Supplementary Material for An Ensemble Approach to Predicting the Impact of Vaccination on Rotavirus Disease in Niger

Jaewoo Park*[1], Joshua Goldstein[2], Murali Haran[1], and Matthew Ferrari[3]

[1]Department of Statistics, The Pennsylvania State University, University Park, PA 16802, USA

[2]Social and Data Analytics Laboratory, 900 N Glebe Rd, Virginia Tech, Arlington, VA 22203. USA

[3]Department of Biology, The Pennsylvania State University, University Park, PA 16802. USA

July 15, 2017

We provide details below about the five different dynamic models along with information about computational methods used to perform inference for each of them. In addition, we describe the implemention of the Bayesian model averaging approach used in the manuscript.

A Model Details

The structure of the models is given in Figure 1, which we explain in detail below. Common to each of the models we describe, we assume a time-varying transmission rate with a period of one year to account for seasonality,

$$\beta_i(t) = \beta_{0i}\left(1 + \omega\cos\left(\frac{2\pi t - 52\phi}{52}\right)\right),$$

where t is time in weeks, β_{0i} is the baseline rate for age class i, and ω and ϕ are the amplitude and offset of the seasonal variation.

We also assume the birth rate $\mu(t)$ varies with time. The mean weekly birthrate is estimated by $\bar{\mu} = 1/(5 \times 52)$. The variation in monthly birth rate is shown in Table 1. Finally, for each model we assume a negative binomial observation process with mean equal to the number of weekly reported cases and dispersion parameter ν.

We describe in detail the dynamics of each of the five models outlined in Figure 1. Model A [13, 2] is an SIRS model in which severe and mild rotavirus are tracked separately. Severe infections have a

*Email addresses: jzp191@psu.edu (J. Park), joshg22@vbi.vt.edu (J. Goldstein), muh10@psu.edu (M. Haran), mjf283@psu.edu (M. Ferrari)

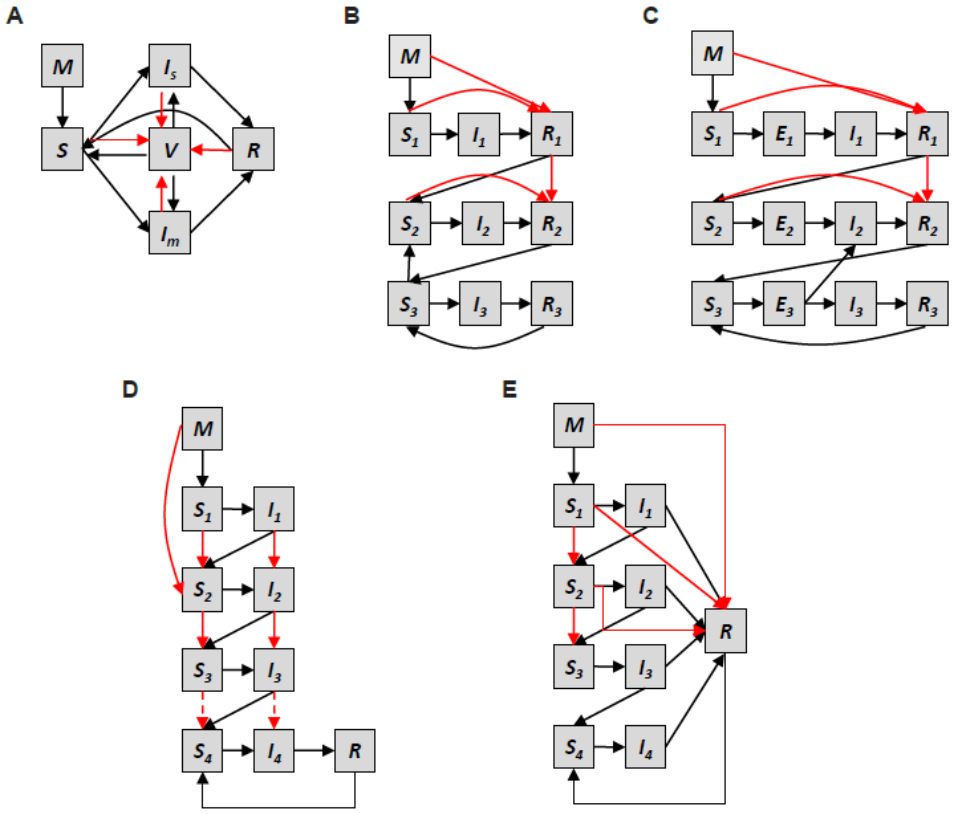

Figure 1: Structure of the compartmental models adapted from [10].

Table 1: Seasonal variation in birth rate in Niger, estimated from 1980-2000 using Demographic and Health Surveys. [5] An amplitude of $-.17$ for January tells us the birth rate is 17% below the mean.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Amplitude	-0.17	0.01	0.03	0.25	0.12	0.03	-0.01	0.09	0.01	0.13	-0.31	-0.17

longer duration and contribute more to the overall force of infection. Following infection, there is a period of temporary immunity that wanes over time. The model is age structured with age groups 0-1 month, 2-3 months, 4-5 months, 6-11 months, 1 year, and 2-5 years indexed by i. The differential equations describing the model dynamics are:

Model A

$$\frac{dM_i}{dt} = \alpha_{i-1}M_{i-1} - \alpha_i M_i + \mu N - \delta M_i \tag{1}$$

$$\frac{dS_i}{dt} = \alpha_{i-1}S_{i-1} - \alpha_i S_i + \delta M_i - \lambda_i S_i + \tau R_i \tag{2}$$

$$\frac{dI_i^{(s)}}{dt} = \alpha_{i-1}I_{i-1}^{(s)} - \alpha_i I_i^{(s)} + \lambda_i^{(s)} S_i - \gamma^{(s)} I_i^{(s)} \tag{3}$$

$$\frac{dI_i^{(m)}}{dt} = \alpha_{i-1}I_{i-1}^{(m)} - \alpha_i I_i^{(m)} + \lambda_i^{(m)} S_i - \gamma^{(m)} I_i^{(m)} \tag{4}$$

Movement between age classes occurs at rates dependent on the length of the interval in weeks, $\boldsymbol{\alpha} = \left\{\frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{24}, \frac{1}{48}, \frac{1}{144}\right\}$. The force of infection for age class i is given by $\lambda_i = \sum_{j=1}^{6} \beta_j(t) C_{ij} \frac{(I_j^{(s)} + 0.5 I_j^{(m)})}{N_j}$, assuming that relative infectiousness for mild infections is less than for severe RVGE. Here C_{ij} represents the frequency of contact from age class i onto class j [6], and satisfies $f_i C_{ij} = f_j C_{ji}$ where f_i is the fraction of the population in class i. We make the simplifying assumption that contact between age groups is homogeneous. With the absence of data on rotavirus infections for children over 5 and adults, we also assume the population of children under 5 is closed and consider child-child transmission only. Infection with rotavirus is typically asymptomatic [10] or unreported for older children and adults, but could potentially play a role in transmission. The contact matrix is

$$C = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 3 & 3 & 3 & 1 & 1 & 1 \\ 6 & 6 & 6 & 2 & 1 & 1 \\ 18 & 18 & 18 & 6 & 3 & 1 \end{pmatrix}$$

.

The differences in our age groups means that the contact matrix is not symmetric, for example we assume the population from 2-5 years is 18 times larger than the population from 0-1 months.

For all models, fixed parameters including infection period, immunity period, and exposed period in the SIR models are estimated from England and Wales data as described in [10]. After a period of maternal immunity (M_i), individuals can be susceptible (S_i), infected with either mild ($I_i^{(m)}$) or severe ($I_i^{(s)}$) rotavirus, or recovered (R_i). These represent the number of individuals in each class. In (1) we see how the number of children protected by maternal immunity change over time. Newborns are added to this class at rate μ and individuals leave this class when maternal immunity wanes with rate δ, where the mean period of maternal immunity is assumed to be 13 weeks ($\delta = \frac{1}{13}$).

When maternal immunity wanes children are susceptible to rotavirus infection. In (2), we see that individuals enter the susceptible class when maternal immunity wanes. They become infected at a rate given by the force of infection λ_i. After recovery, individuals may reenter the susceptible class at rate τ, where the mean period of immunity following infection is fixed at one year ($\tau = 1/52$).

Equation (3) models the change in total infections with severe rotavirus. We assume the proportion of infections with severe rotavirus is lower than mild by setting $\lambda_i^{(s)} = 0.24\lambda_i$. Individuals leave the infected with severe rotavirus for a mean period of one week ($\gamma^{(s)} = 1$) following which they are considered to be recovered. Similarly, (4) tracks the total infections with mild rotavirus, with $\lambda_i^{(m)} = 0.76\lambda_i$ and a mean infectious period of just half a week ($\gamma^{(m)} = 2$).

Only a fraction of infections with rotavirus develop RVGE (fixed at 24%), and we assume only severe cases are reported, so the expected number of reported cases for age class i is given by $\rho\lambda_i^{(s)}S_i$ where ρ is the reporting rate. We make the simplifying assumption for all models that ρ is constant across time and does not vary by age group.

Model B [11] is an SIRS model allowing for successive infections in which a second, third or subsequent infection will have a reduced susceptibility to infection and level of infectiousness. This represents partial immunity granted through repeated infections. Only a fraction of individuals in the a first or second infectious class are assumed to develop severe RVGE. The model dynamics are described by as follows.

Model B

$$
\begin{aligned}
\frac{dM_i}{dt} &= \alpha_{i-1}M_{i-1} - \alpha_i M_i + \mu N - \delta M_i \\
\frac{dS_i^{(1)}}{dt} &= \alpha_{i-1}S_{i-1}^{(1)} - \alpha_i S_i^{(1)} + \delta M_i - \lambda_i S_i^{(1)} \\
\frac{dI_i^{(1)}}{dt} &= \alpha_{i-1}I_{i-1}^{(1)} - \alpha_i I_i^{(1)} + \lambda_i S_i^{(1)} - \gamma^{(1)} I_i^{(1)} \\
\frac{dR_i^{(1)}}{dt} &= \alpha_{i-1}R_{i-1}^{(1)} - \alpha_i R_i^{(1)} + \gamma^{(1)} I_i^{(1)} - \tau R_i^{(1)} \\
\frac{dS_i^{(2)}}{dt} &= \alpha_{i-1}S_{i-1}^{(2)} - \alpha_i S_i^{(2)} + \tau R_i^{(1)} - \lambda_i^{(2)} S_i^{(2)} \\
\frac{dI_i^{(2)}}{dt} &= \alpha_{i-1}I_{i-1}^{(2)} - \alpha_i I_i^{(2)} + \lambda_i^{(2)} S_i^{(2)} - \gamma^{(2)} I_i^{(2)} \\
\frac{dR_i^{(2)}}{dt} &= \alpha_{i-1}R_{i-1}^{(2)} - \alpha_i R_i^{(2)} + \gamma^{(2)} I_i^{(2)} - \tau R_i^{(2)} \\
\frac{dS_i^{(3)}}{dt} &= \alpha_{i-1}S_{i-1}^{(3)} - \alpha_i S_i^{(3)} + \tau R_i^{(2)} + \tau R_i^{(3)} - \lambda_i^{(3)} S_i^{(3)} \\
\frac{dI_i^{(3)}}{dt} &= \alpha_{i-1}I_{i-1}^{(3)} - \alpha_i I_i^{(3)} + \lambda_i^{(3)} S_i^{(3)} - \gamma^{(2)} I_i^{(3)} \\
\frac{dR_i^{(3)}}{dt} &= \alpha_{i-1}R_{i-1}^{(3)} - \alpha_i R_i^{(3)} + \gamma^{(2)} I_i^{(3)} - \tau R_i^{(3)}
\end{aligned}
$$

Here in addition to an initial period of maternal immunity, individuals can be in the susceptible, infected, or recovered classes for their first $(S^{(1)}, I^{(1)}, R^{(1)})$, second $(S^{(2)}, I^{(2)}, R^{(2)})$, or third and subsequent $(S^{(3)}, I^{(3)}, R^{(3)})$ infections. The force of infection for age class i is given by $\lambda_i = \sum_{j=1}^{6} \beta_j(t) C_{ij} \frac{(I_j^{(1)} + 0.5 I_j^{(2)} + 0.2 I_j^{(3)})}{N_j}$, assuming that relative infectiousness decreases for subsequent infections. We assume the relative risk of infection decreases for subsequent infections, setting $\lambda_i^{(2)} = 0.62\lambda_i$ and $\lambda_i^{(3)} = 0.37\lambda_i$ as in [10]. Only 13% of first infections and 3% of second infections are assumed to develop severe RVGE, based on data from a Mexico cohort study [16]. So the expected number of reported cases for age class i is given by $\rho(0.13\lambda_i S_i^{(1)} + 0.03\lambda_i^{(2)} S_i^{(2)})$. Following

[16], we assume that the mean infectious period for the first infection is one week ($\gamma^{(1)} = 1$) and for subsequent infections is half a week ($\gamma^{(2)} = 2$).

Model C [4] is an SEIRS model, similar to Model B but allowing for an additional exposed or incubation period. Individuals in the exposed class are infected but not yet infectious. The dynamic equations are given by:

Model C

$$
\begin{aligned}
\frac{dM_i}{dt} &= \alpha_{i-1}M_{i-1} - \alpha_i M_i + \mu N - \delta M_i \\
\frac{dS_i^{(1)}}{dt} &= \alpha_{i-1}S_{i-1}^{(1)} - \alpha_i S_i^{(1)} + \delta M_i - \lambda_i S_i^{(1)} \\
\frac{dE_i^{(1)}}{dt} &= \alpha_{i-1}E_{i-1}^{(1)} - \alpha_i E_i^{(1)} + \lambda_i S_i^{(1)} - \xi E_i^{(1)} \\
\frac{dI_i^{(1)}}{dt} &= \alpha_{i-1}I_{i-1}^{(1)} - \alpha_i I_i^{(1)} + \xi E_i^{(1)} - \gamma^{(1)} I_i^{(1)} \\
\frac{dR_i^{(1)}}{dt} &= \alpha_{i-1}R_{i-1}^{(1)} - \alpha_i R_i^{(1)} + \gamma^{(1)} I_i^{(1)} - \tau R_i^{(1)} \\
\frac{dS_i^{(2)}}{dt} &= \alpha_{i-1}S_{i-1}^{(2)} - \alpha_i S_i^{(2)} + \tau R_i^{(1)} - \lambda_i^{(2)} S_i^{(2)} \\
\frac{dE_i^{(2)}}{dt} &= \alpha_{i-1}E_{i-1}^{(2)} - \alpha_i E_i^{(2)} + \lambda_i^{(2)} S_i^{(2)} - \xi E_i^{(2)} \\
\frac{dI_i^{(2)}}{dt} &= \alpha_{i-1}I_{i-1}^{(2)} - \alpha_i I_i^{(2)} + \xi E_i^{(2)} - \gamma^{(2)} I_i^{(2)} \\
\frac{dR_i^{(2)}}{dt} &= \alpha_{i-1}R_{i-1}^{(2)} - \alpha_i R_i^{(2)} + \gamma^{(2)} I_i^{(2)} - \tau R_i^{(2)} \\
\frac{dS_i^{(3)}}{dt} &= \alpha_{i-1}S_{i-1}^{(3)} - \alpha_i S_i^{(3)} + \tau R_i^{(2)} - \lambda_i^{(3)} S_i^{(3)} \\
\frac{dE_i^{(3)}}{dt} &= \alpha_{i-1}E_{i-1}^{(3)} - \alpha_i E_i^{(3)} + \lambda_i^{(3)} S_i^{(3)} - \xi E_i^{(3)} \\
\frac{dI_i^{(3)}}{dt} &= \alpha_{i-1}I_{i-1}^{(3)} - \alpha_i I_i^{(3)} + \xi E_i^{(3)} - \gamma^{(2)} I_i^{(3)} \\
\frac{dR_i^{(3)}}{dt} &= \alpha_{i-1}R_{i-1}^{(3)} - \alpha_i R_i^{(3)} + \gamma^{(2)} I_i^{(3)} - \tau R_i^{(3)}
\end{aligned}
$$

The modeling assumptions are the same as Model B but for the addition of an exposed class for the first, second, or subsequent infections ($E^{(1)}, E^{(2)}, E^{(3)}$). We assume a mean exposed period of 1 day ($\xi = 7$).

Model D [15] is an SIS model which also allows for successive infections with different levels of infectiousness, but assumes there is no period of temporary immunity following infection. After four infections individuals are assumed to be fully immune to infection. The dynamics are described as follows.

Model D

$$
\begin{aligned}
\frac{dM_i}{dt} &= \alpha_{i-1}M_{i-1} - \alpha_i M_i + \mu N - \delta M_i \\
\frac{dS_i^{(1)}}{dt} &= \alpha_{i-1}S_{i-1}^{(1)} - \alpha_i S_i^{(1)} + \delta M_i - \lambda_i S_i^{(1)} \\
\frac{dI_i^{(1)}}{dt} &= \alpha_{i-1}I_{i-1}^{(1)} - \alpha_i I_i^{(1)} + \lambda_i S_i^{(1)} - \gamma^{(1)} I_i^{(1)} \\
\frac{dS_i^{(2)}}{dt} &= \alpha_{i-1}S_{i-1}^{(2)} - \alpha_i S_i^{(2)} + \gamma^{(1)} I_i^{(1)} - \lambda_i^{(2)} S_i^{(2)} \\
\frac{dI_i^{(2)}}{dt} &= \alpha_{i-1}I_{i-1}^{(2)} - \alpha_i I_i^{(2)} + \lambda_i^{(2)} S_i^{(2)} - \gamma^{(2)} I_i^{(2)} \\
\frac{dS_i^{(3)}}{dt} &= \alpha_{i-1}S_{i-1}^{(3)} - \alpha_i S_i^{(3)} + \gamma^{(2)} I_i^{(2)} - \lambda_i^{(3)} S_i^{(3)} \\
\frac{dI_i^{(3)}}{dt} &= \alpha_{i-1}I_{i-1}^{(3)} - \alpha_i I_i^{(3)} + \lambda_i^{(3)} S_i^{(3)} - \gamma^{(2)} I_i^{(3)} \\
\frac{dS_i^{(4)}}{dt} &= \alpha_{i-1}S_{i-1}^{(4)} - \alpha_i S_i^{(4)} + \gamma^{(2)} I_i^{(3)} - \lambda_i^{(4)} S_i^{(4)} \\
\frac{dI_i^{(4)}}{dt} &= \alpha_{i-1}I_{i-1}^{(4)} - \alpha_i I_i^{(4)} + \lambda_i^{(4)} S_i^{(4)} - \gamma^{(2)} I_i^{(4)}
\end{aligned}
$$

The force of infection is $\lambda_i = \sum_{j=1}^{6} \frac{\beta_j(t) C_{ij} (I_j^{(1)} + 0.5 I_j^{(2)} + 0.2 I_j^{(3)} + 0.2 I_j^{(4)})}{N_j}$, assuming that relative infectiousness decreases for subsequent infections. We also assume the relative risk of infection decreases for subsequent infections, setting $\lambda_i^{(2)} = 0.62\lambda_i$ and $\lambda_i^{(3)} = \lambda_i^{(4)} = 0.37\lambda_i$. Again, we assume only 13% of first infections and 3% of second infections are assumed to develop severe RVGE. So the expected number of reported cases in age group i is given by $\rho(0.13\lambda_i S_i^{(1)} + 0.03\lambda_i^{(2)} S_i^{(2)})$.

Finally, Model E [1] is an SIR-SIS hybrid wherein following infection, individuals have a chance to either return to the susceptible class or gain full immunity. The equations for the dynamics are as follows.

Model E

$$\frac{dM_i}{dt} = \alpha_{i-1}M_{i-1} - \alpha_i M_i + \mu N - \delta M_i$$
$$\frac{dS_i^{(1)}}{dt} = \alpha_{i-1}S_{i-1}^{(1)} - \alpha_i S_i^{(1)} + \delta M_i - \lambda_i S_i^{(1)}$$
$$\frac{dI_i^{(1)}}{dt} = \alpha_{i-1}I_{i-1}^{(1)} - \alpha_i I_i^{(1)} + \lambda_i S_i^{(1)} - \gamma^{(1)} I_i^{(1)}$$
$$\frac{dS_i^{(2)}}{dt} = \alpha_{i-1}S_{i-1}^{(2)} - \alpha_i S_i^{(2)} + \kappa^{(1)}\gamma^{(1)} I_i^{(1)} - \lambda_i^{(2)} S_i^{(2)}$$
$$\frac{dI_i^{(2)}}{dt} = \alpha_{i-1}I_{i-1}^{(2)} - \alpha_i I_i^{(2)} + \lambda_i^{(2)} S_i^{(2)} - \gamma^{(2)} I_i^{(2)}$$
$$\frac{dS_i^{(3)}}{dt} = \alpha_{i-1}S_{i-1}^{(3)} - \alpha_i S_i^{(3)} + \kappa^{(2)}\gamma^{(2)} I_i^{(2)} - \lambda_i^{(3)} S_i^{(3)}$$
$$\frac{dI_i^{(3)}}{dt} = \alpha_{i-1}I_{i-1}^{(3)} - \alpha_i I_i^{(3)} + \lambda_i^{(3)} S_i^{(3)} - \gamma^{(2)} I_i^{(3)}$$
$$\frac{dS_i^{(4)}}{dt} = \alpha_{i-1}S_{i-1}^{(4)} - \alpha_i S_i^{(4)} + \kappa^{(3)}\gamma^{(2)} I_i^{(3)} - \lambda_i^{(4)} S_i^{(4)}$$
$$\frac{dI_i^{(4)}}{dt} = \alpha_{i-1}I_{i-1}^{(4)} - \alpha_i I_i^{(4)} + \lambda_i^{(4)} S_i^{(4)} - \gamma^{(2)} I_i^{(4)}$$

The chance of returning to the susceptible class varies by number of previous infections. Following [1] we fix $\kappa^{(1)} = 0.62$, $\kappa^{(2)} = 0.65$, $\kappa^{(3)} = 0.85$. The remaining modeling assumptions are the same as for models B-D.

A.1 Computational Details

Denote the observed data by $Y = \{Y_i(t); t \in (1, ..., t_{obs}), i \in (1, ..., 6)\}$ where $Y_i(t)$ is the number of reported cases in age group i during week t. Cases were observed over $t_{obs} = 118$ weeks. Denote the number of cases in age group i during week t predicted by our models by $\xi_i(t)$. For Model A,

$$\xi_i(t) = \rho\lambda_i^{(s)}(t)S_i(t)$$

While for models B-E,

$$\xi_i(t) = \rho(0.13\lambda_i(t)S_i^{(1)}(t) + 0.03\lambda_i^{(2)}(t)S_i^{(2)}(t))$$

For each model, the periodic solution to the system of ODEs specified above determines the number of reported cases in age group i during a given week. Although we assume model dynamics are periodic, the initial solutions in the numerical solver may not be periodic. Therefore, the numerical solver uses an iterative approach, integrating the model dynamics forward until a periodic solution is obtained. To carry this out, we use an iterative numerical solver [9] in the `deSolve` [14] package

in R. Solutions have a period of one year; that is, starting from arbitrary initial conditions, we run the dynamics forward until our expected number of cases is identical from one 52 week period to the next, to within a small tolerance; i.e.

$$\sum_{i=1}^{6} \sum_{t=t^*}^{t^*+52} |\xi_i(t) - \xi_i(t-52)| < \epsilon = 0.01$$

In practice, numerical integration for 20 years was enough to ensure the periodic solution was reached. After reaching a periodic solution, the models are integrated forward an additional 118 weeks to get the expected number of reported cases $(\Xi_i(t); t \in (1, ..., t_{obs}), i \in (1, ..., 6))$.

Define random variables $N_i(t) \sim NB(\Xi_i(t), \nu)$. The likelihood is

$$\mathcal{L}(Y|\Theta) = \prod_{i=1}^{6} \prod_{t=1}^{t_{obs}} f_{N_i(t)}(Y_i(t))$$

The number of observed reported cases is modeled as a Negative Binomial with mean equal to the expected number of cases and dispersion parameter ν.

Inference for our model parameters is done via Markov chain Monte Carlo (MCMC) for models A-E. At each step of the Markov chain, new parameters Θ' are proposed and the model dynamics are integrated forward until the periodic solution $\Xi_i(t; \Theta')$ is reached in order to calculate $\mathcal{L}(Y|\Theta')$. The parameters estimated by MCMC are $\Theta = (\omega, \phi, \nu, \rho, \beta_{0i}; i \in (1, ..., 6))$, including seasonal amplitude ω, seasonal phase ϕ, the dispersion ν of the Negative Binomial observation process, the reporting rate ρ, and the baseline transmission rate for age class β_{0i}.

MCMC samples are obtained from the posterior distribution

$$\pi(\omega, \phi, \nu, \rho, \beta_{01}, ..., \beta_{06}|Y) \propto \mathcal{L}(Y|\omega, \phi, \nu, \rho, \beta_{01}, ..., \beta_{06}) p(\omega) p(\phi) p(\nu) p(\rho) \prod_{i=1}^{6} p(\beta_{0i})$$

where we take priors $p(\beta_{0i}) = N(20, 5)$, $p(\omega) = \text{Unif}(0, 1)$, $p(\phi) = \text{Unif}(2, 2\pi + 2)$, $p(\nu) = \text{Gamma}(0.001, 0.001)$, and $p(\rho) = N(0.117, 0.06)$. The prior of our reporting rate ρ is centered at 11.7%, determined from the estimated reporting rate from the cluster survey (42.9%) and the estimated proportion of the population under 5 in the four districts that is covered by hospital surveillance (27.3%, from 2009 census data). In practice, we find that our estimates are robust to the choice of standard deviation of $p(\rho)$.

Table 2 provides parameter estimates from five different models. Estimates of the strength of transmission are similar for models B-D, higher for Model E and significantly lower for Model A. The same holds true for the reporting rate (Model A's estimate of the reporting rate is dramatically lower, and does not agree with estimates from the cluster survey, evidence that it is performing

Table 2: Posterior means and 95% HPD intervals for estimated parameters

Model	ω	ϕ	ν	ρ
A	0.50 (0.48,0.51)	7.4 (7.3,7.5)	1.5 (1.4,1.5)	0.039 (0.035,0.044)
B	0.38 (0.35,0.41)	7.4 (7.3,7.5)	2.7 (2.6,2.8)	0.108 (0.100,0.117)
C	0.39 (0.33,0.42)	7.4 (7.3,7.5)	2.7 (2.3,2.9)	0.109 (0.101,0.118)
D	0.31 (0.24,0.36)	7.1 (7.0,7.2)	2.6 (2.5,2.7)	0.109 (0.100,0.119)
E	0.33 (0.31,0.36)	7.3 (7.2,7.4)	5.4 (5.3,5.6)	0.119 (0.111,0.126)

poorly). Notably, the estimated phase of the transmission ϕ is similar across all models (Table 2). For reference, an estimated ϕ of 7.4 corresponds to a peak transmission in early March. This is quite close to the period of peak night time brightness in Maradi as measured by satellite imagery [3]. The peak night time brightness has been shown to be related to fluctuation of measles cases. Temporal change of urban population density and measles transmission are highly correlated, and population density can be measured by night time brightness [3].

A.2 Dynamics Accounting for Vaccination

Based on the results of [8] we assume that 63% of vaccinated individuals are successfully seroconvert after a single dose. Our models with vaccination allow for the red transitions in Figure 1. For example, Model B allows for transitions directly from $M_{i=1}$ and $S^{(1)}_{i=1}$ to $R^{(1)}_{i=2}$ on the first dose, and from $R^{(1)}_{i=2}$ and $S^{(2)}_{i=2}$ to $R^{(2)}_{i=3}$ on the second dose. The dynamics equations will be modified by the following terms:

$$
\begin{aligned}
\frac{dM_{i=2}}{dt} &= (1-\sigma\psi)\alpha_1 M_{i=1} + ... \\
\frac{dS^{(1)}_{i=2}}{dt} &= (1-\sigma\psi)\alpha_1 S^{(1)}_{i=1} + ... \\
\frac{dR^{(1)}_{i=2}}{dt} &= (\sigma\psi)\alpha_1 M_{i=1} + (\sigma\psi)\alpha_1 S^{(1)}_{i=1} + ... \\
\frac{dR^{(1)}_{i=3}}{dt} &= (1-\sigma\psi)\alpha_2 R^{(1)}_{i=2} + ... \\
\frac{dS^{(2)}_{i=3}}{dt} &= (1-\sigma\psi)\alpha_2 S^{(2)}_{i=2} + ... \\
\frac{dR^{(2)}_{i=3}}{dt} &= (\sigma\psi)\alpha_2 R^{(1)}_{i=2} + (\sigma\psi)\alpha_2 S^{(2)}_{i=2} + ...
\end{aligned}
$$

Where ψ is the coverage and $\sigma = 0.63$ is the rate of seroconversion [8] for low socio-economic settings. This means that an individual who is vaccinated with a single dose has a lower risk of infection, comparable to the effect of recovering from a natural infection. Vaccination with a second dose further reduces risk of infection.

In Model A, the risk of infection does not decline with the previous number of infections. Therefore,

an additional vaccinated state V_i is added to the model for age group i. Two additional input parameters are required for the vaccine efficacy against severe and mild RVGE. We assume the vaccination happens at 2 months, but the vaccine efficacy is equal to the efficacy predicted under models B-E for the two dose strategy, $\eta^{(s)} = .796$ and $\eta^{(m)} = .609$.

$$
\begin{aligned}
\frac{dM_{i=2}}{dt} &= (1-\psi)\alpha_1 M_{i=1} + ... \\
\frac{dS_{i=2}}{dt} &= (1-\psi)\alpha_1 S_{i=1} + ... \\
\frac{dV_{i=2}}{dt} &= (\psi)\alpha_1 M_{i=1} + (\psi)\alpha_1 S_{i=1} + ... \\
\frac{dV_{i>2}}{dt} &= \alpha_{i-1}V_{i-1} - \alpha_i V_i - (\tau + \lambda_i^{(s)}(1-\eta^{(s)}) + \lambda_i^{(m)}(1-\eta^{(m)}))V_i \\
\frac{dS_{i>2}}{dt} &= \tau V_i + ... \\
\frac{dI_{i>2}^{(s)}}{dt} &= \lambda_i^{(s)}(1-\eta^{(s)})V_i + ... \\
\frac{dI_{i>2}^{(m)}}{dt} &= \lambda_i^{(m)}(1-\eta^{(m)})V_i + ...
\end{aligned}
$$

Given our vaccination strategy for models B-E, the vaccine efficacy for severe RVGE after two doses is 79.6%, in line with efficacy studies of rotavirus vaccines. This is calculated by multiplying the proportion of individuals who are successfully immunized twice, once, or zero times by the expected reduction in RVGE incidence for each case.

$$VE = 1 - \left[0.37^2 \times 1 + 2(0.37)(0.63) \times \left(0.62\frac{0.03}{0.13}\right) + 0.63^2 \times \left(0.37\frac{0}{0.13}\right)\right] = 79.6\%. \tag{5}$$

We assume following [16] that 47% of first infections and 25% of second infections and 32% of third infections are assumed to develop any RVGE (mild RVGE is unreported). Therefore the vaccine efficacy for all RVGE is

$$VE = 1 - \left[0.37^2 \times 1 + 2(0.37)(0.63) \times \left(0.62\frac{0.25}{0.47}\right) + 0.63^2 \times \left(0.37\frac{0.32}{0.47}\right)\right] = 60.9\%.$$

In practice, first model parameters Θ are estimated via MCMC for the models without vaccination. Using the fitted model, the dynamics are then integrated forward at the posterior mean of Θ until the periodic solution has been reached. Then, the dynamics are modified to allow for transitions between compartments by vaccination.

A.2.1 Calculating the Direct Effect of Vaccination

When the vaccine is introduced in the population it leads to decreased transmission, which in turn leads to a reduced force of infection. The direct effect (DE) of vaccination is the expected reduction in cases for vaccinated individuals that is not due to the reduction in force of infection. On the other hand, the indirect effect (IE) of vaccination is the expected reduction in cases for both vaccinated and unvaccinated individuals due to the reduction in force of infection. The total effect (TE) of vaccination includes both DE and IE.

Define S^* to be the updated number of susceptibles after the vaccine has been introduced. Our models assume successive infections except for model A which has 0 weight. If the vaccination has been introduced for a long time (long enough to include all age classes) then any $S^{(1)}$ in the age class would be one who was vaccinated but failed to seroconvert. Therefore, $S^{*(1)}_{i>1} = (1-\sigma\psi)S^{(1)}_{i>1}$, and a reduced amount of $(\sigma\psi)S^{(1)}_{i>1}$ is moved to $S^{(2)}_{i>1}$. This same logic would follow for those leaving $S^{(2)}$ and entering $S^{(3)}$ due to vaccination. Therefore, $S^{*(2)}_{i>1} = (\sigma\psi)S^{(1)}_{i>1} + (1-\sigma\psi)S^{(2)}_{i>1}$.

We estimate DE by using $S^{(1)}$, $S^{(2)}$, and λ from dynamic equations without vaccination. Then the reduced burden is calculated $\rho\lambda_i(0.03S^{*(1)}_i + (0.63)(0.62)S^{*(2)}_i)$. This burden estimate considers movement between susceptibles due to vaccination in the absence of any resulting reduction in force of infections. The reduced burden is estimated for each model and the BMA estimate is evaluated according to the model weights.

A.2.2 Projections Based on Vaccine Efficacy from a Recent Study

Recently [7] estimated that 3 doses of vaccine had 66.7% efficacy against severe RVGE among children in Niger. Though we do not explicitly account for 3 doses of vaccine, we can calculate the effective seroconversion rate for our model above that would yield this observed efficacy after a complete sequence of doses. Thus, we set $\eta^{(s)} = .667$ and use (5) to calculate the effective seroconversion rate as 49%. Then the vaccine efficacy for all RVGE is $\eta^{(m)} = .515$. We estimate the predicted impact of vaccination using different $\eta^{(s)}, \eta^{(m)}$ and σ values with the same dynamic equations.

Although we used the two dose strategy, by using different value of the efficacy, our study can account for uncertainty in the seroconversion rate. Figures 2-4 are matched to Figures 3-5 in the main paper. Because of the lower seroconversion rate, the projected results were qualitatively similar, quantitatively smaller. Vaccination causes a shift in the age distribution across models (Figure 2), with a higher proportion of RVGE cases occurring for older children.

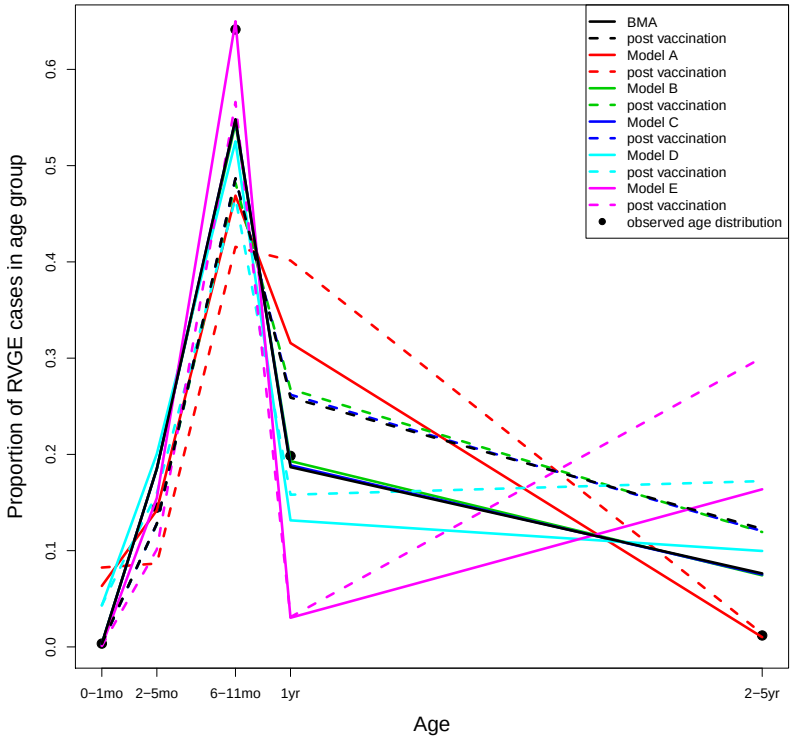

Figure 2: Distribution of cases across age groups observed in the data (black dots), predicted by the models (solid lines), and predicted 20 years after vaccination has been introduced at 70% coverage (dashed lines).

Over the short term, Models A-E predict an overall decline in total burden, but an increase in the magnitude of peak incidence (Figure 3).

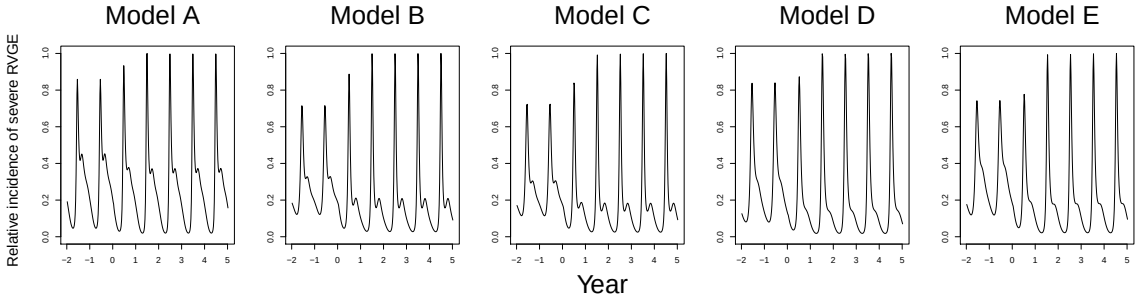

Figure 3: Relative incidence of severe RVGE after vaccination has been introduced into the models assuming 70% coverge, out to five years after vaccination has been introduced. The vaccination has been introduced at 0 year.

Figure 4 indicates that the short term trend of vaccination impacts based on BMA is similar to that of Model C. BMA predicts 31.1% (indrect effect: 1.0%) of long term reduction (99%CI : $(29.4\%, 32.1\%)$).

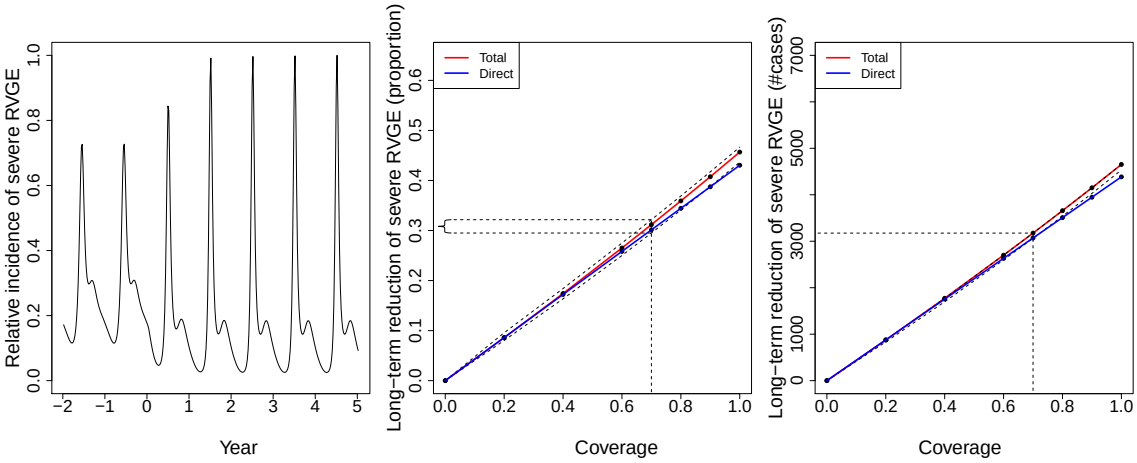

Figure 4: Relative incidence of severe RVGE (Left), percent (Middle) and absolute (Right) long term reduction in cases by coverage for Bayesian model averaging from the five fitted models. Dashed lines denote 99% confidence interval for the total effect. The vaccination has been introduced at 0 year. Variation in reduction for a fixed (70%) level of coverage is demonstrated.

B Bayesian Model Averaging

For $k = 1, ..., 5$, consider M_k, the kth model, with prior $p(\Theta_k|M_k)$ and likelihood function $\mathcal{L}(Y|\Theta_k, M_k)$. Note that we take the uniform model prior for $p(M_l)$ and model evidence $P(Y|M_k)$ is approximated via Bayesian information criterion (BIC) as in [12]. Then the posterior model probability (PMP) for M_k given the observed data C is

$$p(M_k|Y) = \frac{p(Y|M_k)p(M_k)}{\sum_{l=1}^{5} p(Y|M_l)p(M_l)},$$

where

$$p(Y|M_k) = \int \mathcal{L}(Y|\Theta_k, M_k)p(\Theta_k|M_k)d\Theta_k$$

is the model evidence for M_k which measures how well each model is supported by the observed data. Then the BMA estimate of the burden is

$$E[\xi(t)|Y] = \sum_{l=1}^{5} E[\xi_l(t)|Y, M_l]p(M_l|Y).$$

A summary of our implementation of BMA is as follows: (1) We construct a separate MCMC algorithm for each of the models A-E. (2) For each model, the burden estimate $\xi_k(t)$ is evaluated for the MCMC samples of the posterior distribution of that model. (3) The expected burden for model k, $E[\xi_k(t)|Y, M_k]$, is estimated through the sample mean of the $\xi_k(t)$s obtained from Step (2). (4) We take the weighted average of the burden across all models, with the weights equal to the posterior model probabilities, $p(M_k|Y)$, obtained above.

References

[1] Atchison, C., Lopman, B., and Edmunds, W. J. (2010). Modelling the seasonality of rotavirus disease and the impact of vaccination in England and Wales. *Vaccine*, 28(18):3118–3126.

[2] Atkins, K. E., Shim, E., Pitzer, V. E., and Galvani, A. P. (2012). Impact of rotavirus vaccination on epidemiological dynamics in England and Wales. *Vaccine*, 30(3):552–564.

[3] Bharti, N., Tatem, A. J., Ferrari, M. J., Grais, R. F., Djibo, A., and Grenfell, B. T. (2011). Explaining seasonal fluctuations of measles in Niger using nighttime lights imagery. *Science*, 334(6061):1424–1427.

[4] de Blasio, B. F., Kasymbekova, K., and Flem, E. (2010). Dynamic model of rotavirus transmission and the impact of rotavirus vaccination in Kyrgyzstan. *Vaccine*, 28(50):7923–7932.

[5] Dorélien, A. M. (2013). A time to be born: birth seasonality in sub-Saharan Africa. *University of Michigan Population Studies Center Research Report Series Report*, 13(785):1–61.

[6] Hethcote, H. (1996). *Modeling heterogeneous mixing in infectious disease dynamics*. Cambridge: Cambridge University Press.

[7] Isanaka, S., Guindo, O., Langendorf, C., Matar Seck, A., Plikaytis, B. D., Sayinzoga-Makombe, N., McNeal, M. M., Meyer, N., Adehossi, E., Djibo, A., et al. (2017). Efficacy of a low-cost, heat-stable oral rotavirus vaccine in Niger. *New England Journal of Medicine*, 376(12):1121–1130.

[8] Lopman, B. A., Pitzer, V. E., Sarkar, R., Gladstone, B., Patel, M., Glasser, J., Gambhir, M., Atchison, C., Grenfell, B. T., Edmunds, W. J., et al. (2012). Understanding reduced rotavirus vaccine efficacy in low socio-economic settings. *PloS one*, 7(8):e41720.

[9] Petzold, L. (1983). Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations. *SIAM journal on scientific and statistical computing*, 4(1):136–148.

[10] Pitzer, V. E., Atkins, K. E., de Blasio, B. F., Van Effelterre, T., Atchison, C. J., Harris, J. P., Shim, E., Galvani, A. P., Edmunds, W. J., Viboud, C., et al. (2012). Direct and indirect effects of rotavirus vaccination: comparing predictions from transmission dynamic models. *PloS one*, 7(8):e42320.

[11] Pitzer, V. E., Viboud, C., Simonsen, L., Steiner, C., Panozzo, C. A., Alonso, W. J., Miller, M. A., Glass, R. I., Glasser, J. W., Parashar, U. D., et al. (2009). Demographic variability, vaccination, and the spatiotemporal dynamics of rotavirus epidemics. *Science*, 325(5938):290–294.

[12] Raftery, A. E. (1996). Approximate Bayes factors and accounting for model uncertainty in generalised linear models. *Biometrika*, 83(2):251–266.

[13] Shim, E., Feng, Z., Martcheva, M., and Castillo-Chavez, C. (2006). An age-structured epidemic model of rotavirus with vaccination. *Journal of mathematical biology*, 53(4):719–746.

[14] Soetaert, K., Petzoldt, T., and Setzer, R. W. (2010). Solving differential equations in R: Package desolve. *Journal of Statistical Software*, 33(9):1–25.

[15] Van Effelterre, T., Soriano-Gabarro, M., Debrus, S., Claire Newbern, E., and Gray, J. (2010). A mathematical model of the indirect effects of rotavirus vaccination. *Epidemiology and infection*, 138(06):884–897.

[16] Velázquez, F. R., Matson, D. O., Calva, J. J., Guerrero, M. L., Morrow, A. L., Carter-Campbell, S., Glass, R. I., Estes, M. K., Pickering, L. K., and Ruiz-Palacios, G. M. (1996). Rotavirus infection in infants as protection against subsequent infections. *New England Journal of Medicine*, 335(14):1022–1028.