Age-structured pertussis dynamics in Thailand
Supplementary Information

JC Blackwood, DAT Cummings, S Iamsirithaworn, & P Rohani

Contents:

Mean age of infection

To show that the distribution of age-stratified case reports is structurally different between vaccine eras (i.e. 1981–1983 and 1996–2000), we estimate the mean age of infection in each era using bootstrap samples of the data. Given that the range within each age class ranges from 1 year to 10 years or more, we first determined whether it is reasonable to assume that reported cases within each interval are uniformly distributed. Therefore, we used the reported data by simulating 10,000 datasets of that actual ages of each case report under the assumption that reports are uniformly distributed within each age class. We computed the mean age of infection for each sample dataset and the results are displayed in Figure A1. Given that there is relatively small variation in the mean age of infection under this assumption, we conclude that taking bootstrap samples of one simulated dataset is representative of all such sample sets. Therefore, to estimate the mean age of infection, we select a sample data set and look at the distribution of 10,000 bootstrap samples (main text, Figure 4B).

Incidence in infants across provinces

Figure 4A in the main test demonstrates a clear decline in cases among infants between vaccine era. This figure is based on case reports at the national level. As a supplement, Figure **??** compares the distribution of cases in infants between each vaccine era across all provinces. The difference between the two is substantial and the median value of cases per 100,000 in the later vaccine era is 0 (compared to 35.42 in the early vaccine era).

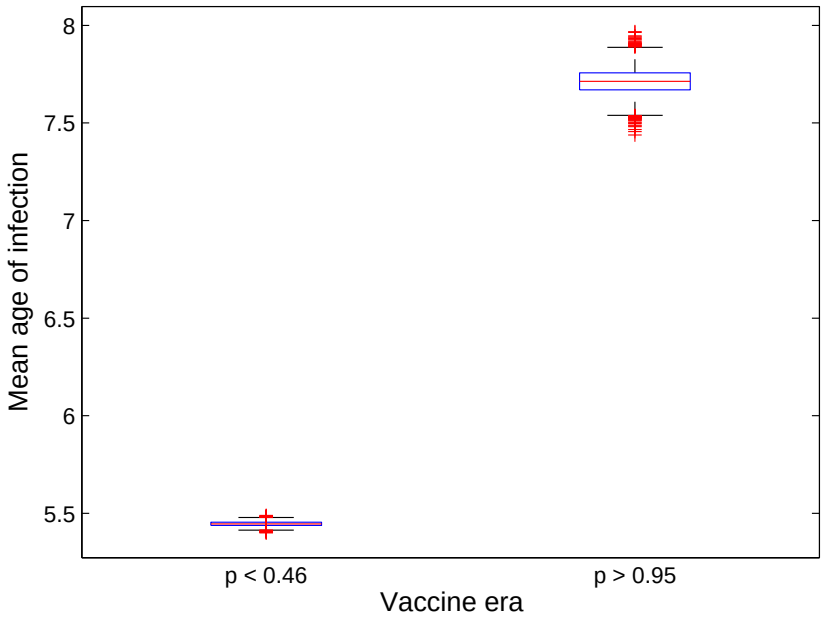

Figure A1. Box plots of the mean age of infection in each vaccine era for 10,000 simulated datasets. Each dataset assumes that the reported cases are distributed uniformly across the associated age group.

Age-structured models

To determine whether vaccination decreases pertussis transmission or only prevents symptomatic disease, we analyze two-age structured models and compare the results to the data. The first model is a simple age-structured SIR model that assumes lifelong immunity. In contrast, the second model assumes that there is a vaccinated class that can become infected without being observed and transmit pertussis, which we call the age-structured SIV^IR model. We divide each model into 5 age classes to mimic the data: $<$1 yr, 1-4 yr, 5-9 yr, 10-14 yr, and $>$15 years old. We assume the high risk group is aged $<$1 year and once individuals leave this class they have a probability p of becoming vaccinated.

We first describe the SIV^IR model in detail because it collapses to the SIR under certain conditions described below. For all age classes, the force of infection from individuals in the infected (I) class is

$$\lambda_i = \sum_{j=1}^{5} \beta_{i,j} I_j \tag{A1}$$

and the force of infection from individuals in the infected class that can transmit without being

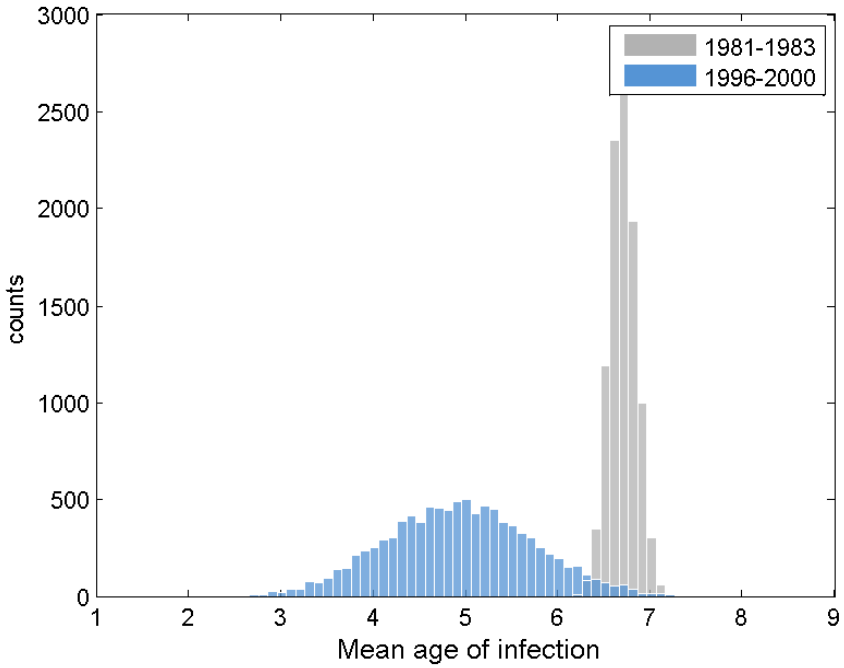

Figure A2. Region 1 (Northern): Distribution of 10,000 bootstrap samples of the mean age of infection ($p < 46\%$ in gray, $p > 95\%$ in color, where the color corresponds to the region as in Figure 2A in the main text. From 1981–1983, the sample mean was found to be 6.6 yrs, with a 95% confidence interval of $(6.33, 6.91)$. From 1996–2000, the sample mean was found to be 4.85 yrs, with a 95% confidence interval of $(3.19, 6.49)$.

observed (I^V) is given by

$$\lambda_i^V = \sum_{j=2}^{5} \beta_{i,j}^V I_j^V \tag{A2}$$

where $\beta_{i,j}$ is the transmission rate from individuals in the j age class to members of the i age class.

The SIV^IR model for the first age class (< 1 yr) includes births and for simplicity we assume that the birth and death rates are equal:

$$\frac{dS_1}{dt} = \mu - (\lambda + \lambda^V)S_1 - (\alpha_1 + \mu)S_1 \tag{A3}$$

$$\frac{dI_1}{dt} = (\lambda + \lambda^V)S_1 - (\gamma + \alpha_1 + \mu)I_1 \tag{A4}$$

$$\frac{dR_1}{dt} = \gamma I_1 - (\alpha_1 + \mu)R_1 \tag{A5}$$

where α_i is the aging rate for each age class i, μ is the birth (death) rate, and $1/\gamma$ is the infectious period.The second age group is 6 mo-4 year olds, and upon entering this class some proportion p become vaccinated:

$$\frac{dS_2}{dt} = (1-p)\alpha_1 S_1 - (\lambda + \lambda^V)S_2 - (\alpha_2 + \mu)S_2 \tag{A6}$$

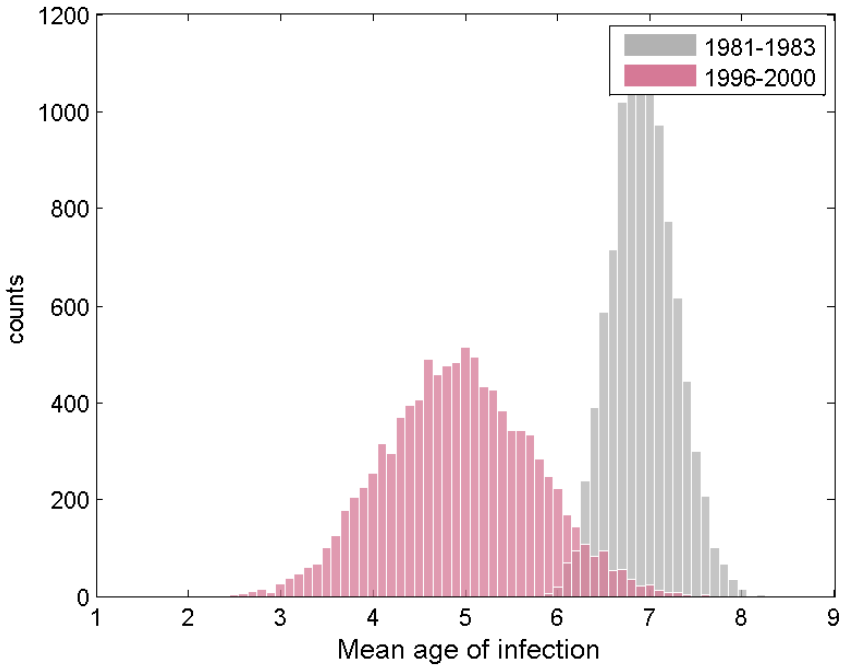

Figure A3. Region 2 (Western): Distribution of 10,000 bootstrap samples of the mean age of infection ($p < 46\%$ in gray, $p > 95\%$ in color, where the color corresponds to the region as in Figure 2A in the main text. From 1981–1983, the sample mean was found to be 6.95 yrs, with a 95% confidence interval of $(6.32, 7.74)$. From 1996–2000, the sample mean was found to be 4.85 yrs, with a 95% confidence interval of $(3.23, 6.48)$.

$$\frac{dV_2}{dt} = p\alpha_1 S_1 - (\rho + \alpha_2 + \mu)V_2 \tag{A7}$$

$$\frac{dS_2^V}{dt} = \rho V_2 - (\lambda + \lambda^V)S_2^V - (\alpha_2 + \mu)S_2^V \tag{A8}$$

$$\frac{dI_2^V}{dt} = (\lambda + \lambda^V)S_2^V - (\gamma + \alpha_2 + \mu)I_2^V \tag{A9}$$

$$\frac{dI_2}{dt} = \alpha_1 I_1 + (\lambda + \lambda^V)S_2 - (\gamma + \alpha_2 + \mu)I_2 \tag{A10}$$

$$\frac{dR_2}{dt} = \alpha_1 R_1 + \gamma(I_2^V + I_2) - (\alpha_2 + \mu)R_2 \tag{A11}$$

where p is the proportion of individuals vaccinated and ρ is the rate at which individuals leave the vaccinated class (V) and enter the susceptible vaccinated class (S^V) that can acquire subclinical infections (I^V). The third and fourth classes will then look like:

$$\frac{dS_i}{dt} = \alpha_{i-1} S_{i-1} - (\lambda + \lambda^V)S_i - (\alpha_i + \mu)S_i \tag{A12}$$

$$\frac{dV_i}{dt} = \alpha_{i-1} V_{i-1} - (\rho + \alpha_i + \mu)V_i \tag{A13}$$

$$\frac{dS_i^V}{dt} = \alpha_{i-1} S_{i-1}^V + \rho V_i - (\lambda + \lambda^V)S_i^V - (\alpha_i + \mu)S_i^V \tag{A14}$$

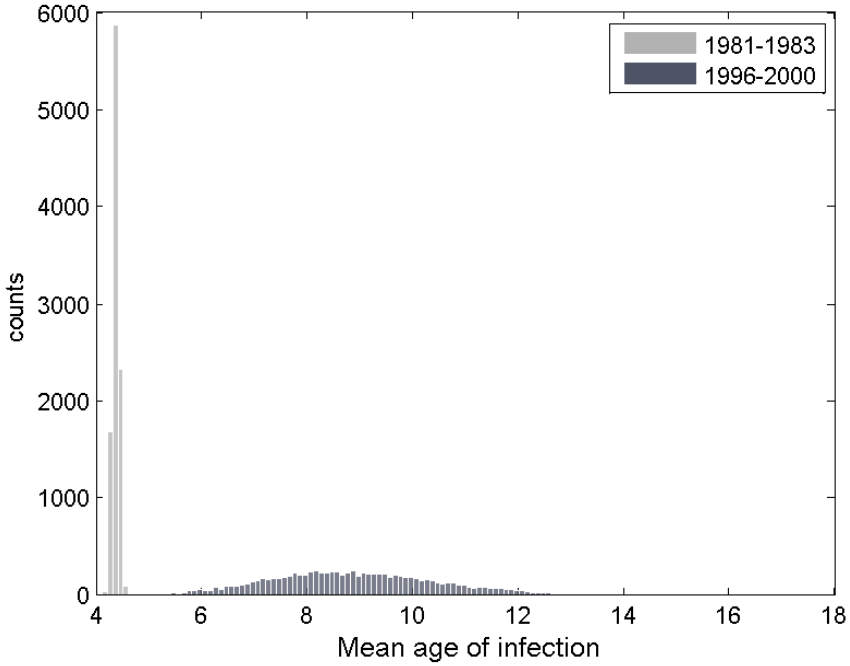

Figure A4. Region 3 (Central): Distribution of 10,000 bootstrap samples of the mean age of infection ($p < 46\%$ in gray, $p > 95\%$ in color, where the color corresponds to the region as in Figure 2A in the main text. From 1981–1983, the sample mean was found to be 4.37 yrs, with a 95% confidence interval of $(4.25, 4.49)$. From 1996–2000, the sample mean was found to be 8.68 yrs, with a 95% confidence interval of $(6.16, 12.7)$.

$$\frac{dI_i^V}{dt} = \alpha_{i-1} I_{i-1}^V + (\lambda + \lambda^V) S_i^V - (\gamma + \alpha_i + \mu) I_i^V \quad \text{(A15)}$$

$$\frac{dI_i}{dt} = \alpha_{i-1} I_{i-1} + (\lambda + \lambda^V) S_i - (\gamma + \alpha_i + \mu) I_i \quad \text{(A16)}$$

$$\frac{dR_i}{dt} = \alpha_{i-1} R_{i-1} + \gamma(I_i^V + I_i) - (\alpha_i + \mu) R_i \quad \text{(A17)}$$

and the final class will be:

$$\frac{dS_5}{dt} = \alpha_4 S_4 - (\lambda + \lambda^V) S_5 - \mu S_5 \quad \text{(A18)}$$

$$\frac{dV_5}{dt} = \alpha_4 V_4 - (\rho + \mu) V_5 \quad \text{(A19)}$$

$$\frac{dS_5^V}{dt} = \alpha_4 S_4^V + \rho V_5 - (\lambda + \lambda^V) S_5^V - \mu S_5^V \quad \text{(A20)}$$

$$\frac{dI_5^V}{dt} = \alpha_4 I_4^V + (\lambda + \lambda^V) S_5^V - (\gamma + \mu) I_5^V \quad \text{(A21)}$$

$$\frac{dI_5}{dt} = \alpha_4 I_4 + (\lambda + \lambda^V) S_5 - (\gamma + \mu) I_5 \quad \text{(A22)}$$

$$\frac{dR_5}{dt} = \alpha_4 R_4 + \gamma(I_5^V + I_5) - \mu R_5 \quad \text{(A23)}$$

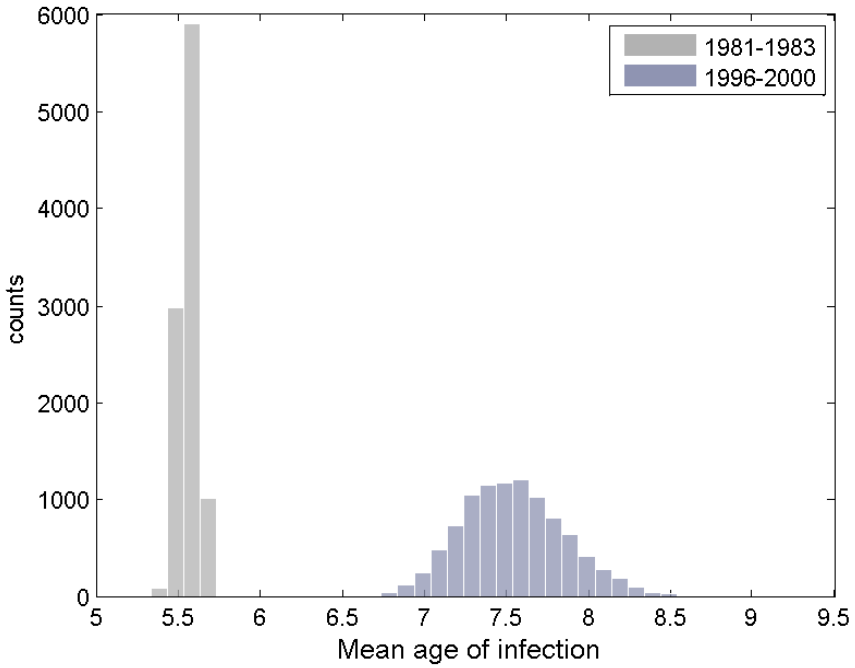

Figure A5. Region 4 (Northeastern): Distribution of 10,000 bootstrap samples of the mean age of infection ($p < 46\%$ in gray, $p > 95\%$ in color, where the color corresponds to the region as in Figure 2A in the main text. From 1981–1983, the sample mean was found to be 5.57 yrs, with a 95% confidence interval of $(5.46, 5.68)$. From 1996–2000, the sample mean was found to be 7.64 yrs, with a 95% confidence interval of $(7.14, 8.44)$.

Parameters chosen correspond to a basic reproductive number, R_0, of 16 [1]. Here, R_0 is defined as the average number of secondary infections arising from the introduction of an infective (I) individual in an otherwise entirely susceptible (S) population. All parameters for this model are listed in Table A-1. We first assume that mixing is homogeneous (Figure 4C in the main text), and then assume there is non-homogeneous mixing according to contact patterns reported in [3]. The analysis in [3] reported contact patterns for several European countries and for simplicity, we took the average of these matrices (see Figure A9). Further, to account for reciprocity in reporting we use the methods of [4]. Further, the matrix is appropriately scaled so that R_0 is consistent across both models. The results using each assumption on contact structure is qualitatively similar, so the main text displays the results for homogeneous mixing, and Figure A10 displays the results that use the data in [3].

We compare model to the age-structured SIR model which is identical to the above model, but without λ^V and the S^V, V, and I^V classes. Instead, vaccinated individuals move directly into the recovered R class. To compare the models to the date, we plot the equilibrium prevalence of infected individuals in each age class (i.e. I_i^*/n_i where n_i is the proportion of the population in age class i) when vaccination is at 46% and 95% (Figure 4 in main text). To find equilibrium prevalence, we simulate each model for 100 years and average I_i^*/n_i over the final year of the simulation.

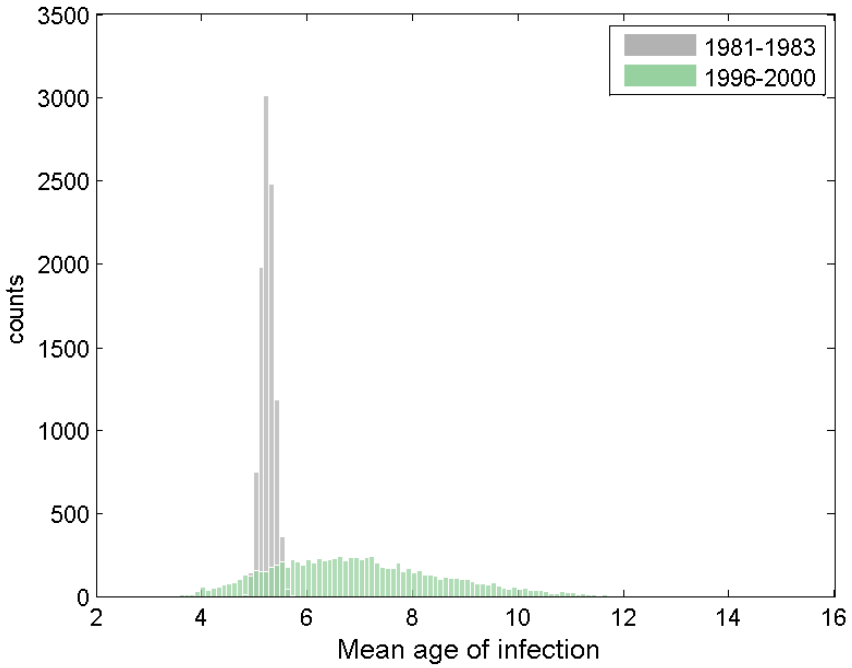

Figure A6. Region 5 (Eastern): Distribution of 10,000 bootstrap samples of the mean age of infection ($p < 46\%$ in gray, $p > 95\%$ in color, where the color corresponds to the region as in Figure 2A in the main text. From 1981–1983, the sample mean was found to be 5.28 yrs, with a 95% confidence interval of $(5.03, 5.55)$. From 1996–2000, the sample mean was found to be 7.21 yrs, with a 95% confidence interval of $(4.93, 12.72)$.

Table A-1. SIR and SIV^IR model parameters

PARAMETER	DESCRIPTION	VALUE	Source
μ	Per capita death (birth) rate	1/(365*68.5) day^{-1}	[6]
β,β_V	Transmission rate, corresponds to $R_0 = 16$	16/21 day^{-1}	[1, 5]
$1/\gamma$	Infectious period	21 days	[2]
$1/\rho$	Duration of vaccine-derived immunity	Variable	*NA*

References

[1] RM Anderson and RM May. *Infectious diseases of humans*. Oxford University Press, New York, 1991.

[2] CJE Metcalf, ON Bjørnstad, BT Grenfell, and V Andreasen. Seasonality and comparative dynamics of six childhood infections in pre-vaccination Copenhagen. *Proc R Soc B*, 276:4111–4118, 2009.

[3] Jol Mossong, Niel Hens, Mark Jit, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco

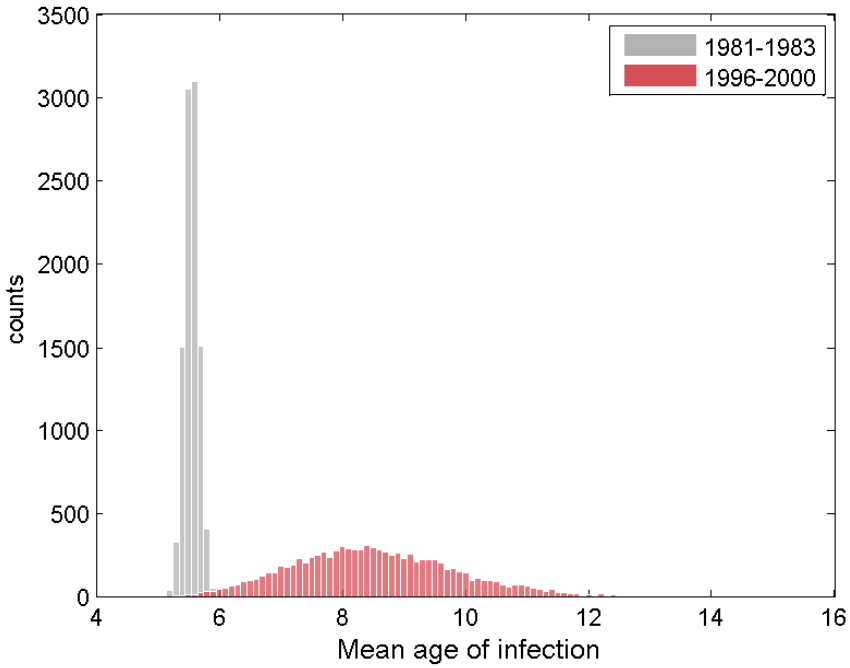

Figure A7. Region 6 (Southern): Distribution of 10,000 bootstrap samples of the mean age of infection ($p < 46\%$ in gray, $p > 95\%$ in color, where the color corresponds to the region as in Figure 2A in the main text. From 1981–1983, the sample mean was found to be 5.47 yrs, with a 95% confidence interval of $(5.25, 5.69)$. From 1996–2000, the sample mean was found to be 8.34 yrs, with a 95% confidence interval of $(6.34, 11.8)$.

Massari, Stefania Salmaso, Gianpaolo Scalia Tomba, Jacco Wallinga, Janneke Heijne, Malgorzata Sadkowska-Todys, Magdalena Rosinska, and W. John Edmunds. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med*, 5(3):e74, 03 2008.

[4] M.A. Riolo, A.A. King, and P. Rohani. Can vaccine legacy explain the british pertussis resurgence? *Vaccine*, 32:5903–5908, 2013.

[5] P. Rohani, D.J.D. Earn, and B.T. Grenfell. Opposite patterns of synchrony in sympatric disease metapopulations. *Science*, 286:968–971, 1999.

[6] The World Bank. Life expectancy at birth, total (years). *Accessed via: http://data.worldbank.org/indicator/SP.DYN.LE00.IN/countries/*, 2011.

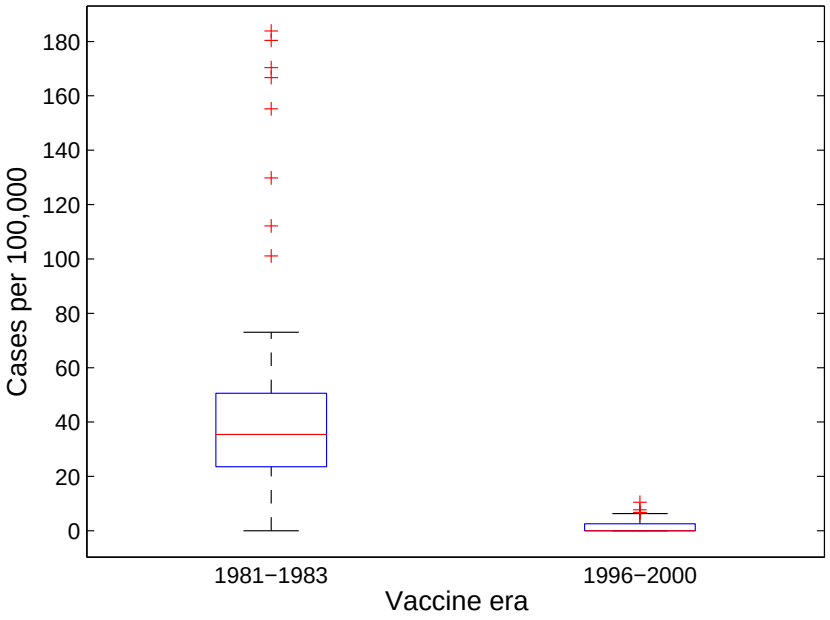

Figure A8. Box plot displaying the distribution of cases per 100,000 in infants across all provinces between vaccine eras.

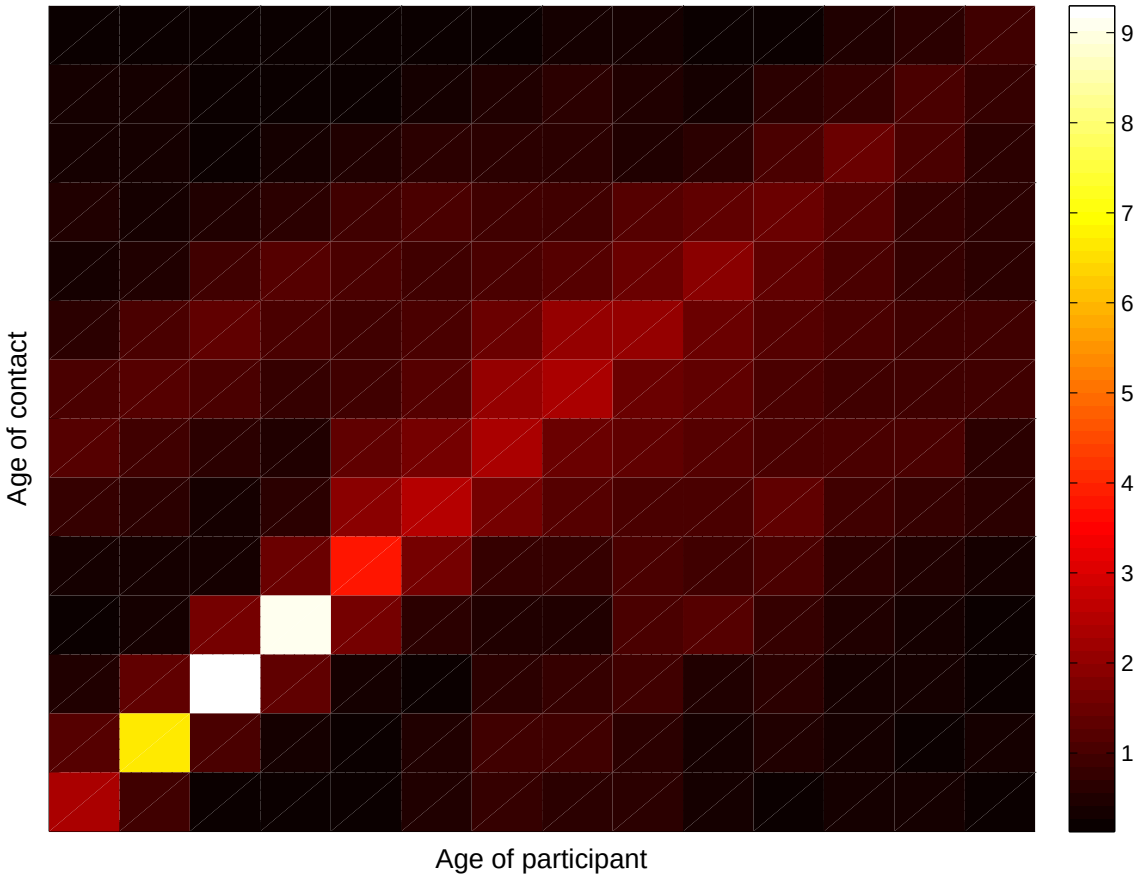

Figure A9. The average of the contact matrices for each country studied as found in [3]. Each row and column correspond to an age group. Age groups start at 0-4 years of age and increase by intervals of five years through age 69. Ages >70 are in the last row/column.

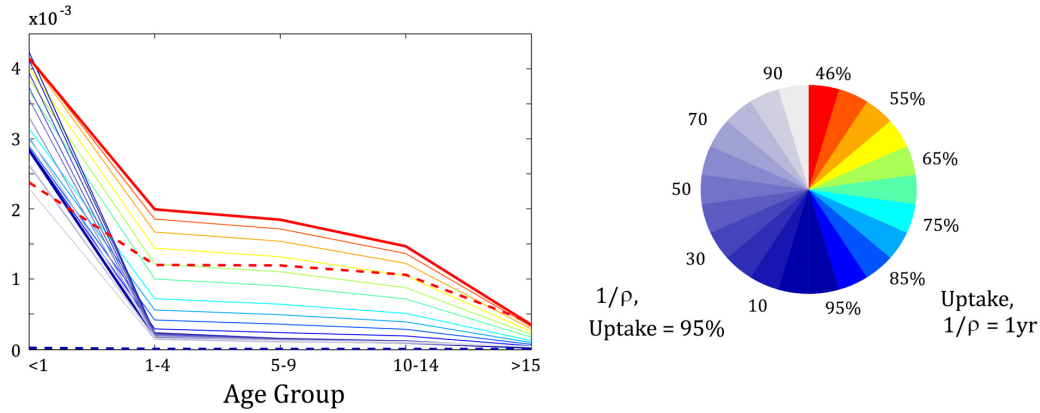

Figure A10. Similar to Fig. 5A in the main text, but this figure uses mixing according to contact patterns reported in [3]. Results from the SIV_IR are displayed in as the equilibrium prevalence by age group, I_i^*/n_i, where n_i is the fraction of the total population in age group i. Prevalence is corrected to account for the difference between the modeled infectious period (21 d) versus actual case notifications which are reported monthly. This is shown for various levels of vaccine uptake while holding the duration of immunity $1/\rho$ at one year (right side of colormap circle) and various values $1/\rho$ while holding vaccine uptake at 95% (left side of colormap circle). The dashed lines display the results from the age-structured SIR model for low (red) and high (blue) vaccine uptake.