Supplemental Methods for Model-1

Elizabeth C. Lee, Andrew S. Azman, and Justin Lessler
John Hopkins Bloomberg School of Public Health

1 Model Description

In this study, we adapted a deterministic compartmental model that had been used previously to compare the potential impact of one-dose and two-dose oral cholera vaccine regimens in Port-au-Prince for use at the country-level in Haiti (*1*). In brief, the original version of this model made simple assumptions about cholera transmission and disease course; the model assumed Susceptible - Exposed - Infectious - Recovered (SEIR) dynamics; the city population of Port-au-Prince experienced equal potential disease contact with all other individuals in the population (homogeneous mixing); and disease transmission occurred through person-to-person contact with added variability according to a seasonally-varying transmission term. Due to the outbreak-response-focus of the initial study, waning immunity was not considered in this original model. Importantly, there were no explicit model components capturing transmission due to contact with contaminated environments, seasonality in transmission due to climatic changes (e.g., rainfall), or spatial heterogeneities in transmission.

1.1 Base Model Structure

For this study, we constructed a continuous-time, discrete-state Susceptible - Exposed - Infectious - Asymptomatic - Recovered - Susceptible (SEIARS) compartmental model with population dynamics. The Infectious state represents symptomatic, infectious individuals, while the Asymptomatic state represents asymptomatic individuals, who are only a fraction as infectious as Infectious individuals. We model loss of immunity among recovered individuals, births into the susceptible population, and deaths by natural causes from all compartments due to the long-term projection of vaccine effects. Compared to the original model, we allow for the possibility of non-homogeneous mixing among potential disease-causing contacts in the population, using methods similar to those used in classic time series Susceptible - Infectious - Recovered (TSIR) models (*2, 3*), due to observed differences in disease dynamics across departments in Haiti and the country-level scale of our model. The original model included a term for introductions into the modeled city from the rest of the country, but we removed this term since introductions into Haiti are thought to be extremely rare (*4*).

The underlying dynamics of our model may be represented by the following set of differential equations:

$$\begin{aligned}
\frac{dS}{dt} &= -\lambda S - \delta S + \alpha R + \mu N \\
\frac{dE}{dt} &= -\sigma(1-\theta_0)E - \sigma\theta_0 E - \delta E + \lambda S \\
\frac{dI}{dt} &= -\gamma I - \delta I + \sigma(1-\theta_0)E \\
\frac{dA}{dt} &= -\gamma A - \delta A + \sigma\theta_0 E \\
\frac{dR}{dt} &= -\alpha R - \delta R + \gamma I + \gamma A
\end{aligned}$$

where λ is the force of infection, δ is the natural death rate, α is the rate of recovery from infection, μ is the birth rate, $1/\sigma$ is the latent period, θ_0 is the proportion of Exposed individuals that become Asymptomatic infections, $1/\gamma$ represents the mean infectious period, and S, E, I, A, R, N represent the Susceptible, Exposed, symptomatic

Infectious, Asymptomatic infectious, Recovered, and total populations, respectively. The force of infection is modeled as:

$$\lambda = (I + (1 - \kappa)A)^{\nu}\beta/N,$$

where κ is the relative reduction in infectiousness for Asymptomatic individuals as compare to Infectious individuals, ν is the population mixing coefficient that captures deviations from homogeneous mixing, N is the (current) total population size, and β is the seasonal transmission term. We note that only vaccinated individuals may develop Asymptomatic infections.

Seasonal transmission is modeled as:

$$\beta = \sum_{j=1}^{6} \beta_j s_j$$

from j periodic basis spline terms (s_j) with six basis splines each with six degrees of freedom (*1*). We construct the seasonal transmission term with basis splines in order to calibrate the shape, phase, and magnitude of the seasonal component simultaneously; this enables variable complexity in the seasonality of cholera transmission.

We assumed that our cholera incidence surveillance data did not represent all symptomatic cholera infections in Haiti, and thus would be best represented by a partially-observed Markov Process. Thus, the likelihood of our model was

$$L(\boldsymbol{X}|cases) = NegBinom(\rho\xi, \tau)$$

where $\boldsymbol{X}$ is the vector of model parameters and the observation process is parameterized with the mean, which was the product of reporting rate ρ and incidence of new symptomatic infections ξ, and the inverse dispersion τ.

1.2 Modeling Vaccination Campaigns

1.2.1 Vaccine Effects

Vaccination has two direct effects in the model:

1. Vaccinated individuals in the Exposed compartment had some probability of becoming Asymptomatic infections; this probability ψ_t was equal to adult vaccine efficacy t months after vaccination. Multiple waning vaccine efficacy scenarios were pre-determined by the modeling exercise.

2. Asymptomatic individuals are assumed to be less infectious than Infectious individuals; the force of infection term λ modifies the relative reduction in infectiousness with the parameter κ, which was fixed to 0.95 in our model (Asymptomatic individuals were 95% less infectious than Infectious individuals).

1.2.2 Tracking Vaccinated Individuals

Vaccination campaigns are tracked as cohorts for each department, and each cohort is represented by another set of Susceptible - Exposed - Infectious - Asymptomatic - Recovered compartments. Thus, the base model with vaccination is modified to:

$$\frac{dS}{dt} = -\lambda S - \delta S + \alpha R + \mu N - \sum_{i=1}^{d_i} \eta_{ik} S$$
$$\frac{dE}{dt} = -\sigma(1-\theta_0)E - \sigma\theta_0 E - \delta E + \lambda S - \sum_{i=1}^{d_i} \eta_{ik} I$$
$$\frac{dI}{dt} = -\gamma I - \delta I + \sigma(1-\theta_0)E - \sum_{i=1}^{d_i} \eta_{ik} E$$
$$\frac{dA}{dt} = -\gamma A - \delta A + \sigma\theta_0 E - \sum_{i=1}^{d_i} \eta_{ik} A$$
$$\frac{dR}{dt} = -\alpha R - \delta R + \gamma I + \gamma A - \sum_{i=1}^{d_i} \eta_{ik} R$$

where i indicates the department, d_i is the number of departments vaccinated in a given scenario, and η_{ik} is the vaccination rate in a department in model week k. While a real departmental vaccination campaign would likely take place over many weeks, we push all of the vaccines for a given department into our model in a single week at the end of the rollout period. This model implementation enables easier tracking of waning vaccine immunity and is conservative in the effects of the vaccine since vaccine effects are experienced only at the very end of the vaccination campaign period. Thus, the vaccination rate η_{ik} is only non-zero during a single week that represents the rollout of a given departmental campaign, according to Table 5. Each of the d_i vaccination cohorts may be represented as:

$$\frac{dSv_i}{dt} = -\lambda Sv_i - \delta Sv_i + \alpha Rv_i + \mu Nv_i + \eta_{ik} S$$
$$\frac{dEv_i}{dt} = -\sigma(1-\theta_{vk})Ev_i - \sigma\theta_{vk} Ev_i - \delta Ev_i + \lambda Sv_i + \eta_{ik} E$$
$$\frac{dIv_i}{dt} = -\gamma Iv_i - \delta Iv_i + \sigma(1-\theta_{vk})Ev_i + \eta_{ik} I$$
$$\frac{dAv_i}{dt} = -\gamma Av_i - \delta Av_i + \sigma\theta_{vk} Ev_i + \eta_{ik} A$$
$$\frac{dRv_i}{dt} = -\alpha Rv_i - \delta Rv_i + \gamma Iv_i + \gamma Av_i + \eta_{ik} R$$

where θ_{vk} represents is the population vaccine efficacy at week k. Population vaccine efficacy was calculated as

$$\theta_{vk} = \psi_t * (1 - (1 - 0.4688) * 0.11),$$

where ψ_t, the adult vaccine efficacy at month t after vaccination, is scaled to the population-weighted average of vaccine efficacy among people over and under five years old. Vaccine efficacy among children under five was assumed to be 0.4688 times as effective as vaccine efficacy among adults (*5*), and 11% of the Haitian population was identified as under five years old (*6*). A single dose of cholera vaccine was assumed to confer the same level of protection as two doses in the first year and no protection in subsequent years of vaccination.

The force of infection λ is the same for Susceptible individuals in all vaccination cohorts

$$\lambda = ((I + \sum_{i=1}^{d_i} Iv_i) + (1-\kappa)(A + \sum_{i=1}^{d_i} Av_i))^{\nu} \beta / N,$$

so the vaccination cohort structure is used solely for tracking the time of vaccination and applying the appropriate vaccine efficacy for a given time point in the model. Figure 1 depicts the model structure for a two department vaccination campaign.

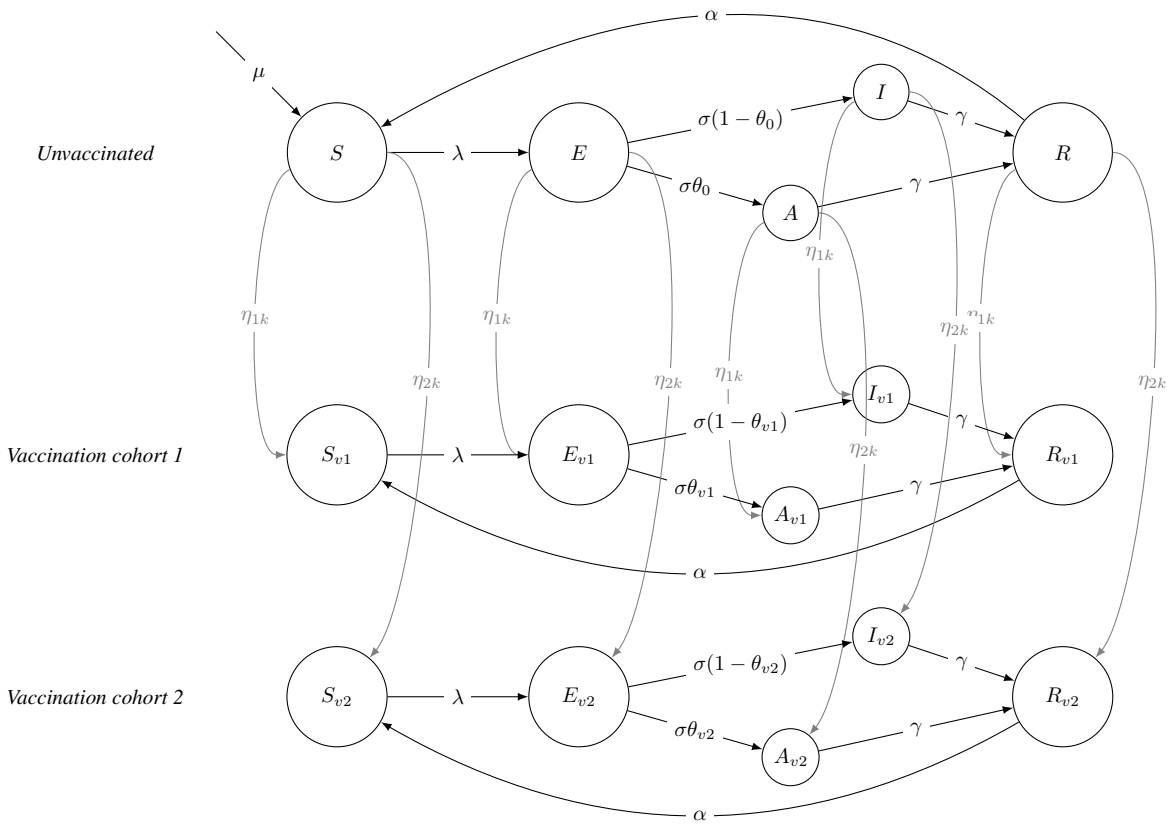

Figure 1: Demonstrative flow diagram of the model structure for a two department vaccination campaign. Death rates (δ) from all compartments are not shown to improve diagram readability.

2 Model Calibration

We calibrated our model in two stages using the R package POMP (v1.19) (*7, 8*). The first stage calibrated the model during the "epidemic" phase of the data, which we assumed was from October 2010 through March 2015 with the second "epidemic" stage going from April 2015 through January 2019 (end of the calibration period). The model was parameterized with parameters scaled to week-long time steps and model states were extracted at weekly intervals, in line with data reports.

2.1 Epidemic phase

We profiled four parameters using Latin Hypercube Sampling, reporting rate (ρ), the first component of the basis spline (β_1), population mixing coefficient (ν), and inverse dispersion (θ), to generate 300 sets of starting values that covered a wide range of parameter space (Table 1). Fixed and fitted parameters are described in Tables 2 and 3.

We first used trajectory matching to perform maximum likelihood estimation on our model's deterministic skeleton. We passed the parameter estimates from trajectory matching as starting values for the improved iterated filtering algorithm (*9*). We iterated 100 particles 100 times for each set of starting values and the resulting parameter estimates among parameter sets with zero filtering failures are reported in Table 3.

Table 1: Ranges for starting parameters profiled in the epidemic phase.

Parameter	Range	Rationale	Reference
ρ	$[10^{-8}, 1]$	reporting proportion must exceed zero	
β_1	$[10^{-9}, 10]$	beta spline terms are dependent; for a given beta term, a value of 10 corresponds to a basic reproductive number (R0) of roughly 2.85	
ν	$[.95, 1]$	mixing coefficients in other model contexts are usually close to unity	(*2, 3*)
τ	$[1, 20]$	dispersion in the observation process likely exceeds that of a Poisson distribution if correctly specified	

Table 2: Parameters that were fixed across both model calibration stages.

Parameter	Value	Description	Reference
$N.0$	10,911,819	population of Haiti at the start of the study period	(*10*)
μ	.43	birth rate per 1000 population per week	(*11*)
δ	.14	natural death rate per 1000 population per week	(*11*)
$1/\sigma$	1.4	latent period (days)	(*12*)
$1/\gamma$	2	infectious period (days)	(*13*)
$1/\alpha$	8	mean duration of natural immunity (years)	range (*14, 15*)
θ_0	0	proportion of non-vaccinated Exposed individuals that develop Asymptomatic infection	assumed
κ	.95	proportion of reduction in transmissibility among Asymptomatic infections	assumed

Table 3: Parameters estimates or values for the epidemic phase of the data after model calibration.

Parameter	Calibrated	Median Estimate [Range]	Description
β_1	yes	5.8 [.04,12.1]	transmission parameter (1st component of basis spline)
β_2	yes	3.5 [0,10.3]	transmission parameter (2nd component of basis spline)
β_3	yes	4.4 [.18,9.7]	transmission parameter (3rd component of basis spline)
β_4	yes	2.8 [0,8.3]	transmission parameter (4th component of basis spline)
β_5	yes	6.2 [.8,12.2]	transmission parameter (5th component of basis spline)
β_6	yes	2.4 [0,7.7]	transmission parameter (6th component of basis spline)
ν	yes	.96 [.90,1.0]	population mixing coefficient (1 represents homogeneous mixing)
ρ	yes	.34 [.13,1]	reporting fraction
τ	yes	4.0 [1.7,5.9]	inverse dispersion for the negative binomial measurement model
$E.0$	yes	405 [1,1752]	initial number of Exposed individuals
$I.0$	yes	518 [1,2271]	initial number of Infected individuals

2.2 Endemic phase

After calibrating the epidemic phase of the data, we passed the final epidemic parameter estimates and state values (number of individuals in each compartment) at the final time point as starting parameters for the endemic model. The starting state values were fixed for each parameter set, but all other parameters calibrated in the epidemic phase were re-calibrated in the endemic phase with the iterated filtering algorithm (Table 4).

Table 4: Parameters estimates or values for the endemic phase of the data after model calibration.

Parameter	Calibrated	Median Estimate [Range]	Description
β_1	yes	5.4 [.06,80.4]	transmission parameter (1st component of basis spline)
β_2	yes	3.3 [0,36.2]	transmission parameter (2nd component of basis spline)
β_3	yes	4.3 [.07,63.1]	transmission parameter (3rd component of basis spline)
β_4	yes	3.5 [0,12.3]	transmission parameter (4th component of basis spline)
β_5	yes	5.0 [.3,98.8]	transmission parameter (5th component of basis spline)
β_6	yes	2.7 [0,9.7]	transmission parameter (6th component of basis spline)
ν	yes	.96 [.64,1.0]	population mixing coefficient (1 represents homogeneous mixing)
ρ	yes	.34 [.11,1]	reporting fraction
τ	yes	4.1 [2.1,7.0]	inverse dispersion for the negative binomial measurement model
$E.0$	no	248 [1,1693]	initial number of Exposed individuals, simulated from the epidemic model fits
$I.0$	no	324 [1,2160]	initial number of Infected individuals, simulated from the epidemic model fits

3 Assessment of Model Fit

We assessed the model fit visually over the entire time series and for the tail-end of the endemic phase due to its importance in realistic forward projections (Figure 2). While individual realizations simulated from a single parameter set retained substantial variability, summarizing the fits across multiple realizations and parameter sets presented a reasonable fit across both epidemic and endemic phases of the data. We also examined the reasonableness of all fitted parameter values (Figure 3 and Figure 4). In this examination, we note that estimates for reporting rate ρ are bimodal in both phases, with modes at 0.25 and near 1. We removed parameter sets in the endemic phase with very low log-likelihoods (log-likelihood ¡ -3000) because they corresponded to outlying calibrated parameter values. We used all remaining parameter sets to project the long-term effects of mass oral cholera vaccination campaigns in Haiti.

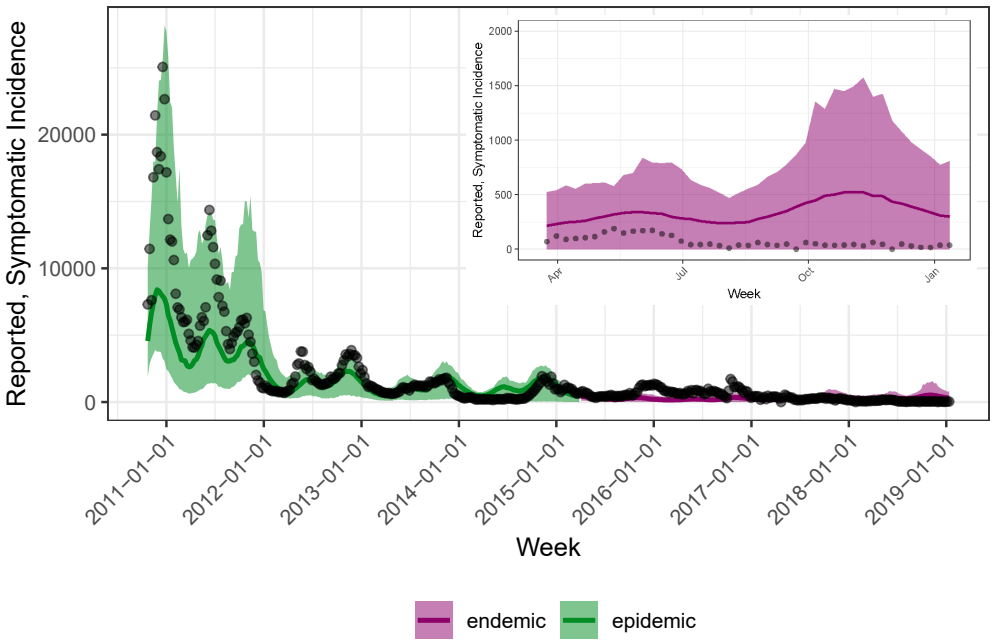

Figure 2: Model calibration of observed symptomatic cases for epidemic and endemic phases together. The inset represents the model calibration of observed symptomatic cases for the end of the endemic phase. Colored lines represent the median of the model calibrations across parameter sets and colored ribbons represent the range of model calibration values from the 2.5 to 97.5 percentile. Black dots represent Haiti cholera surveillance data.

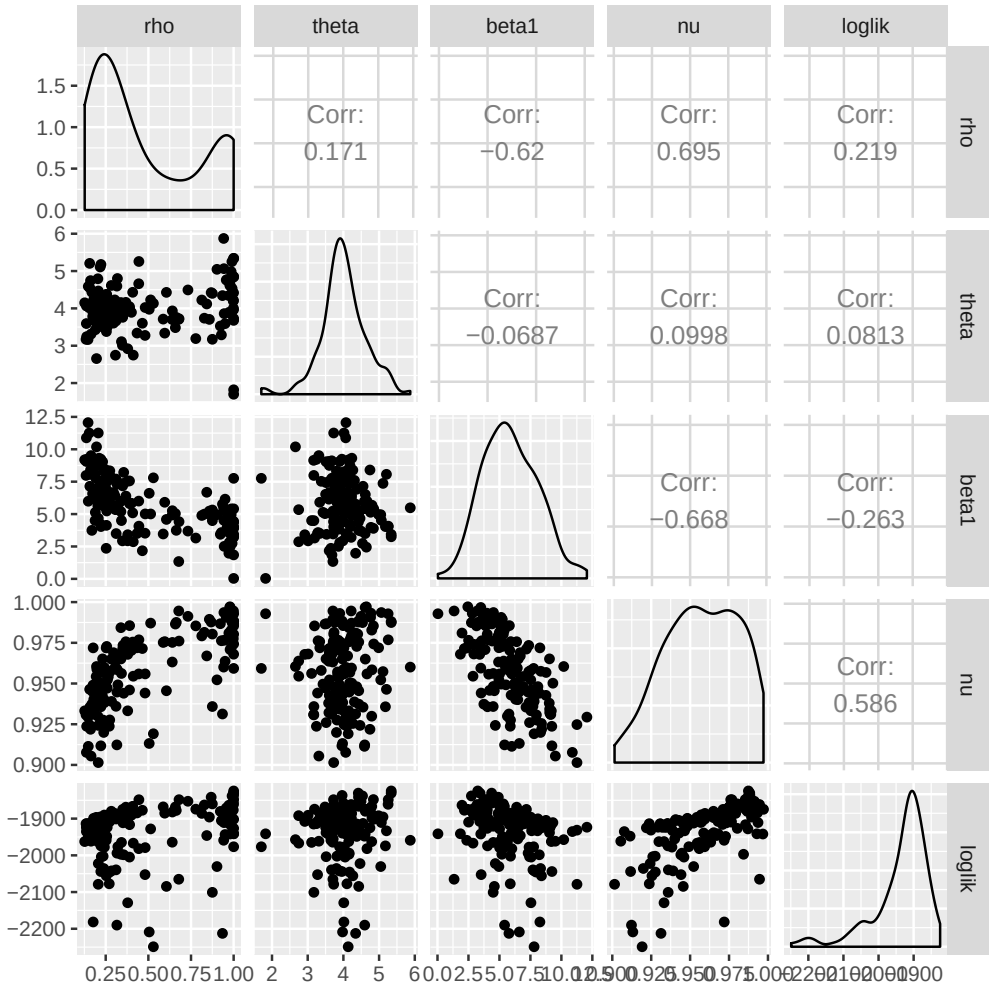

Figure 3: Distribution of parameter estimates and log likelihood for epidemic phase.

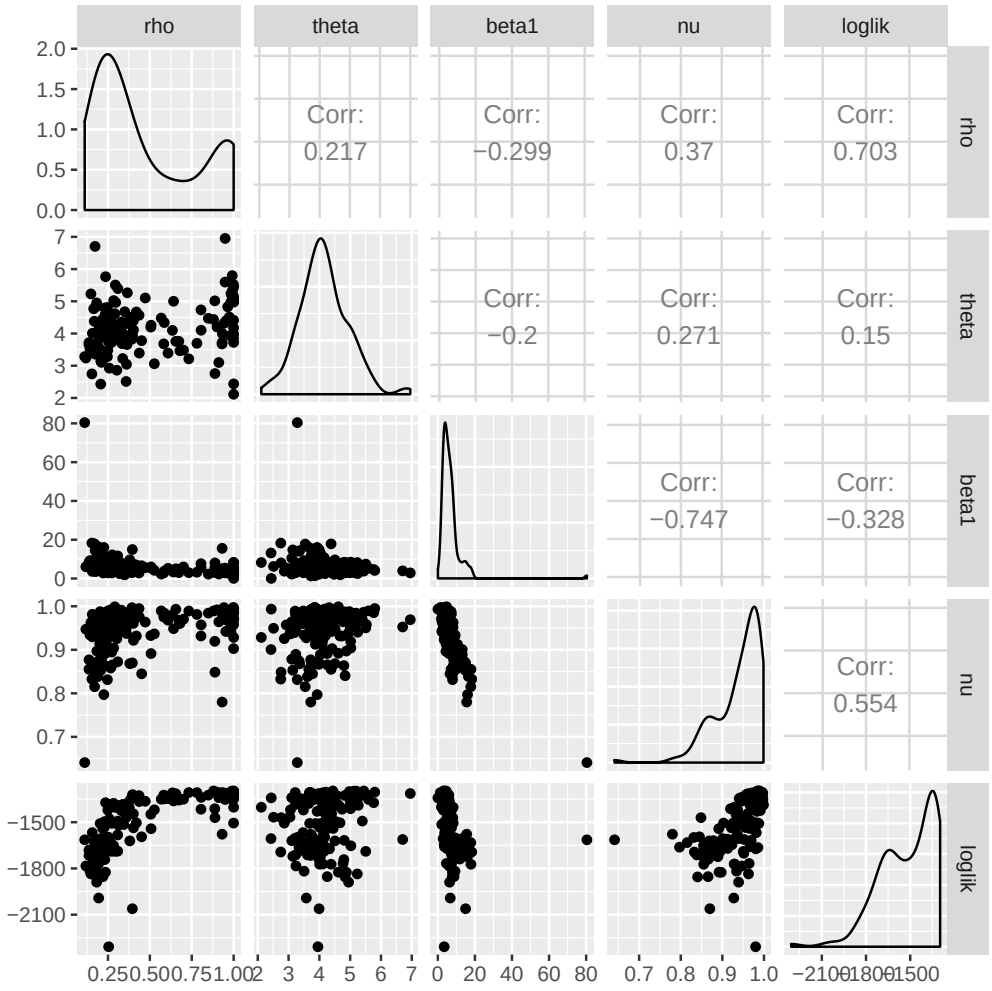

Figure 4: Distribution of parameter estimates and likelihood for endemic phase.

4 Simulation of Vaccination Campaigns

We used the estimated endemic phase parameters and the simulated states at the final time point to generate stochastic forecasts of observed and true cholera incidence in Haiti for 10 years. We compared model projections representing the status quo, where current control measures continue constantly over the 10 year forecast, to 36 vaccination scenarios that vary in their assumptions of the speed and geographic coverage of vaccination campaigns, levels of 2-dose and 1-dose vaccination coverage, and the efficacy of vaccine-induced immunity over time. Projections were simulated in POMP using an Euler approximation at day-long time intervals (delta.t = 1/7; models were calibrated at week-long time intervals). The observation process was stochastically simulated on top of the latent Markov process where the number of observed cases was assumed to follow a negative binomial distribution, as follows:

$$cases = NegBinom(\rho\xi, \tau)$$

4.1 Deployment of Vaccination Campaigns

Our model had only country-level disease dynamics, yet the pre-determined vaccination scenarios called for department-specific vaccination campaigns. As mentioned above, we simplified our deployment so that vaccination occurs in

cohorts that represent vaccination of in a given department target, and all vaccines for a campaign are deployed in a single week. These rollouts are evenly spaced in time over two or five year periods, as appropriate to the vaccination scenario, according to the schedule described in Figure 5 and Table 5.

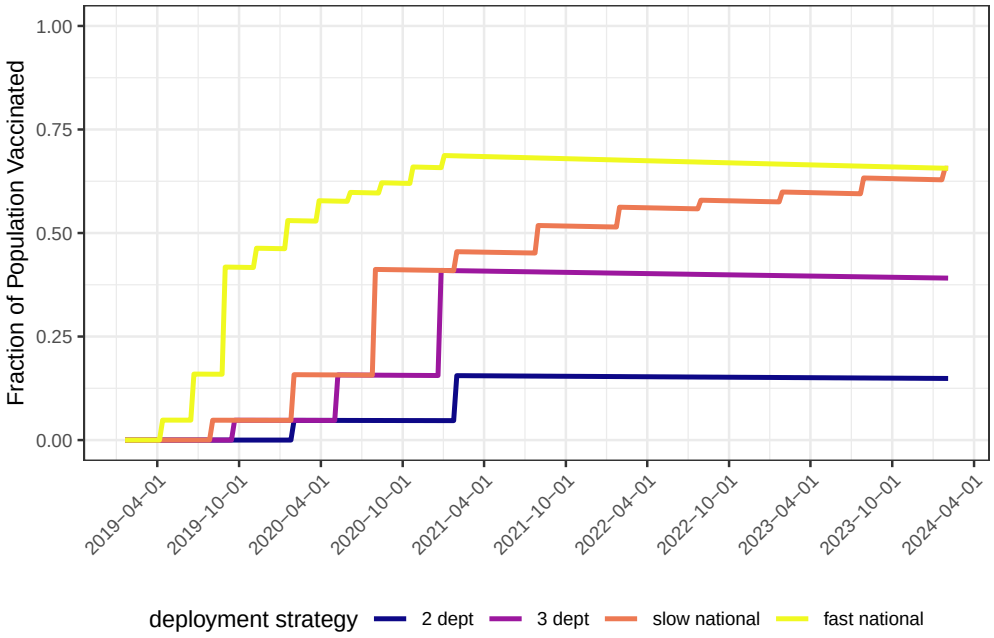

Figure 5: Influx of vaccines into the model projections, by vaccine deployment scenario. The fraction of the population vaccinated declines after the end of vaccination campaigns because the population is growing over time.

Table 5: Deployment schedule of vaccination campaigns in model projections.

Department	2 Dept	3 Dept	Slow National	Fast National
Centre	Feb 1, 2020	Sep 21, 2019	Aug 3, 2019	Apr 13, 2019
Artibonite	Jan 30, 2021	May 9, 2020	Jun 22, 2019	Feb 1, 2020
Ouest		Dec 26, 2020	Aug 31, 2019	Aug 1, 2020
Nord-Ouest			Nov 9, 2019	Jan 30, 2021
Nord			Jan 18, 2020	Jul 31, 2021
Sud			Mar 28, 2020	Jan 29, 2022
Nippes			Jun 6, 2020	Jul 30, 2022
Nord-Est			Aug 15, 2020	Jan 28, 2023
Sud-Est			Oct 24, 2020	Jul 29, 2023
Grand'Anse			Jan 2, 2021	Jan 27, 2024

5 Calculating Elimination Measures

We assessed the probability of elimination at three, five, and ten years after the start of the vaccination campaign. Elimination is defined as achieving a true incidence below one case for at least 52 consecutive weeks and remaining at this level through to the end of the 10 year projection period. The probability of elimination at a given time point is the proportion of simulations that achieved the start of an elimination state before the time point of interest (e.g., began the 52 consecutive weeks below one case before three years after vaccination start). True incidence includes all symptomatic and asymptomatic cases; this is different than the observed, symptomatic cases, which were the subset of data that had been calibrated to the Haiti surveillance data.

Among simulations that achieved elimination, we calculated the median time to elimination, which is the median number of weeks between the start of vaccination and elimination across simulations.

References

1. A. S. Azman, *et al.*, *PLOS Medicine* **12**, e1001867 (2015).
2. O. N. Bjørnstad, B. F. Finkenstadt, B. T. Grenfell, *Ecological Monographs* **72**, 169 (2002).
3. B. T. Grenfell, O. N. Bjornstad, B. F. Finkenstadt, *Ecological Monographs* **72**, 185 (2002).
4. J. A. Lewnard, *et al.*, *PLoS Medicine* **13**, 1 (2016).
5. Q. Bi, *et al.*, *Lancet Infect Dis* **17**, 1080 (2017).
6. United Nations DESA, Population Division, *World Population Prospects 2019, Haiti*.
7. A. A. King, D. Nguyen, E. L. Ionides, *Journal of Statistical Software* **69**, 1 (2016).
8. A. A. King, *et al.*, *pomp: Statistical Inference for Partially Observed Markov Processes* (2018). R package, version 1.19.
9. E. L. Ionides, D. Nguyen, Y. Atchadé, S. Stoev, A. A. King, *Proceedings of the National Academy of Sciences of the United States of America* **112**, 719 (2015).
10. Institute Haïtien de Statistique et d'Informatique, *Estimation de la population de 2015*.
11. Central Intelligence Agency, *The World Factbook 2016–17* (2016).
12. A. S. Azman, K. E. Rudolph, D. A. T. Cummings, J. Lessler, *Journal of Infection* **66**, 432 (2013).
13. A. A. Weil, *et al.*, *Clin Infect Dis* **49**, 1473 (2009).
14. M. M. Levine, *et al.*, *The Journal of Infectious Diseases* **143**, 818 (1981).
15. K. Koelle, M. Pascual, *The American Naturalist* **163**, 901 (2004).