Supporting Text

Model equations. The system of differential equations corresponding to the model is

$$\begin{aligned}
\dot{S} &= -\pi(\frac{V}{C(a)k+V})^a S - \delta S + \delta N \\
\dot{I}_- &= \pi(\frac{l}{l+P})(\frac{V}{C(a)k+V})^a S - (\mu_- + \delta) I_- \\
\dot{I}_+ &= \pi(\frac{P}{l+P})(\frac{V}{k+V})^a S - (\mu_+ + \delta) I_+ \\
\dot{R} &= \mu_- I_- + \mu_+ I_+ - \delta R \\
\dot{V} &= [m(1-\frac{V}{K_v}) - \gamma P]V + c(I_- + I_+) \\
\dot{P} &= (\beta\gamma V - \omega)P + \alpha c I_+,
\end{aligned}$$

where $N = S + I_- + I_+ + R$, $k = K/a$, $l = L/a$, $C(a) = 2^{\frac{1}{a}} - 1$, the dot represents the derivative with respect to time t in days, and parameters are as in Table 1. By making the transformation

$$\begin{aligned}
\{x, y, w, z\} &= \frac{c}{\pi k}\{S, I_-, I_+, R\} \\
u &= \frac{\omega}{{K_v}^2 \beta\gamma} V \\
v &= \frac{\gamma}{m} P \\
\tau &= \pi t
\end{aligned}$$

and writing

$$\begin{aligned}
f(u; k, a) &= (\frac{u}{C(a)k+u})^a \\
g(v; l) &= \frac{v}{l+v}
\end{aligned}$$

we obtain a nondimensionalized version of the original model equations:

$$\begin{aligned}
\dot{x} &= -f(u)x - \tilde{\delta}(x-n) \\
\dot{y} &= (1-g(v))f(u)x - (\tilde{\mu}_- + \tilde{\delta})y \\
\dot{w} &= g(v)f(u)x - (\tilde{\mu}_+ + \tilde{\delta})w \\
\dot{z} &= \tilde{\mu}_- y + \tilde{\mu}_+ w - \tilde{\delta} z \\
\dot{u} &= \tilde{m}(1-\phi u - v)u + \kappa(y+w) \\
\dot{v} &= \tilde{\omega}(\phi^2 u - 1)v + \xi w,
\end{aligned}$$

where $\{\tilde{\mu}_-, \tilde{\mu}_+, \tilde{m}, \tilde{\omega}, \tilde{\delta}\} = \pi^{-1}\{\mu_-, \mu_+, m, \omega, \delta\}$, $n = x+y+w+z$, and the dot now represents the derivative with respect to τ. The composite parameters are given by

$$\phi = K_v \frac{\beta\gamma}{\omega},\ \kappa = \frac{k\omega}{K_v^2\,\beta\gamma},\ \lambda = \frac{l\gamma}{m},\ \text{ and } \xi = \frac{\alpha k\gamma}{m}.$$

Note that the prevalence is given by

$$\frac{I_- + I_+}{N} = \frac{y+w}{n},$$

so that the dynamics of prevalence are completely specified by the nondimensional system. Because the shedding parameter c does not appear there, the dynamics are independent of c, as mentioned in the paper.

The R routines evaluate the nondimensionalized equations and provide scalings to recover the dimensional versions of the solutions.

Stability analysis of bacteria/phage dynamics. We will analyze the nondimensional equations and drop the tildes for convenience. Consider first the situation when $\delta = 0$. In this case, no new susceptibles are introduced. Then for all initial conditions, the infected classes $y, w \to 0$, as all

susceptible individuals are eventually infected and recover or die. Then the bacteria-phage equlibria of the entire system are identical to the equilibria of the following system:

$$\begin{aligned} \dot{u} &= m(1-\phi u - v)u \\ \dot{v} &= \omega(\phi^2 u - 1)v. \end{aligned}$$

Setting $\dot{u} = \dot{v} = 0$, we find the following possible equilibrium states:

$$(\hat{u}, \hat{v}) = \{(0,0), (\frac{1}{\phi}, 0), \text{ and } (\frac{1}{\phi^2}, 1 - \frac{1}{\phi})\}.$$

We infer stability by adding a small positive perturbation term (ϵ_u, ϵ_v) to each equilibrium in turn, substitute into the reduced equations, and consider the eigenvalues of the linearized system in (ϵ_u, ϵ_v). For the degenerate equilibrium $(0,0)$, the linearized system is

$$\begin{aligned} \dot{\epsilon_u} &= m\epsilon_u \\ \dot{\epsilon_v} &= -\omega\epsilon_v, \end{aligned}$$

so that the eigenvalues are simply m and $-\omega$, and stability ensues if both are negative, i.e, only if $m < 0$ and $\omega > 0$. Because we consider only $m > 0$ in this work, the degenerate equilibrium is always unstable.

For the resource-control equilibrium $(\frac{1}{\phi}, 0)$, first note that $\hat{V} = (K_v\phi)\hat{u} = K_v$, the carrying capacity, by the above variable transformation. The linearized system is

$$\begin{aligned} \dot{\epsilon_u} &= -\frac{m}{\phi}(\phi\epsilon_u + \epsilon_v) \\ \dot{\epsilon_v} &= \omega(\phi - 1)\epsilon_v \end{aligned}$$

and the eigenvalues are $-m$ and $\omega(\phi-1)$. Then, for $\omega > 0$, we have stability only if $m > 0$ and $\phi < 1$, and the nondegenerate phage-control equilibrium $(\frac{1}{\phi^2}, 1-\frac{1}{\phi})$ must be unstable. Stability of these two equilibria must switch when $\phi > 1$, so that this condition specifies phage control. Note that when $\phi > 1$, $\frac{1}{\phi^2} < \frac{1}{\phi}$. That is, equilibrium bacterial density is reduced below carrying capacity by phage predation.

For small $\delta > 0$, as in our numerical analysis, these conditions are good approximations for most of the levels of human bacterial and phage shedding we consider, and equilibria are shifted by small relative amounts. Under higher rates of infection of humans by phage, we find that in the resource control regime, $\hat{P} >> 0$ by virtue of a constant input of phage from infected humans (see paper).

Epidemic equilibrium and basic reproductive number. Consider the nondimensional system above, dropping the tildes, and setting $\mu_+ = \mu_- = \mu$. Setting the derivatives equal to zero, we can solve for the equilibrium values of the state variables, $\hat{x}, \hat{y}, \hat{w}, \hat{z}$, in terms of the equilibrium bacterial and phage densities $\hat{u}$ and $\hat{v}$:

$$\begin{aligned}
\hat{x} &= n(1+\frac{\hat{f}}{\delta})^{-1} \\
\hat{y} &= n\frac{1-\hat{g}}{\mu+\delta}(\frac{1}{\hat{f}}+\frac{1}{\delta})^{-1} \\
\hat{w} &= n\frac{\hat{g}}{\mu+\delta}(\frac{1}{\hat{f}}+\frac{1}{\delta})^{-1} \\
\hat{z} &= n\frac{\mu}{\delta}\frac{1}{\mu+\delta}(\frac{1}{\hat{f}}+\frac{1}{\delta})^{-1}
\end{aligned}$$

where $\hat{f} = f(\hat{u})$ and $\hat{g} = g(\hat{v})$. Setting $\dot{u}$ and $\dot{v}$ to zero and substituting the above expressions for $\hat{x}, \hat{y}, \hat{w}, \hat{z}$ gives exact equations for $\hat{u}$ and $\hat{v}$ that can be

solved numerically. These were used in establishing the initial conditions for the model outbreaks in the paper.

Using the observation (ref. 1, p. 17) that the basic reproductive number R_0 multiplied by the equilibrium fraction of susceptibles is equal to unity, we have

$$R_0 = \frac{n}{\hat{x}} = 1 + \frac{\hat{f}}{\delta}.$$

R_0 is plotted as a function of the bacterial median infectious dose k in Fig. 6.

Sensitivity analyses. Here we consider the effects of changes in bacterial growth rate and phage decay rate, over emipirical determined ranges, on the magnitude of simulated cholera epidemics.

Fig. 7 illustrates the effect of bacterial growth rate and the bacterial bloom size on the severity of an outbreak, in the absence of phage, in the case where phage cannot coexist with bacteria ($\phi = 2/3$). The severity is given in terms of the number of new disease cases exceeding the number expected under constant equilibrium cholera prevalence (the "excess cases"), over 1 year. Excess cases are calculated as follows: if the prevalence is p, the number of excess cases from the beginning of the epidemic to a time T equals $I(T) - I(0) + (\mu_- + \delta) \int_0^T (I_-(t) - pN)dt$, where N is the (constant) total number of individuals.

In Fig. 8 we consider the number of excess cases prevented by the introduction of exogenous phage into the environmental reservoir, over a range of initial phage densities. These results demonstrate that phage can affect epidemic course for the situation in which phage do not stably persist, but

that both bacterial growth and phage decay rates must be low to obtain a substantial effect.

Fig. 9 shows the effects of decreasing levels of phage instability, for the case in which bacteria and phage coexist ($\phi > 1$). Initial phage densities are given as fractions of the equilibrium phage density. Note that, in contrast to the resource-controlled case, outbreak severity is insensitive to changes in the bacterial growth rate over the experimental range.

Parameter estimates. Below we set out our rationale behind the choices of values for unknown parameters.

c, daily bacterial shedding per individual: In our model, the bacterial shedding parameter affects the dynamics only as a linear scaling of the epidemic compartments. Thus, when compartments are analyzed as proportions of the total population, c cancels out; it does not affect the endemic prevalence or the basic reproductive number of the epidemic, for example. We also found that, in the scenarios we consider, changes in initial human population numbers over several orders of magnitude do not strongly affect the dynamics of the outbreak. Thus we chose $c = 10$ for the purposes of explicit calculations, but the outbreak dynamics, and hence our conclusions, do not depend strongly on the value of this parameter. Because severely infected individuals excrete enormous numbers of vibrio [$5 \times 10^{11} - 2 \times 10^{12}$ cells per day (2)], a contribution of 10 cells per liter of reservoir is a modest estimate even for an extremely large body of water.

We did not explicitly model the growth of bacteria and phage within the individual, but supposed that each infected individual contributed a constant density of bacteria and/or phage to the reservoir each day over the course

of infection. The underlying assumption here is that bacterial growth within the infected person is very rapid relative to the time scales of the outbreak and bacterial growth in the reservoir. We believe this is justifiable, since the time between infection and onset of symptoms is short (3) relative to the length of outbreaks, and shedding and symptomatic infection are coincident. Numerical experiments incorporating a constant delay between infection and shedding of up to 10 d indicate that neither magnitudes nor times of epidemic and density peaks shift appreciably compared to the simpler model presented here (data not shown).

μ_+, daily recovery rate of phage-positive individuals: While the model was formulated to allow for differences in recovery rates between phage-negative and phage-positive infecteds, we set the recovery rates for both classes equal to 0.1. We justified this simplification primarily on the basis of phage therapy experiments (2) indicating that densities of phage much greater than arise through within-host replication alone are necessary to shorten the symptomatic period. Monsur (2), citing unpublished data, noted that the gut transit time of vibriophage is probably shorter than the eclipse phase (the time from phage adsorption to phage burst) in symptomatic individuals, so that multiple rounds of phage replication within the gut are unlikely.

α , phage/bacteria ratio in phage-positive individuals: Faruque *et al.* (5) observed between 10^2 and 10^8 virions/ml stool in their study; assuming 108 cells/ml in the stool of diarrheal patients (2), this gives between 10^{-6} and 1. We chose $\alpha = 1$ for our numerical experiments, to maximize the possible contribution of phage within the estimated range.

k, median infectious vibrio dose: Using human volunteers, Cash *et al.* (3)

estimated that this parameter lay between 10^6 and 10^8 cells. In that study, the lowest dose was effective only when buffered by sodium bicarbonate. We chose $k = 4 \times 10^7$ for our simulations. Figure S1 shows the sensitivity of the basic reproductive number R_0 to changes in k, given the other parameter values we chose. This gives a rough idea of how the magnitude of the model outbreaks would respond to different k.

a, threshold parameter: When $a = 1$, as in ref. 6, under the other parameter constraints, a low equilibrium prevalence of cholera (0.5 in 1,000) can be achieved only with a very low proportion of susceptible individuals, too low to give epidemics of the size observed in Dhaka. When $a > 1$, the rate of immigration can be adjusted to give larger numbers of susceptibles at equilibrium and consequently larger epidemic sizes. Although $a \sim 7$ is somewhat arbitrary, it gives epidemic sizes near that observed in Dhaka, without requiring unreasonable immigration rates. The effect of changes in a on the reproductive number R_0 can be seen in Fig. 6.

Data limitations and model "fitting". As comprehensive as they are, these data have features that make it difficult to estimate the model parameters using numerical procedures. Only relative bacterial densities were reported, while the model we consider was constructed in terms of absolute densities. The reason for this limitation is technical: viable pathogenic V. cholerae are difficult to recover from environmental sources by culture, requiring enrichment in selective medium (4, 7, 8). Comparison of fluorescent antibody staining and culture results suggest a detection limit of about 10^3 cells/ml (7), but even at higher densities, enriched cultures are frequently negative for cholera bacteria. Also, as we note in the Discussion, it is likely

that the environmental densities would be too low during most of the epidemic to account for cholera prevalence by consumption of environmental water alone. This leads us to suppose that local reservoirs exist that are actually responsible for most infections, but within which the microbial dynamics are unknown.

While the data of Faruque *et al.* (5) may not be amenable to numerical fitting, we can assume that the sampled densities of bacteria and phage are directly proportional to those in the reservoirs responsible for disease. The relative changes in microorganismal densities and cholera infections over time can then be studied with our mathematical reformulation of their conceptual model. To this end, we seek sets of parameters such that the behavior of the model reproduces the following qualitative and semi-quantitative features of the data: 1) a synchronous rise in environmental bacteria density and outbreak cases, 2) a delayed rise and fall of environmental phage density relative to the outbreak, 3) a reduction of bacterial and phage densities to undetectable levels (after putative dilution), and 4) a rise in phage-positive infecteds from zero to a high proportion of all infecteds. These dynamics should play out on a time scale comparable to that of the actual outbreak.

Access to model code. R code for numerical solution of the model equations and calculation of equilibria and excess cases can be obtained by contacting the authors.

1. Anderson, R. M. & May, R. M. (1991) *Infectious Diseases of Humans: Dynamics and Control* (Oxford Univ. Press, Oxford, U.K.).
2. Monsur, K. A., Rahman, M. A., Huq, F., Islam, M. N., Northrup, R. S. &

Hirschhorn, N. (1970) *Bull. World Health Org.* **42**, 723-732.

3. Cash, R. A., Music, S. I., Libonati, J. P., Snyder, M. J., Wenzel, R. P. & Hornick, R. B. (1974) *J. Infect. Dis.* **129**, 45-52.
4. Monsur, K. A. (1961) *Trans. R. Soc. Trop. Med. Hyg.* **55**, 440-442.
5. Faruque, S. M., Islam, M. J., Ahmad, Q. S., Faruque, A. S., Sack, D. A., Nair, G. B. & Mekalanos, J. J. (2005) *Proc. Natl. Acad. Sci. USA* 102, 6119-6124.
6. Codeco, C. T. (2001) *BMC Infect. Dis.* **1**, 1.
7. Huq, A., Colwell, R. R., Rahman, R., Ali, A., Chowdhury, M. A., Parveen, S., Sack, D. A. & Russek-Cohen, E. (1990) *Appl. Environ. Microbiol.* **56**, 2370-2373.
8. Brayton, P. R., Tamplin, M. L., Huq, A. & Colwell, R. R. (1987) *Appl. Environ. Microbiol.* **53**, 2862-2865.