners of the former British East Indian province of Bengal[9]. The data consist of monthly cholera death counts in each of 26 districts over the period 1891–1940 (Supplementary Fig. 1). To analyse these data, we formulated a series of models incorporating known or hypothesized mechanisms of transmission and immunity. A

parsimonious model for cholera dynamics is of susceptible–infectious–recovered–susceptible (SIRS) form (Fig. 1a). A novel feature of this model is that it incorporates both transmission tied to human prevalence (using a traditional mass-action term) and transmission from an environmental reservoir (where the pathogen is commonly living in aquatic environments)[5,6,10–12]. This model is a
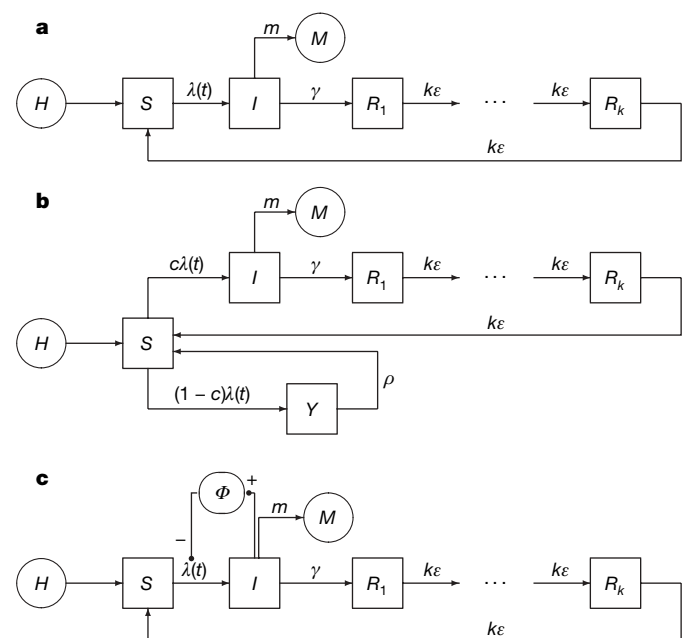

Figure 1 | The mechanistic models used. a, SIRS model; **b**, two-path model; **c**, environmental-phage model. Births, related to the total population size H, are assumed to feed the pool of susceptibles, S. Individuals are susceptible to infection when born. Exposure to the pathogen occurs at time-dependent rate $\lambda(t)$. c is the probability that an exposure leads to a contagious infection (class I). Note that when $c = 1$ and $\rho = \infty$, the two-path model (**b**) reduces to the SIRS model (**a**); when $c < 1$, some exposures result in short-term immunity (class Y). Infected individuals die at an excess rate m and recover at a rate γ; the time an individual spends within the I class is exponentially distributed. We assume that an individual remains immune to reinfection for a duration gamma-distributed with mean $1/\varepsilon$ and variance $1/k\varepsilon^2$. Once immunity has waned, an individual re-enters the susceptible pool (S). The measured variable is monthly deaths, M. The mean duration of short-term immunity is $1/\rho$. Individuals in each class are subject to constant background mortality at rate 0.02 yr^{-1}. The force of infection, $\lambda(t)$, includes terms for environmental and human sources of infection and is assumed to vary seasonally. Because the seasonality of cholera dynamics in Bengal is complex, we used a semi-mechanistic approach: transmission was modelled by a flexible periodic function of time. In the environmental-phage model (**c**), as infected hosts shed pathogen, phage Φ builds up in the environment and reduces transmissibility. The equations specifying these models are given in the Supplementary Equations.

---

[1]Department of Ecology and Evolutionary Biology, [2]Department of Mathematics, [3]Department of Statistics, University of Michigan, Ann Arbor, Michigan 48109, USA. [4]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA. [5]Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, University of London, London WC1E 7HT, UK.

partially observed, nonlinear, stochastic dynamical system in continuous time and as such is not amenable to analysis by standard statistical techniques. Inference for such systems has recently been facilitated by a new likelihood maximization procedure[7] which allows for measurement error, non-stationarity, irregular sampling intervals and the mechanistic inclusion of covariates (see Methods). We fitted the SIRS model to each district's data, assessed its explanatory power using log likelihood and the Akaike information criterion (AIC_c) and compared it with a semi-mechanistic time-series model for cholera[13] and two non-mechanistic time-series models.

The SIRS model explained the data dramatically better than the previous best fit[13] ($\Delta AIC_c > 58$; Supplementary Table 1). Parameter estimates reveal some surprises. A prediction of the model, robustly consistent across districts, is that immunity must wane on a timescale of weeks to months (Fig. 2 and Supplementary Table 3). This is in stark contrast to the widely held belief that infection-derived immunity to cholera wanes on a timescale of 3–10 yr (refs 13–17). The model also robustly predicts low case fatality (0.004 ± 0.002, mean ± s.d. across 26 districts). Because fatality among hospitalized cases was historically in excess of 50%[18], this implies a very large asymptomatic ratio. The implication is that most exposures result not in severe cholera, but in mild or asymptomatic infection, the immunological memory of which is short-lived. This prediction of a high asymptomatic ratio is consistent with the only intensive field studies of inapparent infections of which we are aware[3,4]. However, because in this model all infected individuals are equally infectious, the prediction is that the vast majority of infectives are 'silent shedders'—infectious but without symptoms. The evidence for such a conclusion is mixed but weak at best[2]. Moreover, because in this model all infections are alike, it is unclear whether the improved dynamical description afforded by the SIRS model over the previous best fit is due to the higher proportion of asymptomatic infections or to the shorter duration of immunity.

To tease apart the immunological consequences of exposure to the pathogen from the degree of infectiousness induced, and to allow for heterogeneity in infectiousness, we formulated a second mechanistic model in which exposed individuals can follow an alternate pathway, deriving immunity to infection while shedding a negligible quantity of vibrio and suffering a negligible disease-induced mortality (Fig. 1b). In essence, this model allows for the possibility that exposure to the pathogen effectively vaccinates against reinfection. We leave it to the data to determine how frequently such natural vaccination occurs and how long the resulting immunity lasts.

Under the two-path model, the timescale of short-term immunity (9.9 ± 4.7 weeks, mean ± s.d. across 26 districts) agrees with that of the SIRS model. The estimated case fatality under this model is 0.34 ± 0.21—consistent with conditions during this period[1]—and survivors of severe infections receive longer-lasting immunity

(1.5 ± 0.7 yr), consistent with experimental evidence[14,15,19]. Although in some districts the evidence strongly favours the two-path model, the data taken as a whole are equivocal on the subject of which model is better (see Supplementary Table 2). The implication is that it is the short-term immunity and high asymptomatic ratio predicted by both models, and not a large fraction of silent shedders, that accounts for the dramatically improved description afforded by our models.

A second robust prediction of both models is that, in typical epidemics, relatively few cases are due to the environmental reservoir. More precisely, though our models cannot speak to the exact route of transmission (that is, food-borne, water-borne, fomites, etc.), our results clearly indicate the relative importance of the positive feedback associated with the human-associated source of infection over the human-independent environmental source of infection. Nevertheless, the seasonally averaged basic reproductive number, R_0, for human-associated transmission is estimated to be quite low (SIRS model, 1.5 ± 0.2; two-path model, 1.5 ± 0.2). This contrasts sharply with the much higher values of R_0 recently proposed[20] and suggests, paradoxically, that for *V. cholerae*, in this the region where cholera is most persistent and outbreaks most frequent, humans are potentially a marginal habitat: the persistence of the disease may be largely due to the environmental reservoir which provides a small extrinsic force of infection. The strength of this environmental force of infection varies geographically in a pattern (Supplementary Fig. 2) that mirrors the earlier suggestion that the coastal regions of Bangladesh are the native habitat of classical cholera[21]. Both the SIRS and two-path models make the simplifying assumption that the force of infection due to the environmental reservoir has negligible seasonal fluctuation. To relax this assumption[22], we fit a third model which extends the SIRS model but allows for seasonality in the environmental reservoir. This model explains the data significantly better in many districts, but our conclusions about low R_0, rapid loss of immunity and high prevalence of inapparent infection remain unchanged (see Supplementary Information).

A central challenge in cholera epidemiology is to explain how the outbreaks can be so explosive at the outset and yet self-limiting to the point that sizable epidemics can recur twice annually. Our explanation contrasts sharply with those proposed earlier[13,17]. Simulations of the various models (Fig. 3) make plain the contrasts between the explanations. In the previous view, infection-derived immunity tends to wane over the scale of years, the explosiveness of outbreaks is controlled by spatial effects, and it is the seasonal drop in transmissibility (presumably related to environmental drivers) that stops the epidemics. Because in the previous view the ratio of deaths to infections is comparatively modest, the susceptible population is depleted only slightly by each epidemic and is continually replenished by births and the waning of immunity. In the new view, epidemics are again triggered by rising seasonal transmissibility and the presence of a reservoir, but the high asymptomatic ratio means that many more individuals are exposed. The vast majority of these receive short-term protection from infection so that depletion of the susceptible pool brings the epidemic to a halt. As immunity wanes, on the timescale of weeks to months, the susceptible pool is replenished, setting the stage for the next outbreak. In the new view, severe cases play little role in shaping the dynamics.

It has been proposed that the seasonality of cholera in this region may be due to the interaction of free-living vibrio with lytic bacteriophage in surface waters[23–25]. According to this hypothesis, as an epidemic progresses, the density of phage in the environment increases and survivorship of vibrio consequently declines, which in turn leads to attenuation of transmissibility. To determine whether such an effect might explain the data, we formulated a fourth model, which extends the SIRS model by allowing for phage to build up in the environment (as infected individuals shed pathogen into surface waters) and ultimately interrupt transmission (Fig. 1c and Supplementary Equations). Using the same methods, we estimated
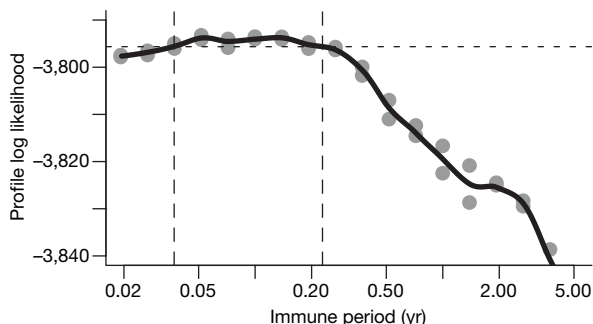

Figure 2 | Profile likelihood of the duration of immunity, $1/\varepsilon$, in the SIRS model for the data from Dacca district. Dacca (now spelled Dhaka) is the only district for which comparable analyses have been performed[13,27]. To compute the profile likelihood, ε is fixed and the likelihood is maximized over the remaining parameters. The vertical dashed lines embrace the approximate 95% confidence interval (2–12 weeks).

the strength and timescale of this effect; the results are presented in the Supplementary Information. In brief, the main conclusions described above are robust to the inclusion of the environmental phage effect: the estimated duration of immunity remains short (9.3 ± 8.3 weeks), fatality low (0.004 ± 0.002, compare the corresponding value for the SIRS model) and R_0 small (1.6 ± 0.3). Overall, the historical mortality data are not strong evidence in favour of the environmental-phage hypothesis. Furthermore, a critical implication of the environmental-phage hypothesis is that the phage cycles must lag behind cholera cycles, but not too far: the initial drop in cholera cases is predicted to be roughly coincident with high levels of ambient phage. In the one instance where concurrent data on cholera cases and ambient phage densities are available, however, the phage appear to lag approximately 180 degrees behind cholera cases[24]. Thus cholera cases begin to fall while phage concentrations remain small, and cholera cases begin to climb again while phage densities remain high. This is further, independent, evidence against the notion that phage–vibrio interaction in the environment is responsible for cholera seasonality.

The models presented here make several testable predictions. In particular, our results suggest that population surveys using sensitive techniques to detect the presence of *V. cholerae*, cholera-specific bacteriophage and/or coproantibodies in stools of healthy and mildly symptomatic individuals should reveal a much higher prevalence of inapparent infection than is currently understood to be the case. Such inapparent infections are predicted to be associated with reduced susceptibility to infection. Studies that centre on cholera patients, on the other hand, because they focus on the tail of the distribution of disease severity, can be expected to yield severely biased results when extrapolated to the population level. Further implications of our results for public health are discussed in the Supplementary Information.

In all models of epidemiological dynamics, critical quantities include the fraction of individuals susceptible to infection, immune from infection and asymptomatically infected. Yet it is almost never the case that available data include direct measurements of these quantities. Moreover, it is frequently true that many asymptomatic—or merely unrecorded—infections exist for each case that is recorded. Our results demonstrate that conclusions about the biological mechanisms underlying a pattern of disease outbreaks can depend sensitively on the prevalence of inapparent infections. New modelling approaches that allow indirect inference in the face of hidden variables and incomplete information are therefore likely to revise our understanding of the ecology of many infectious diseases.

METHODS SUMMARY

All models were formulated as stochastic differential equations, which were integrated using the Euler–Maruyama algorithm. Details of the equations and simulation methods are given in the Methods. The iterated filtering algorithm[7] was used to fit the models to the data. This algorithm is described in detail in the Supplementary Methods and is implemented as part of an open-source package, pomp, within the R statistical computing environment[26]. All codes will be made available by the authors on request.
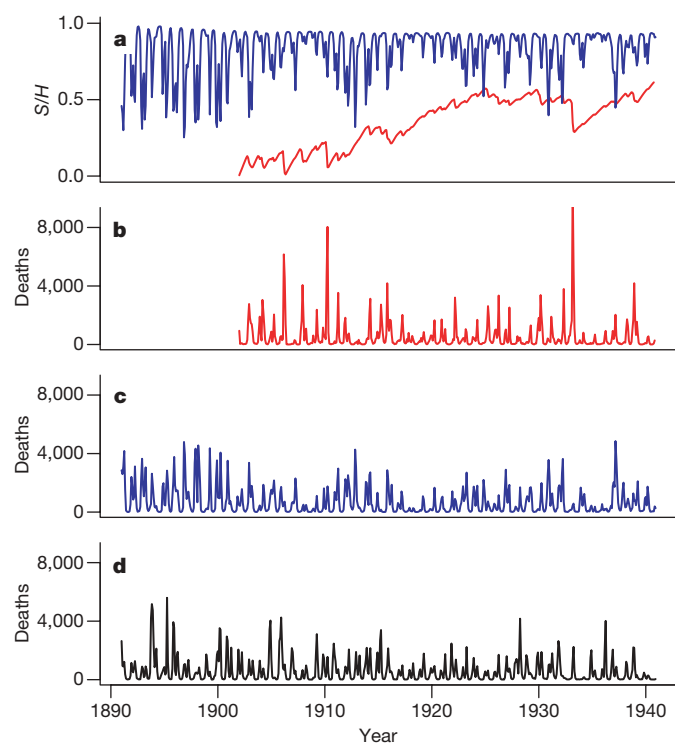
Figure 3 | Typical model simulations versus data for the district of Dacca. a, Simulated susceptible fraction (S/H) from our SIRS model with seasonal reservoir (blue) and from the time-series SIRS (TSIRS) model of Koelle *et al.*[13,17] (red). The estimated susceptible fraction of the population is much lower under the TSIRS model, and grows with time. By contrast, under the models introduced here, rapid waning of immunity allows the susceptible fraction to return seasonally to high levels. The predicted high prevalence of inapparent infection implies that the susceptible pool is rapidly depleted during an outbreak. **b**, Simulated monthly deaths from the TSIRS model. Despite the low susceptible fraction and that model's non-mass-action transmission assumption, the relatively low prevalence of inapparent infection leads to overly explosive outbreaks and overly deep crashes. **c**, Simulated monthly deaths from the SIRS model with seasonal reservoir. **d**, Cholera death data from Dacca district. Parameter estimates for the seasonal-reservoir SIRS model are reported in the Supplementary Information; parameters for the TSIRS model are from K. Koelle (personal communication). The semi-parametric regression used to fit the TSIRS model is initialized using the first 11 yr of data: no simulated data are therefore generated over that interval.

1.   Kaper, J. B., Morris, J. G. & Levine, M. M. Cholera. *Clin. Microbiol. Rev.* **8**, 48–86 (1995).
2.   Feachem, R. G. Environmental aspects of cholera epidemiology. III. Transmission and control. *Trop. Dis. Bull.* **79**, 1–47 (1982).
3.   Van De Linde, P. A. M. & Forbes, G. I. Observations on the spread of cholera in Hong Kong, 1961–63. *Bull. World Health Organ.* **32**, 515–530 (1965).
4.   McCormack, W. M., Islam, M. S., Fahimuddin, M. & Mosley, W. H. A community study of inapparent cholera infections. *Am. J. Epidemiol.* **89**, 658–664 (1969).
5.   Glass, R. I. & Black, R. E. in *Cholera* (eds Barua, D. & Greenough, W.B. III) 129–154 (Plenum Medical Book Co., New York, 1992).
6.   Islam, M. S., Drasar, B. S. & Sack, R. B. The aquatic environment as a reservoir of *Vibrio cholerae*: a review. *J. Diarrhoeal Dis. Res.* **11**, 197–206 (1993).
7.   Ionides, E. L., Bretó, C. & King, A. A. Inference for nonlinear dynamical systems. *Proc. Natl Acad. Sci. USA* **103**, 18438–18443 (2006).
8.   Sack, D. A., Sack, R. B., Nair, G. B. & Siddique, A. K. Cholera. *Lancet* **363**, 223–233 (2004).
9.   Sanitary Commissioner for Bengal Reports and Bengal Public Health Reports. (Bengal Secretariat Press, Calcutta and Bengal Government Press, Alipore, 1891–, 1942).
10. Glass, R. I., Lee, J. V., Huq, M. I., Hossain, K. M. & Khan, M. R. Phage types of *Vibrio cholerae* O1 biotype El Tor isolated from patients and family contacts in Bangladesh: epidemiologic implications. *J. Infect. Dis.* **148**, 998–1004 (1983).
11. Sack, R. B. *et al.* A 4-year study of the epidemiology of *Vibrio cholerae* in four rural areas of Bangladesh. *J. Infect. Dis.* **187**, 96–101 (2003).
12. Huq, A. *et al.* Critical factors influencing the occurrence of *Vibrio cholerae* in the environment of Bangladesh. *Appl. Environ. Microbiol.* **71**, 4645–4654 (2005).
13. Koelle, K. & Pascual, M. Disentangling extrinsic from intrinsic factors in disease dynamics: a nonlinear time series approach with an application to cholera. *Am. Nat.* **163**, 901–913 (2004).
14. Levine, M. M. *et al.* in *Acute Enteric Infections in Children: New Prospects for Treatment and Prevention* (eds Holme, T., Holmgren, J., Merson, M. H. & Mollby, R.) 443–459 (Elsevier/North-Holland Biomedical Press, Amsterdam, 1981).
15. Cash, R. A. *et al.* Response of man to infection with Vibrio cholerae. II. Protection from illness afforded by previous disease and vaccine. *J. Infect. Dis.* **130**, 325–333 (1974).
16. Glass, R. I. *et al.* Endemic cholera in rural Bangladesh, 1966–1980. *Am. J. Epidemiol.* **116**, 959–970 (1982).
17. Koelle, K., Rodo, X., Pascual, M., Yunus, M. & Mostafa, G. Refractory periods and climate forcing in cholera dynamics. *Nature* **436**, 696–700 (2005).

18. World Health Organization. *Wkly Epidemiol. Rec.* ⟨http://www.who.int/wer⟩ (1950–, 1965).
19. Cash, R. A. *et al.* Response of man to infection with *Vibrio cholerae*. I. Clinical, serologic, and bacteriologic responses to a known inoculum. *J. Infect. Dis.* **129**, 45–52 (1974).
20. Hartley, D. M., Morris, J. G. & Smith, D. L. Hyperinfectivity: a critical element in the ability of *V. cholerae* to cause epidemics? *PLoS Med.* **3**, 63–69 (2006).
21. Siddique, A. K. *et al.* Survival of classic cholera in Bangladesh. *Lancet* **337**, 1125–1127 (1991).
22. Bouma, M. J. & Pascual, M. Seasonal and interannual cycles of endemic cholera in Bengal 1891–1940 in relation to climate and geography. *Hydrobiologia* **460**, 147–156 (2001).
23. Faruque, S. M. *et al.* Self-limiting nature of seasonal cholera epidemics: Role of host-mediated amplification of phage. *Proc. Natl Acad. Sci. USA* **102**, 6119–6124 (2005).
24. Faruque, S. M. *et al.* Seasonal epidemics of cholera inversely correlate with the prevalence of environmental cholera phages. *Proc. Natl Acad. Sci. USA* **102**, 1702–1707 (2005).
25. Jensen, M. A., Faruque, S. M., Mekalanos, J. J. & Levin, B. R. Modeling the role of bacteriophage in the control of cholera outbreaks. *Proc. Natl Acad. Sci. USA* **103**, 4652–4657 (2006).
26. R Development Core Team. R: a language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria, 2007).
27. Rodo, X., Pascual, M., Fuchs, G. & Faruque, A. S. G. ENSO and cholera: a nonstationary link related to climate change? *Proc. Natl Acad. Sci. USA* **99**, 12901–12906 (2002).

METHODS

Inference for dynamical systems. A stochastic dynamical system is generated by a suitable function $f(x, s, t, \theta, W)$. Here, W is a stochastic quantity that is drawn independently each time the function is evaluated and θ is a vector of parameters. Thus, for a sequence of times $t_0 \leq t_1 \leq \cdots \leq t_N$, the recursion

$$X(t_n) = f(X(t_{n-1}), t_{n-1}, t_n, \theta, W),$$

together with the initial value $X(t_0)$, specifies a Markov process. The unspecified nature of $f(\cdot)$ reflects the idea that the evolution of the system could correspond to an essentially arbitrary stochastic simulation algorithm. In the context of this paper, $X(t)$ is a vector of counts in each disease class and $f(\cdot)$ corresponds to a (numerical) solution of any of the stochastic differential equation systems in the Supplementary Equations. The data consist of observations $y_1, \cdots, y_n$ at times t_1, $\cdots$, t_n. We suppose that, conditional on $\{X(t_k) = x_k\}_{k=1}^{n}$, y_n is drawn from a density $g(y_n | x_n, t_n, \theta)$.

In the Supplementary Methods, we present a pseudocode description of an iterated filtering methodology (maximization by iterated filtering, MIF) which has recently been shown to enable likelihood-based inference for this framework[7]. This procedure allows inference for rather general partially observed, stochastic, multivariate, continuous-time dynamical systems. The MIF algorithm also accommodates measurement error, non-stationarity, irregular observation times and the mechanistic inclusion of covariates. It has been implemented in the R statistical language[26] as part of the package pomp. Note that the algorithmic parameters play a role in the computational efficiency of the method but do not affect the scientific conclusions once likelihood maximization has been confirmed by diagnostic plots[7]. The package and all other codes will be made available by the first author on request.

Fitting the Bengal data. Our statistical inference technique is based on simulation of the stochastic dynamical system. To simulate the stochastic dynamical systems, we used the Euler–Maruyama numerical scheme[28] with a timestep of 0.05 months. We fit the parameters of each model to the data from each district using the algorithm given in the Supplementary Methods. We used $J = 10,000$ particles, a fixed lag $L = 60$ and $M = 80$ iterations with a cooling factor, $a = 0.95$. Using the same algorithm, we then constructed profile likelihoods against the parameters ε and c. This analysis revealed the existence of weak non-identifiabilities among the parameters $c, \gamma, m, \varepsilon$ and ρ. Despite this, certain combinations of parameters, such as $R_0 = c\langle\beta_{seas}(t)\rangle/(\gamma + \delta + m)$ and the case fatality $m/(\gamma + \delta + m)$, are well-identified. Moreover, although estimates of the durations of short-term immunity vary considerably among districts, it is the case that, for all districts, the duration of immunity is predicted to be far shorter than that derived from severe infections induced experimentally[13–17]. Construction of the profile likelihoods also facilitated the location of the global maximum likelihood estimates, which, after further refinement using the algorithm above with 30,000 particles, are reported in the Supplementary Tables.

Review of likelihood-based inference. We give a review of likelihood-based inference in the context of this paper. This review focuses on practical considerations; we refer the reader to Casella & Berger[29] or Rice[30] for a complete introduction, and to Barndorff-Nielsen & Cox[31] for a more advanced treatment. For a discussion of the role of likelihood, profile likelihood ratio tests applied to stochastic models for biological systems, we direct the reader to Hilborn & Mangel[32]. Reasons to use likelihood as a basis for inference include statistical efficiency (that is, making the strongest conclusions available from limited data), objectivity (that is, providing a single criterion applicable to a wide range of models and data) and the availability of useful techniques for comparing rival hypotheses.

For time-series data, written as $x_{1:T} = x_1, \cdots, x_T$, a stochastic model consists of a density function $f(x_{1:T}|\theta)$ depending on a vector θ of unknown parameters. The likelihood function arises from considering the density as a function of θ, treating the data as fixed. Thus the log likelihood is defined to be

$$\log \mathcal{L}(\theta) = \log f(x_{1:T}|\theta),$$

where log is the natural logarithm. Rival hypotheses, labelled H_0 and H_1, concerning the appropriate structure for a stochastic model can be formalized in terms of whether θ takes a value in Θ_0 or Θ_1 respectively, where Θ_0 and Θ_1 are subsets of the set Θ of all possible parameter vectors. We refer to H_0 and H_1 as hypotheses or models interchangeably.

Comparison of nested models. Suppose that H_0 and H_1 are nested, meaning that Θ_0 is a subset of Θ_1. In practical terms, this is taken to mean that Θ_0 is described by d_0 freely varying parameters, whereas Θ_1 is described by these d_0 parameters together with a number $d_1 - d_0$ of additional parameters. Thus Θ_0 can be written in terms of a vector of length d_0 and Θ_1 in terms of a vector of length d_1. The

maximum likelihood parameter estimate for model H_0 is denoted by $\hat{\theta}_0$, and corresponds to the parameter value giving rise to the largest value of $\log \mathcal{L}(\theta)$ among all values of θ in Θ_0. A corresponding definition applies to $\hat{\theta}_1$. Because Θ_1 includes Θ_0, it follows that $\log \mathcal{L}(\hat{\theta}_1) \geq \log \mathcal{L}(\hat{\theta}_0)$. Under standard regularity conditions, twice the improvement in the log likelihood that occurs when moving from the simpler model H_0 to the more complex model H_1 can be compared with a χ^2 random variable on $d_1 - d_0$ degrees of freedom. Thus, Supplementary Table 1 presents a P-value of the form

$$P = \text{Prob}\left[\chi^2_{d_1-d_0} > 2\left\{\log \mathcal{L}(\hat{\theta}_1) - \log \mathcal{L}(\hat{\theta}_0)\right\}\right]. \tag{1}$$

The P-value given by equation (1) is an approximation. More precise P-values can in principle be obtained by simulation, but this χ^2 approximation is a widely used technique that has been found to be reliable for many practical purposes[33]. It is both conceptually and operationally simpler to calculate initially the standard χ^2 approximation in equation (1); if this gives rise to an unequivocal answer, there is little reason to present a simulation study as additional evidence.

Modifications to equation (1) are required when the models under consideration are nested in a non-standard way[34]. This arises in the context of the analysis presented in Supplementary Table 1, owing to the non-standard nesting of the two-path model within the SIRS model. Specifically, the two-path model reduces to the SIRS model when $c = 1$ at which point ρ is unidentified. Anisimova *et al.*[35] investigated similar situations and demonstrated that the chi-squared approximation to the likelihood ratio test is conservative in such cases. In the caption of Supplementary Table 1, we report the P-value associated with such a test on three degrees of freedom: ρ, c and the initial condition Y_0.

Non-nested models may still be compared by their likelihood, for example by the use of Akaike's information criterion[36,37], but formal testing of hypotheses becomes more awkward. Thus, in Supplementary Table 1, we contrast the best-fitting SIRS model with the nested SIRS model constrained to match the duration of immunity found by Koelle & Pascual[13]. Another consideration motivating this comparison is that the nested SIRS model has a similar likelihood to the model of Koelle & Pascual[13], while having fewer parameters to estimate. Additionally, this approach avoids the issue that the comparison with previously published models might be unfair because the procedure used to fit the models in this paper may be more effective than that used by previous research.

Profile likelihood. The profile likelihood is a description of the weight of evidence about the value of a single component of the vector of unknown parameters. Setting $\theta = (\theta_1, \cdots, \theta_d)$, the log likelihood can be written as $\log \mathcal{L}(\theta_1, \ldots, \theta_d)$ and the profile log likelihood of θ_1 is

$$\log \mathcal{L}_{\text{profile}}(\theta_1) = \max_{\theta_2, \cdots, \theta_d} \log \mathcal{L}(\theta_1, \cdots, \theta_d).$$

From equation (1), a 95% confidence region for θ_1 is given by the range of values of θ_1 for which

$$2\left(\log \mathcal{L}(\hat{\theta}) - \log \mathcal{L}_{\text{profile}}(\theta_1)\right) < c,$$

where c is defined by $\text{Prob}\left[\chi^2_1 < c\right] = 0.95$. This is how the confidence interval in Fig. 2 was constructed.

28. Kloeden, P. E. & Platen, E. *Numerical Solution of Stochastic Differential Equations* 3rd edn (Springer, New York, 1999).
29. Casella, G. & Berger, R. L. *Statistical Inference* (Wadsworth, Pacific Grove, 1990).
30. Rice, J. A. *Mathematical Statistics and Data Analysis* (Wadsworth, Belmont, California, 1988).
31. Barndorff-Nielsen, O. E. & Cox, D. R. *Inference and Asymptotics* (Chapman and Hall, London, 1994).
32. Hilborn, R. & Mangel, M. *The Ecological Detective* (Princeton University Press, Princeton, New Jersey, 1997).
33. McCullagh, P. & Nelder, J. A. *Generalized Linear Models* 2nd edn (Chapman and Hall, London, 1989).
34. Self, S. G. & Liang, K.-Y. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J. Am. Stat. Assoc.* **82**, 605–610 (1987).
35. Anisimova, M., Bielawski, J. P. & Yang, Z. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. *Mol. Biol. Evol.* **18**, 1585–1592 (2001).
36. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* **19**, 716–723 (1974).
37. Burnham, K. P. & Anderson, D. R. *Model Selection and Inference: A Practical Information–Theoretic Approach* 2nd edn (Springer, New York, 2002).