Supplemental Methods for Model-2

Laura Matrajt, Jonathan D. Sugimoto, M. Elizabeth Halloran, and Ira M. Longini Jr.
Fred Hutchinson Cancer Research Center and University of Florida

1 Model Description

We constructed a deterministic meta-population model to simulate cholera transmission and vaccination in Haiti. Our model consists of a network of ten nodes representing the ten administrative departments of Haiti. For each department, susceptible individuals S_i get infected upon contact with an infected person (at a rate β) or with a water source (at a rate β_W, modulated by a sinusoidal function representing the dry and rainy season). Upon infection, individuals become exposed E_i (infected but not yet infectious) for $1/\gamma_E$ weeks. A proportion k of these individuals will develop symptoms and transitions to the infectious symptomatic (I_i) compartment. These individuals shed bacteria into the environment at a rate μ. The remaining $1-k$ infectious individuals are assumed to be asymtomatic (A_i), and are assumed to be significantly less infectious and shed significantly less cholera in the environment (represented by multipliers red_β and red_μ respectively). Infected individuals remain infectious for a period of $1/\gamma$ weeks after which they are assumed to be fully recovered and to have transient natural immunity to cholera infections. Natural immunity (both for symptomatic and asymptomatic infections) is assumed to last $1/\sigma$ weeks, after which recovered individuals become fully susceptible to infection again.

In addition, our model tracks the amount of bacteria present in a water reservoir W_i. Bacteria present in water decays at a rate δ. Finally, departments are connected at two levels: through a network of major highways, representing the flow of people in the country, and through a network of rivers, representing the flow of water (see figure 4). A full description of the human mobility network and water mobility network is given below.

Vaccination: We simulate vaccination with one or two doses of Oral Cholera Vaccine (OCV), by explicitly adding vaccination compartments (for one or two doses) to our model. We have the following nomenclature for the vaccination compartments: (V_{ij}, E_{ij}, I_{ij}, A_{ij}, R_{ij} and RA_{ij} for vaccinated susceptible, exposed, infectious symptomatic or asymptomatic and recovered symptomatic or asymptomatic. The index i ($i \in \{1,..10\}$) represents the department and $j = \{1,2\}$ represents vaccination with one or two doses respectively. For bookkeeping purposes, we further separated vaccinated individuals who are younger than five years old in different compartments, where we assumed that the proportion of children under five years old would stay constant over the whole simulation period, but we do not explicitly model aging. Hence, V_{ij_5}, E_{ij_5}, I_{ij_5}, A_{ij_5}, R_{ij_5} and RA_{ij_5} represent the analogous vaccination compartments for children under five years old. Vaccine protection is assumed to be "leaky" (*1*), so that vaccinated individuals see their force of infection reduced by θ_j ($j = 1,2$ represents vaccination with one or two doses respectively and θ_{j_5} represents vaccination with one or two doses for children under five years old). Vaccination was implemented in weekly pulses, and vaccinated individuals are assumed to have immediate protection with those vaccinated with a single dose being protected for one year, while those fully vaccinated being protected for five years (*2*). For example, to model a vaccination campaign with 80% coverage in all departments over a two-year period, we first computed the total number of doses to be given in a single week, and then modeled a vaccination campaign where each week that number of people would be vaccinated in a single department instantaneously. We vaccinated departments in order according to the 2018 reported cholera incidence, with those with the highest incidence being vaccinated first. Figure 1 provides the fraction of the total population vaccinated each week for each vaccination scenario (see the meta supplement for full descriptions of each vaccination scenario). We modeled a conservative scenario where only susceptible individuals will gain protection from the vaccine, so that recovered, infected or latent individuals do not benefit in the model from any vaccine protection. Vaccine efficacy was derived from (*2*). Because this is a deterministic model and vaccination was administered weekly, we could not model the vaccine efficacy in a dynamic way, instead, we utilized the median over five years (two doses) or over one year (one dose) of the log-linear weighted regression model presented in the main supplement. This resulted in $VE_2 = 51.9\%$ for two doses and $VE_1 = 42.9\%$ for a single dose. OCV was

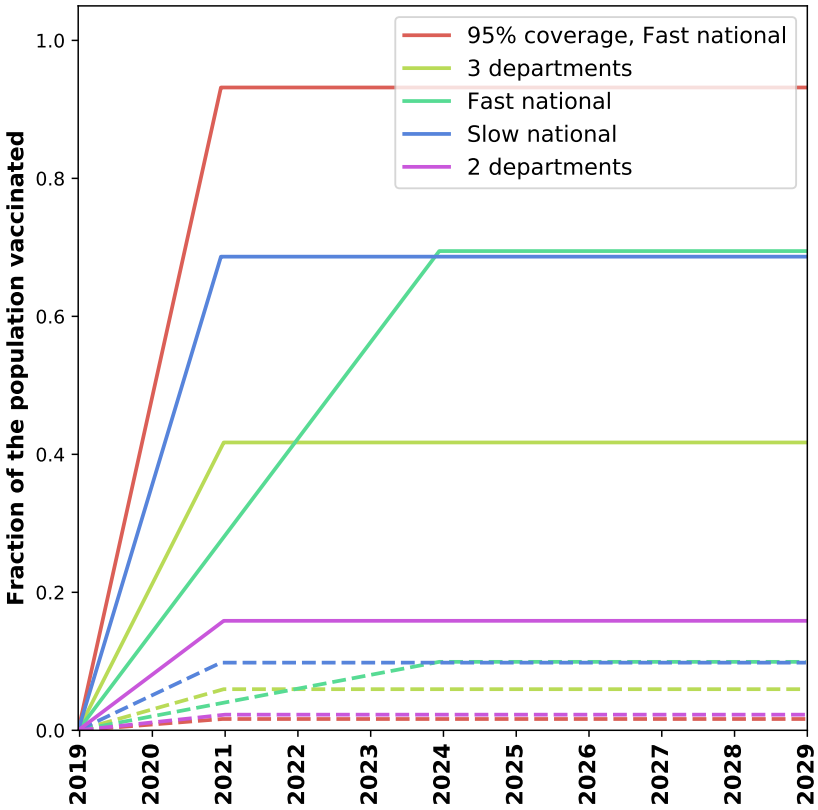

Figure 1: Fraction of the population vaccinated for each scenario considered with two doses (solid lines) or one dose (dashed lines). A full description of the vaccination scearios can be found in the meta supplement.

assumed to be less effective (0.4688*$VE_{1,2}$) in children under five. Table 1 shows all the parameters and the values used in the model.

Network structure: Political boundaries, road networks, and river systems shapefiles and an interpolated population-density surface were obtained from the dataverse website maintained by Harvard University (*3*). An elevation surface raster was also obtained and processed from (*4*). We generated a matrix D describing the average distance that individuals would be expected to travel by major road between each pair of departments as follows. We overlaid the political boundary shapefile on the interpolated population-density raster and we calculated the population-density weighted centroid for each department. Each weighted centroid was then related to nearest point along the road network shapefile, which allowed us to calculate the distance (in km) along the shortest interdepartmental path across primary, primary link, secondary, or secondary link highways. Finally, we used a standard gravity model (*5*) to describe the movement of people through the network, hence obtaining the matrix T used in the equations below, with

$$T_{ij} = v_{rate} * \frac{Pop_i * Pop_j}{D_{ij}^2}$$

where Pop_i is the total population for department i and v_{rate} is a free parameter.

As rivers and streams typically flow in one direction, the matrix characterizing the potential fluvial flow of free-living V. cholerae between a given pair of departments will not necessarily be symmetric. To generate this matrix, we used the elevation raster to determine the direction of flow for each segment of river and the total loss in elevation along its length. Since contamination of the river water depends upon the deposition of untreated human waste, we also used

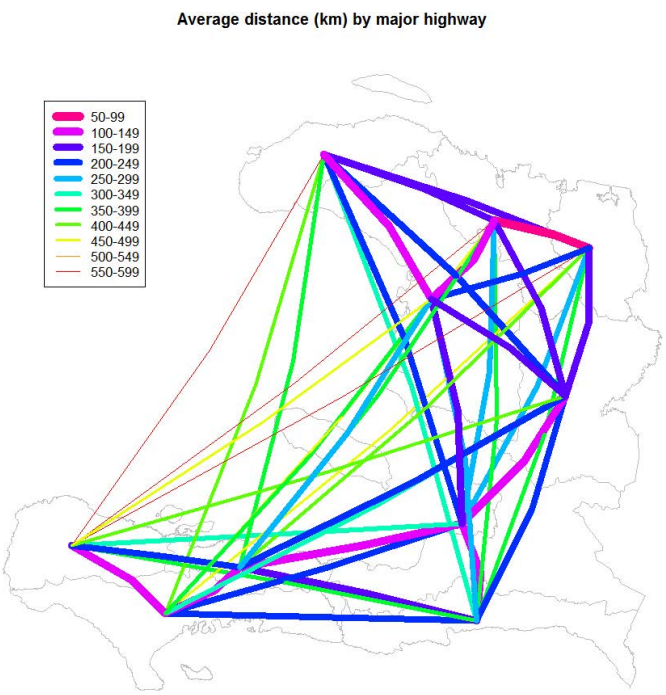

Figure 2: Schematic representation of the human mobility network used in the model. Lines between two departments represent the distance (kilometers) by major highway between the geographic centers of each pair. Line thickness is inversely proportional to distance between departments, so that lines between the closest departments are the thickest, representing the departments the most connected via human mobility.

population density surface to calculate the average population density along the banks of each segment of river. Using this information we calculated an index of potential exposure for each segment of river equal to its length in meters multiplied by the average population density along its length and divided by the amount of elevation (meters) loss. This index should take on higher values where more people are present along the banks of a river segment or where the river segment is longer, and it should decrease in value where the grade is steeper (i.e., the river is likely flowing faster and more likely to carry V. cholerae away faster). To generate the final matrix WMat describing the department to department flow of river water, we sum up the potential exposure index values for all segments of any river that flows out of department A and into department B, even if a river flows through another department in between. Where one river flows into another river, the segments of the second river downstream from the junction are considered to inherit the potential exposure index level of all segments of both rivers that are upstream from the junction. Figure 3 provides a visual representation of the water matrix obtained through this procedure.

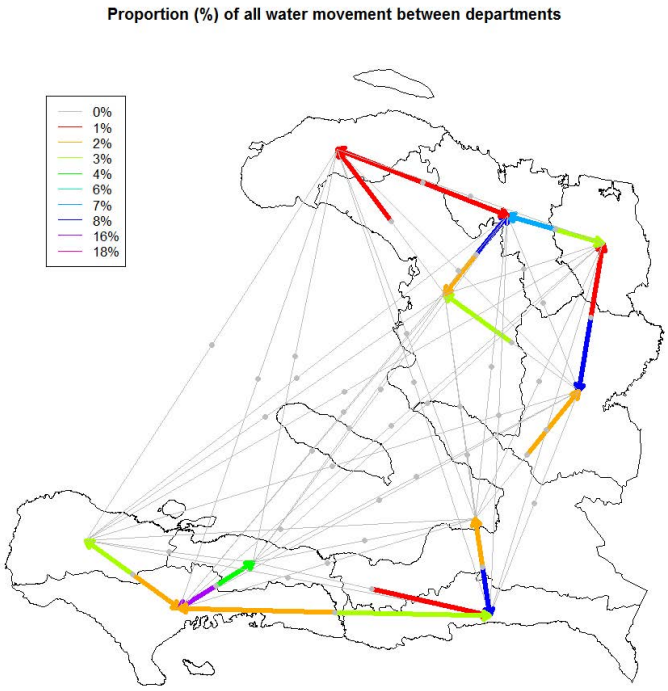

Figure 3: Schematic representation of the water network used in the model. Each arrow represents the proportion of all inter-departmental river water flow that occurs between each pair of departments. The asymmetric nature of the flow is depicted as arrows pointing toward the receiving department with the midpoint (gray circles along lines) of the line between each pair of departments as the beginning of each arrow. Gray lines indicate a zero level of water flow.

With all these assumptions, we obtain the following system of differential equations for each department (index i has been omitted from the equations for simplicity, except in the transport terms where the indices are left for clarity):

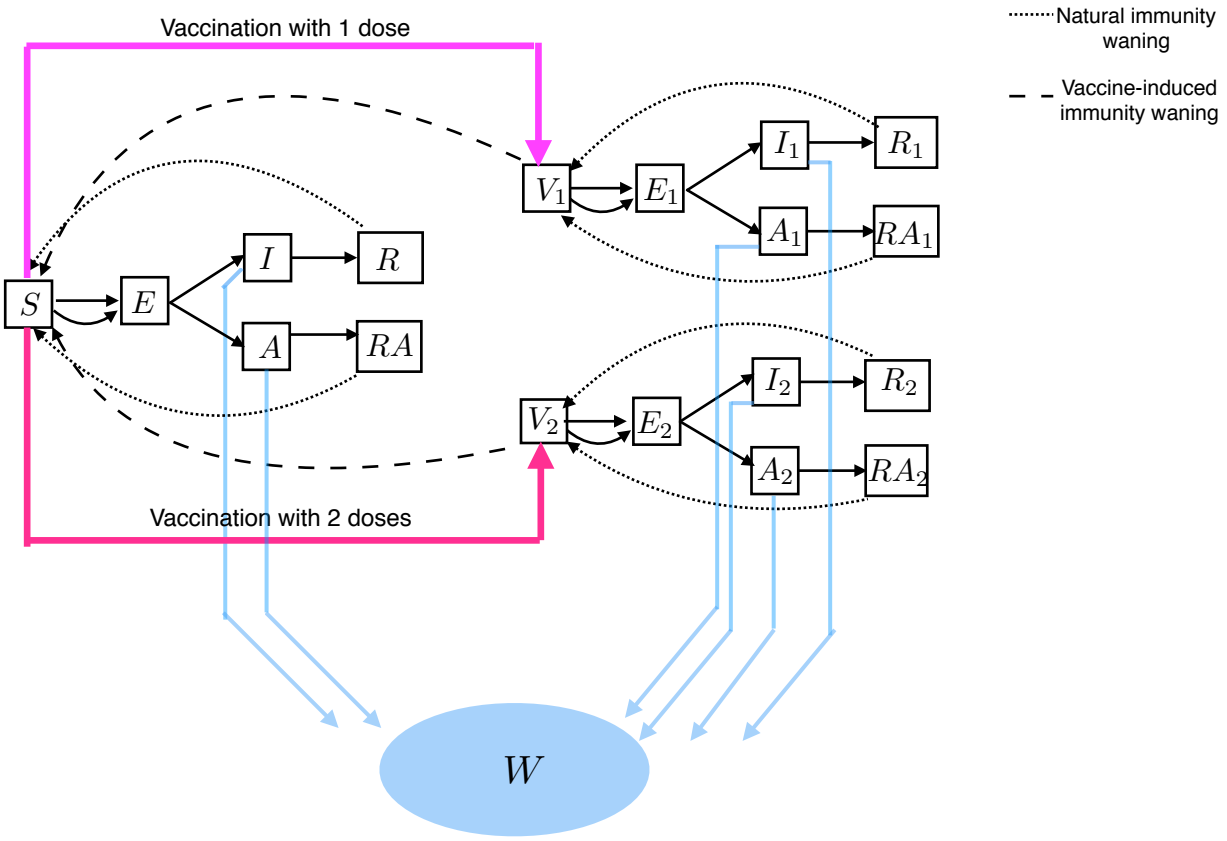

Figure 4: Schematic diagram of the cholera transmission and vaccination dynamics in a single department.

Unvaccinated individuals:

$$\frac{dS}{dt} = -\lambda S + \sigma(R + R_A) + \omega_1(V_1 + V_{1_5}) + \omega_2(V_2 + V_{2_5}) + \Big\{ \sum_{j=1}^{n} T_{ji} S_j - \sum_{j=1}^{n} T_{ij} S_i \Big\}$$

$$\frac{dE}{dt} = \lambda S - \gamma_E E + \Big\{ \sum_{j=1}^{n} T_{ji} E_j - \sum_{j=1}^{n} T_{ij} E_i \Big\}$$

$$\frac{dI}{dt} = k\gamma_E E - (\gamma) I + \Big\{ \sum_{j=1}^{n} T_{ji} I_j - \sum_{j=1}^{n} T_{ij} I_i \Big\}$$

$$\frac{dA}{dt} = (1-k)\gamma_E E - \gamma A + \Big\{ \sum_{j=1}^{n} T_{ji} A_j - \sum_{j=1}^{n} T_{ij} A_i \Big\}$$

$$\frac{dR}{dt} = \gamma I - \sigma R + \Big\{ \sum_{j=1}^{n} T_{ji} R_j - \sum_{j=1}^{n} T_{ij} R_i \Big\}$$

$$\frac{dR_A}{dt} = \gamma A - \sigma R_A + \Big\{ \sum_{j=1}^{n} T_{ji} R_{Aj} - \sum_{j=1}^{n} T_{ij} R_{Ai} \Big\}$$

Vaccinated individuals:

$$\frac{dV_1}{dt} = -\theta_1(\lambda V_1) + \sigma(R_1 + R_{A1}) - \omega_1 V_1 + \Big\{ \sum_{j=1}^{n} T_{ji} V_{1j} - \sum_{j=1}^{n} T_{ij} V_{1i} \Big\}$$

$$\frac{dE_1}{dt} = \theta_1(\lambda V_1) - \gamma_E E_1 + \Big\{ \sum_{j=1}^{n} T_{ji} E_{1j} - \sum_{j=1}^{n} T_{ij} E_{1i} \Big\}$$

$$\frac{dI_1}{dt} = k\gamma_E E_1 - \gamma I_1 + \Big\{ \sum_{j=1}^{n} T_{ji} I_{1j} - \sum_{j=1}^{n} T_{ij} I_{1i} \Big\}$$

$$\frac{dA_1}{dt} = (1-k)\gamma_E E_1 - \gamma A_1 + \Big\{ \sum_{j=1}^{n} T_{ji} A_{1j} - \sum_{j=1}^{n} T_{ij} A_{1i} \Big\}$$

$$\frac{dR_1}{dt} = \gamma I 1 - \sigma R_1 + \Big\{ \sum_{j=1}^{n} T_{ji} R_{1j} - \sum_{j=1}^{n} T_{ij} R_{1i} \Big\}$$

$$\frac{dR_{A1}}{dt} = \gamma A_1 - \sigma R_{A1} + \Big\{ \sum_{j=1}^{n} T_{ji} R_{A1j} - \sum_{j=1}^{n} T_{ij} R_{A1i} \Big\}$$

$$\frac{dV_2}{dt} = -\theta_2(\lambda V_2) + \sigma(R_2 + R_{A2}) - \omega_2 V_2 + \Big\{ \sum_{j=1}^{n} T_{ji} V_{2j} - \sum_{j=1}^{n} T_{ij} V_{2i} \Big\}$$

$$\frac{dE_2}{dt} = \theta_2(\lambda V_2) - \gamma_E E_2 + \Big\{ \sum_{j=1}^{n} T_{ji} E_{2j} - \sum_{j=1}^{n} T_{ij} E_{2i} \Big\}$$

$$\frac{dI_2}{dt} = k\gamma_E E_2 - \gamma I_2 + \Big\{ \sum_{j=1}^{n} T_{ji} I_{2j} - \sum_{j=1}^{n} T_{ij} I_{2i} \Big\}$$

$$\frac{dA_2}{dt} = (1-k)\gamma_E E_2 - \gamma A_2 + \Big\{ \sum_{j=1}^{n} T_{ji} A_{2j} - \sum_{j=1}^{n} T_{ij} A_{2i} \Big\}$$

$$\frac{dR_2}{dt} = \gamma I_2 - \sigma R_2 + \Big\{ \sum_{j=1}^{n} T_{ji} R_{2j} - \sum_{j=1}^{n} T_{ij} R_{2i} \Big\}$$

$$\frac{dR_{A2}}{dt} = \gamma A_2 - \sigma R_{A2i} + \Big\{ \sum_{j=1}^{n} T_{ji} R_{A2j} - \sum_{j=1}^{n} T_{ij} R_{A2i} \Big\}$$

$$\frac{dV_{1_5}}{dt} = -\theta_{1_5}(\lambda V_{1_5}) + \sigma(R_{1_5} + R_{A1_5}) - \omega_1 V_{1_5} + \Big\{ \sum_{j=1}^{n} T_{ji} V_{1_5 j} - \sum_{j=1}^{n} T_{ij} V_{1_5 i} \Big\}$$

$$\frac{dE_{1_5}}{dt} = \theta_{1_5}(\lambda V_{1_5}) - \gamma_E E_{1_5} + \Big\{ \sum_{j=1}^{n} T_{ji} E_{1_5 j} - \sum_{j=1}^{n} T_{ij} E_{1_5 i} \Big\}$$

$$\frac{dI_{1_5}}{dt} = k\gamma_E E_{1_5} - \gamma I_{1_5} + \Big\{ \sum_{j=1}^{n} T_{ji} I_{1_5 j} - \sum_{j=1}^{n} T_{ij} I_{1_5 i} \Big\}$$

$$\frac{dA_{1_5}}{dt} = (1-k)\gamma_E E_{1_5} - \gamma A_{1_5} + \Big\{ \sum_{j=1}^{n} T_{ji} A_{1_5 j} - \sum_{j=1}^{n} T_{ij} A_{1_5 i} \Big\}$$

$$\frac{dR_{1_5}}{dt} = \gamma I_{1_5} - \sigma R_{1_5} + \Big\{ \sum_{j=1}^{n} T_{ji} R_{1_5 j} - \sum_{j=1}^{n} T_{ij} R_{1_5 i} \Big\}$$

$$\frac{dR_{A1_5}}{dt} = \gamma A_{1_5} - \sigma R_{A1_5} + \Big\{ \sum_{j=1}^{n} T_{ji} R_{A1_5 j} - \sum_{j=1}^{n} T_{ij} R_{A1_5 i} \Big\}$$

$$\frac{dV_{2_5}}{dt} = -\theta_{2_5}(\lambda V_{2_5}) + \sigma(R_{2_5} + R_{A2_5}) - \omega_2 V_{2_5} + \Big\{ \sum_{j=1}^{n} T_{ji} V_{2_5 j} - \sum_{j=1}^{n} T_{ij} V_{2_5 i} \Big\}$$

$$\frac{dE_{2_5}}{dt} = \theta_{2_5}(\lambda V_{2_5}) - \gamma_E E_{2_5} + \Big\{ \sum_{j=1}^{n} T_{ji} E_{2_5 j} - \sum_{j=1}^{n} T_{ij} E_{2_5 i} \Big\}$$

$$\frac{dI_{2_5}}{dt} = k\gamma_E E_{2_5} - \gamma I_{2_5} + \Big\{ \sum_{j=1}^{n} T_{ji} I_{2_5 j} - \sum_{j=1}^{n} T_{ij} I_{2_5 i} \Big\}$$

Water equations:

$$\frac{dW}{dt} = (\mu(I + I_1 + I_2 + I_{1_5} + I_{2_5}) + \mu_A(A + A_1 + A_2 + A_{1_5} + A_{2_5})) - \delta W + w_r\Big\{\sum_{j=1}^{n}(\text{WMat}_{ji}W_j - \text{WMat}_{ij}W_i)\Big\}$$

where $\theta_1 = 1 - VE_1$ and $\theta_2 = 1 - VE_2$. In addition, T is the travel matrix representing travel through major highways and WMat is the river matrix representing the flow of water through Haiti. We assumed that $\beta_A = red_\beta\beta$ and $\mu_A = red_\mu\mu$ so that the force of infection λ is given by:

$$\lambda = 0.5\big(1 + \alpha_s \cos(2\pi\frac{t}{p_s})\big)\frac{\beta_W W}{Sat + W} + \big\{\beta(I + I_1 + I_2 + I_{1_5} + I_{2_5}) + \beta_A(A + A_1 + A_2 + A_{1_5} + A_{2_5})\big\}$$

Table 1 provides a list of the parameters used and indicates whether they were fit to the data or assumed.

Table 1: Parameters definitions and values.

Model parameter	Symbol	Value	Units	Reference
Decay of cholera in water	δ	$1/3$	weeks^{-1}	mean of range given in (*6*)
Fraction of symptomatic infections	k	0.2	–	(*7, 8*)
Infectious period	$1/\gamma$	1	weeks	mean of range given in (*6*)
Latent period	$1/\gamma_E$	0.18	weeks	mean of range given in (*6*)
Relative infectiousness of asymptomatic infections	red_β	10^{-3}	–	(*9*)
Relative shedding of cholera for asymptomatic infections	red_μ	10^{-7}	–	(*9*)
Duration of natural immunity	$1/\sigma$	5 [a]	years	–
Concentration of *V. cholerae* at which the infection rate is 50% of the maximum infection rate	Sat	10^5	cells/ml	mean of range given in (*6*)
Amplitude of seasonal forcing	α_s	0.4 [b]	–	–
Period of seasonal forcing	p_s	52[c]	weeks	–
Person-to-person infection rate	β	$9.9 * 10^{-7}$	week^{-1}	fit
Environmental infection rate	β_W	$4.03 * 10^{-2}$		fit
Symptomatic shedding rate	μ	$3.98 * 10^3$, 2.58		fit
Fraction of susceptible population in 03/2014	f	0.75	–	
Travel matrix multiplier	v_{rate}	10^{-12} [d]	–	
Rate of loss of vaccine-induced immunity for one dose	ω_1	1	years^{-1}	(*2*)
Rate of loss of vaccine-induced immunity for two doses	ω_2	$1/5$	years^{-1}	(*2*)
Vaccine efficacy for one dose	VE_1	varied		see text
Vaccine efficacy for two doses	VE_2	varied		see text
Reporting rate (symptomatic infections)	ρ	0.2	–	(*10*)

[a] Natural immunity was assumed to last as long as vaccine-induced immunity.
[b] α_s was chosen to mimic the observed cholera seasonal patterns in Haiti
[c] p_s was chosen to mimic the observed cholera seasonal patterns in Haiti
[d] v_rate was chosen to mimic the distribution of infections in Haiti

2 Model Fitting/Calibration

We fit our model to the weekly reported cases provided by Haiti MSPP in two stages. First, we utilized reported cases from 10/2010 to 03/2014 and used Ordinary Least Squares (OLS, leastsq function from optimize package in Scipy, python (*11*)) to fit weekly incidence provided by the model to weekly reported cases. We assumed that only 20% of

cholera infections become symptomatic (*7,8*) and out of those, only 20% would be reported (*10*). All model parameters were fixed except for three parameters: the person-to-person infection rate, the environmental infection rate, and the symptomatic shedding rate (β, β_W and μ). When possible, fixed parameters were directly taken from data, or they were taken as averages of known ranges (see Table 1). The time period utilized in this first stage corresponds to first phase of the explosive cholera epidemic in Haiti, which occurred after the introduction of cholera in Haiti. We ran our OLS algorithm with 1000 different random starting points and used the mean of the resulting best parameter sets to run the simulations for our "best fit" (see figure 5A). We then focus on the endemic dynamics of cholera transmission, going from 03/2014 to 01/2019. There were several interventions against cholera in Haiti during this period, including vaccination campaigns, WASH interventions and rapid response teams (*12*), all of which might have altered the course of the cholera dynamics. Hence, we re-fit our model for this part, but we only re-fit the parameter μ while keeping the infection rates (β and β_W) constant. We could have chosen any of these three parameters to refit our data. The parameter μ was chosen as a way to represent that most of these interventions might have decreased the rate at which symptomatic infected people would have shed cholera in the environment.

In a serosurvey conducted by Jackson *et al* (*7*), it was found that 39% of the population were seropositive for V. cholerae antibody tests. However, only 28% and 23% of the population were positive for specific antibody tests for the Ogawa and Inaba strains respectively. The study was conducted in the Grande Saline Commune, which was particularly hit by the cholera epidemic during 2010 and 2011. Based on these numbers, we performed sensitivity analysis and found that our model fit the data best when we assumed only 25% of the population was immune to cholera. Hence, we assumed that 75% of the population was susceptible to cholera on March 2014 (see figure 5B).

3 Assessment of Model Fit

To assess the model fit we performed extensive sensitivity analysis. We used an interval centered around each fit parameter (we used 25% of the parameter's value as the 1/2 width of the interval) and around the fraction susceptible in March 2014 (we used 10% of this value as the 1/2 width of the sampling interval) to obtain a hyper cube around our fit values. We then used a Latin hypercube sampling scheme to obtain 1000 parameter sets. These parameter sets were used to run all the vaccination scenarios, and were used to compute the lower and upper bounds presented in the main results. Briefly, our results were mostly sensitive to the fraction of the population susceptible on March 2014, but were not sensitive to the other parameters. More importantly, vaccination strategies where all the departments were vaccinated resulted in cholera elimination for over 90% of the parameters tested.

Code available at https://github.com/lulelita/HaitiCholeraMultiModelingProject.

References

1. M. E. Halloran, C. J. Struchiner, I. M. Longini, *American Journal of Epidemiology* **146**, 789 (1997).
2. Q. Bi, *et al.*, *Lancet Infect Dis* **17**, 1080 (2017).
3. L. Berman, "Haiti Earthquake Data (VECTORS)" (2015).
4. DIVA-GIS, Haiti Elevation Grid File.
5. J. J. Lewer, H. V. den Berg, *Economics Letters* **99**, 164 (2008).
6. I. C.-H. Fung, *Emerg Themes Epidemiol* **11**, 1 (2014).
7. B. R. Jackson, *et al.*, *Am J Trop Med Hyg* **89**, 654 (2013).
8. WHO, Prevention and control of cholera outbreaks: WHO policy and recommendations, `https://www.who.int/cholera/technical/prevention/control/en/` (2010).
9. E. J. Nelson, J. B. Harris, J. G. J. Morris, S. B. Calderwood, A. Camilli, *Nat Rev Microbiol* **7**, 693 (2009).
10. WHO, The global burden of cholera, https://www.who.int/bulletin/volumes/90/3/11-093427/en/ (2012).

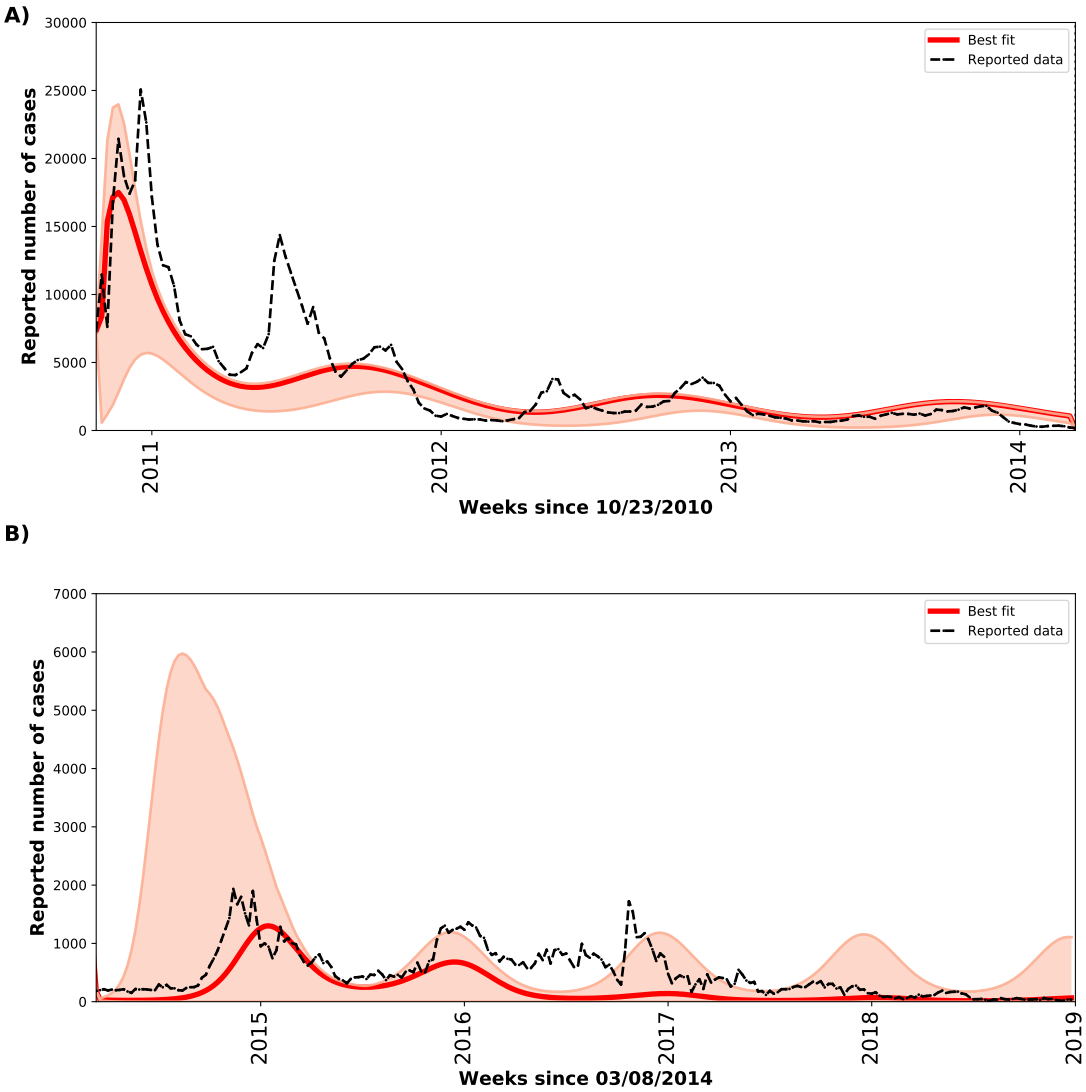

Figure 5: Model fit (red line) and reported number of cases (black dashed line) for the cholera epidemic in Haiti. Case reporting started on October $22^n d$, 2010. The light orange area represents the range in the number of cases generated by the sensitivity analysis.

11. Eric Jones and Travis Oliphant and Pearu Peterson and others, SciPy: Open source scientific tools for Python, http://www.scipy.org/ (2001–).

12. WHO, Haiti's Ministry of Health successfully vaccinates 729,000 persons against cholera, https://www.paho.org/hq/index.php?option=com_content&view=article&id=12771:haiti-ministry-of-health-vaccinates-729000-persons-against-cholera&Itemid=135&lang=en (2016).