

An Introduction to Infectious Disease Modelling

Solutions to exercises

Emilia Vynnycky and Richard G. White
with an introduction by Paul E.M. Fine

Chapter 2 (Solutions)

How are models set up?

I. An introduction to difference equations

2.1 a) The difference equations are as follows:

$$\begin{aligned}\text{Humans: } S_{t+1}^h &= S_t^h - \lambda_t^h S_t^h + r i_t^h \\ i_{t+1}^h &= i_t^h + \lambda_t^h S_t^h - r i_t^h \\ \text{Mosquitoes } S_{t+1}^v &= S_t^v + b - \lambda_t^v S_t^v - \mu S_t^v \\ i_{t+1}^v &= i_t^v + \lambda_t^v S_t^v - \mu i_t^v\end{aligned}$$

Note that the compartments are defined to be the proportions, rather than the numbers of mosquitoes or humans that are susceptible or infected. Since b is the *per capita* birth rate into the mosquito population, we just need to add b into the equation for susceptible mosquitoes to account for births into the population.

b) The risk of infection among humans needs to account for the prevalence of infected mosquitoes, the number of mosquitoes per human, the biting rates of mosquitoes, and the probability that a bite by a mosquito leads to infection in a human.

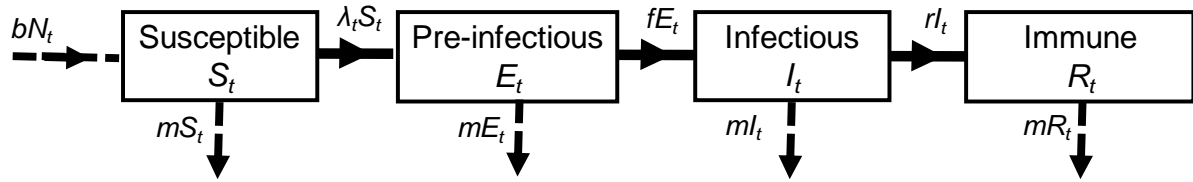
The risk of infection among mosquitoes needs to account for the biting rate of mosquitoes, the probability that a bite by a mosquito leads to infection in a mosquito, and the prevalence of infectious humans.

2.2 We use the symbols b and m to denote the *per capita* birth and death rates respectively, and the symbol N_t to denote the population size at time t .

a) The equations can be rewritten as follows:

$$\begin{aligned}S_{t+1} &= bN_t + S_t - \lambda_t S_t - mS_t \\ E_{t+1} &= E_t + \lambda_t S_t - f E_t - mE_t \\ I_{t+1} &= I_t + f E_t - r I_t - mI_t \\ R_{t+1} &= R_t + r I_t - mR_t\end{aligned}$$

The diagram for this model is provided below, where the expressions next to or above the arrows reflect the number of individuals who move between categories per unit time:



b) The model describes the transmission dynamics of an immunizing infection, and is therefore sufficient for describing the general patterns in incidence for measles and rubella, which are both immunizing infections. There are several ways of making the model more realistic than it is at present:

1. Stratify the compartments by age;
2. Include age-dependent contact between individuals;
3. Include changes in mixing patterns during the course of a year because of school holidays and school terms;
4. Assume that infectious individuals have a different mortality rate from those who are susceptible or are immune;
5. Include maternal immunity.

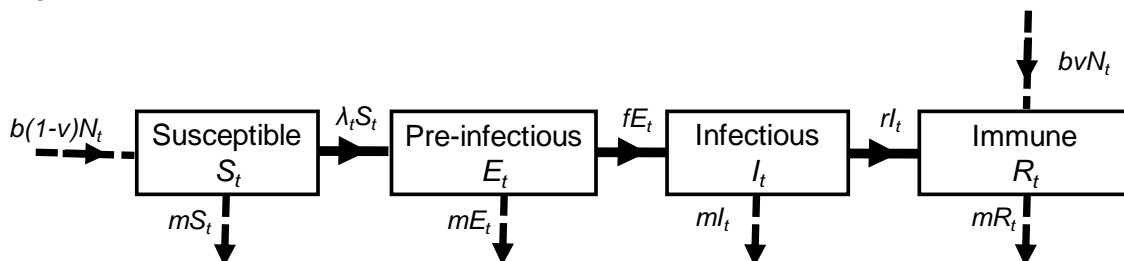
These issues are discussed in later chapters.

c) The equations would be rewritten as follows:

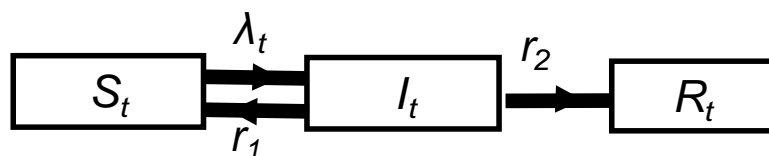
$$\begin{aligned}
 S_{t+1} &= S_t + b(1-v)N_t - \lambda_t S_t - mS_t \\
 E_{t+1} &= E_t + \lambda_t S_t - fE_t - mE_t \\
 I_{t+1} &= I_t + fE_t - rI_t - mI_t \\
 R_{t+1} &= R_t + bvN_t + rI_t - mR_t
 \end{aligned}$$

i.e. the proportion of newborns that is immunized enters the immune compartment and the remainder enters the susceptible compartment.

The diagram for this model is provided below, where the expressions next to or above the arrows reflect the number of individuals who move between categories per unit time:



2.3 The following shows the general structure of the model.



This model doesn't fall naturally into any of the categories presented in Figure 2.2. In fact, it has been called a "compound model" and has been used to describe hookworm data (see chapter 5).

Chapter 3 (Solutions)

How are models set up?

II. An introduction to differential equations

3.1 The differential equations are as follows:

$$\frac{dS(t)}{dt} = -\lambda(t)S(t) - vS(t)$$

$$\frac{dS_p(t)}{dt} = -\lambda(t)S_p(t) + w_v V(t) + w_n R(t)$$

$$\frac{dI(t)}{dt} = \lambda(t)S(t) - rI(t)$$

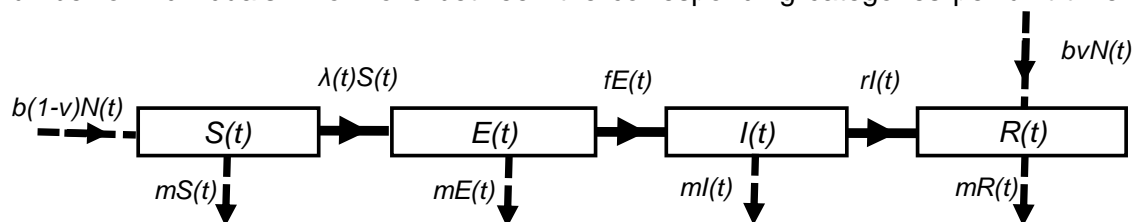
$$\frac{dI_p(t)}{dt} = \lambda(t)S_p(t) - rI_p(t)$$

$$\frac{dV(t)}{dt} = vS(t) - w_v V(t)$$

$$\frac{dR(t)}{dt} = rI(t) - w_n R(t) + rI_p(t)$$

b) The authors would have chosen to use this model rather than an SIRS model so that they could allow the duration of immunity and the infectiousness of infectious persons to depend on whether or not individuals have been infected naturally or vaccinated. The model structure used also allows the authors to allow the susceptibility to infection to differ between those who have been vaccinated and those who have been neither vaccinated nor infected. However, a drawback of having such a high level of detail in the model is that not all of the input parameters that are needed may be known.

3.2 a) The model diagram is as follows; the expressions next to the arrows reflect the number of individuals who move between the corresponding categories per unit time:



b) i) m is interpretable as the *per capita* mortality rate, which is assumed to be identical for all individuals.

ii) $N(t)$ is the total population size at time t .

c) Since both vaccinated individuals and those who are immune because of natural infection have been put into the same compartment, the model assumes that natural infection and vaccination provide the same level of protection..

3.3 a) The differential equations are as follows:

$$\frac{ds(t)}{dt} = -\lambda s(t)$$

$$\frac{dz(t)}{dt} = \lambda s(t)$$

Notice that we have used the symbol λ for the force of infection in these equations, rather than $\lambda(t)$. This reflects the fact that the force of infection is assumed to be the same over time.

b) i) The equations are: $s(t)=e^{-\lambda t}$ or $s_t=(1-\lambda_r)^t$ where λ_r is the risk of infection in each year of life.

ii) Assuming that the proportion (ever) infection is just 1-proportion susceptible, then the proportion ever infected is given by the following:

$$z(t) = 1 - e^{-\lambda t}, \text{ or}$$

$$z_t = 1 - (1 - \lambda_r)^t$$

c) The following table provides the corresponding values for the proportion ever infected by different ages:

Age (years)	Force of infection (% per year)		
	1%	10%	20%
5	0.0488	0.3935	0.6321
10	0.0952	0.6321	0.8647
20	0.1813	0.8647	0.9817
60	0.4512	0.9975	1.0000

Almost all individuals are predicted to have been infected by age 20 years in the high transmission setting. Since rubella is an immunizing infection, i.e. once infected, individuals are immune for life, very few individuals are infected as adults in high transmission settings. The burden of rubella among adults is therefore likely to be smallest in the high transmission setting. These issues are discussed further in chapter 5.

Chapter 4 (Solutions)

What do models tell us about the dynamics of infections?

4.1 a) Figure S4.1a plots the observed data for Gothenburg. This shows that two pandemic waves occurred, with the first occurring in July 1918 and the second occurring in September-October 1918. The cumulative numbers of cases for these waves are shown in Figure S4.1b. The natural log of the cumulative numbers of cases for these two waves are shown in Figure S4.1c.

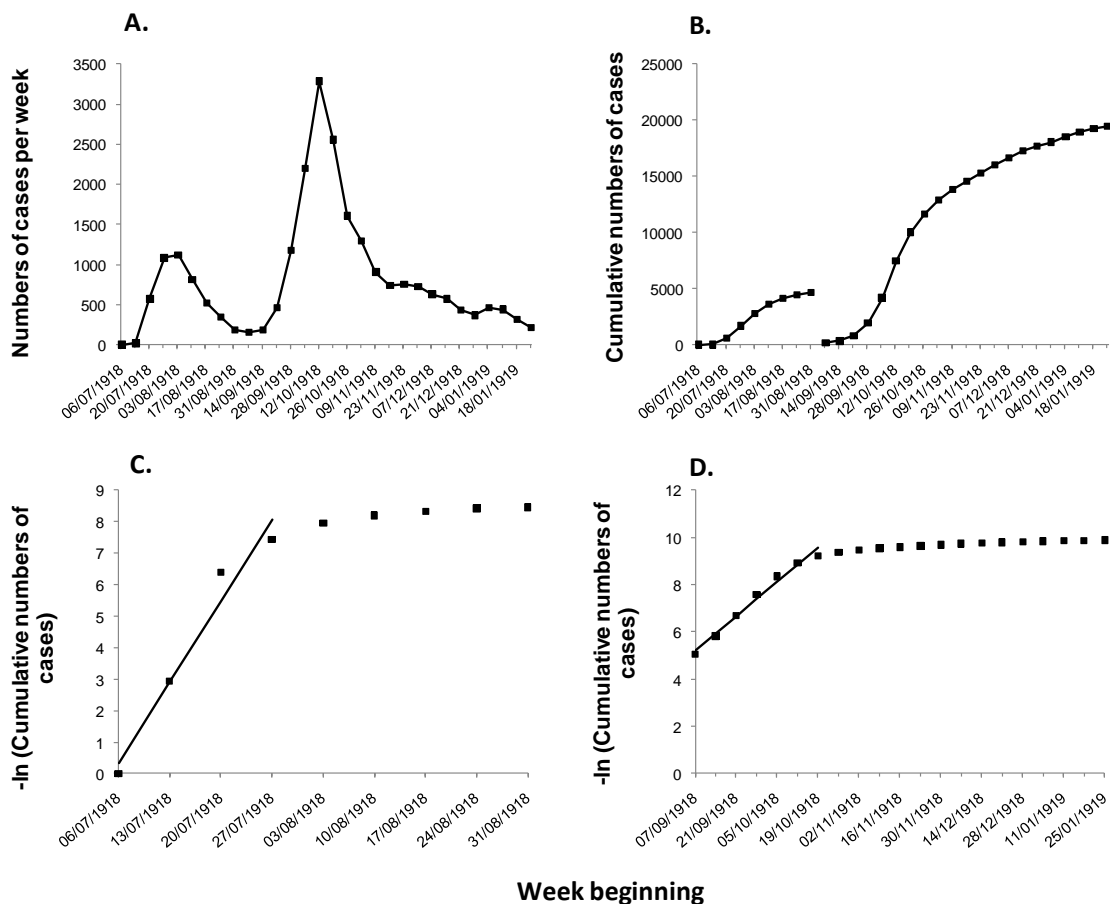


Figure S4.1: Summary of A. The numbers of cases reported each week, B. The cumulative numbers of reported cases and C. and D. the natural log of the cumulative numbers of cases observed during the first and second waves (C. and D. respectively) of the influenza pandemic in Gothenberg, Sweden in 1918.

A straight line can be drawn through the first 4 points of the natural log of the cumulative numbers of cases for the first wave (corresponding to the period 6/7/1918-27/7/1918); this line (drawn either by eye or formally by regression) has a slope of 0.367 per day.

A straight line can be drawn through the first 7 points of the natural log of the cumulative numbers of cases for the second wave (corresponding to the period 7/9/1918-19/10/1918). This line has a slope of 0.104 per day.

The following summarizes the estimates for the net and basic reproduction numbers obtained using the different formulae in Table 4.1, with R_0 estimated to be about 4 for the first wave and just under 3 for the second wave of the pandemic. These estimates are slightly higher than the values that have typically been estimated for the 1918 (Spanish) influenza pandemic (see references in the book for details). Notice that

estimates obtained using the formula $(1+\Delta D)(1+\Delta D')$ and $\frac{\Delta D \left(\frac{\Delta D'}{m} + 1 \right)^m}{\left(1 - \left(\frac{\Delta D}{n} + 1 \right)^{-n} \right)}$ are very similar.

Equation used to calculate the reproduction number:		R_n		R_0^*	
		1 st wave	2 nd wave	1 st wave ($s=0.7$)	2 nd wave ($s=0.5$)
$1+\Delta D$		1.73	1.21	2.48	2.42
$(1+\Delta D)(1+\Delta D')$		3.01	1.46	4.30	2.92
$\frac{\Delta D \left(\frac{\Delta D'}{m} + 1 \right)^m}{\left(1 - \left(\frac{\Delta D}{n} + 1 \right)^{-n} \right)}$	$m=n=10$	2.94	1.37	4.20	2.75
	$m=n=100$	2.94	1.36	4.20	2.73

* Calculated using the expression $R_n/(proportion\ susceptible\ (s))$ at the start of the wave)

b) It would be sensible to apply the epidemic size formula to data from the two waves separately.

Considering the first wave of the pandemic (taken to be during the period 6/7/1918-31/8/1918), 4,657 individuals were reported to have experienced disease. If 70% of individuals were susceptible at the start of the first wave ($s_0=0.7$), the proportion that were susceptible at the end of the first wave is given by the difference between 0.7 and the proportion of the population who experienced disease during the first wave. This calculation assumes that all of those who were reported as cases became immune (see below). We therefore obtain the following result:

$$s_f = 0.7 - 4,657/196,943 \approx 0.676$$

Substituting for s_0 and s_f into the equation $R_0 = \frac{\ln(s_f) - \ln(s_0)}{s_f - s_0}$ implies that R_0 equals the following:

$$R_0 = \frac{\ln(0.676) - \ln(0.7)}{0.676 - 0.7} = 1.45$$

Considering the second wave of the pandemic (the period after 7/9/1918), 19,484 individuals were reported to have experienced disease. Assuming that 50% of individuals were susceptible at the start of this wave ($s_0=0.5$), then applying a similar reasoning to that used to calculate s_f for the first wave, we obtain the following for the proportion of the population that was susceptible at the end of the second wave:

$$s_f = 0.5 - 19,484/196,943 \approx 0.401$$

Substituting for s_0 and s_f into the equation $R_0 = \frac{\ln(s_f) - \ln(s_0)}{s_f - s_0}$ implies that R_0 equals the following:

$$R_0 = \frac{\ln(0.401) - \ln(0.5)}{0.401 - 0.5} = 2.23$$

c) The estimates of R_0 that are based on the growth rate are likely to be more reliable than are those based on the final epidemic size, since they are independent of the proportion of cases that are reported (unless this changes over time). It is unlikely that all cases were reported during the pandemic, and therefore R_0 based on the epidemic size is likely to have been underestimated. However, estimates based on both methods need to make assumptions about the proportion of individuals that are susceptible at the start of the first and second waves. Whilst the values assumed (70% and 50% for the start of the first and second waves respectively) are plausible, it is unclear as to whether they are correct.

d) The lower estimate of R_0 for the second wave (calculated using the epidemic growth rate), as compared with that for the first wave suggests that in Gothenberg, the transmissibility decreased between the first and second waves. However, the value for R_0 during the first wave seems somewhat high in contrast with estimates obtained elsewhere (see references cited in the main text) and it seems plausible that the proportion of cases that were reported changed during the early stages of the first wave of the pandemic. Such changes in the proportion of cases that were reported would have led to an overestimate in R_0 .

4.2 a) According to equation 4.31, the inter-epidemic period is given by the following:

$$T = 2\pi \sqrt{\frac{L(D + D')}{R_0 - 1}}$$

We can rearrange this equation to obtain the following equation for R_0 :

$$R_0 = 1 + \frac{4\pi^2 L(D + D')}{T^2}$$

Substituting for $L=70 \times 365$ days, $T=2 \times 365$ days, $D'=8$ days and $D=7$ days into this equation leads to the following:

$$R_0 = 1 + \frac{4\pi^2 \times 70 \times 365 \times (7 + 8)}{(2 \times 365)^2} \approx 29$$

b) According to equation 4.32, the inter-epidemic period (T) is given by the equation:

$$T = 2\pi \sqrt{A(D + D')}$$

This equation can be rearranged to give the following for the average age at infection:

$$A = \frac{T^2}{4\pi^2(D + D')} \quad \text{S4.1}$$

Substituting for $T=3 \times 365$ days, $D'=8$ days and $D=7$ days into this equation implies that

$$A = \frac{(3 \times 365)^2}{4\pi^2(7 + 8)} \approx 2,025 \text{ days} = 2,025/365 \approx 5.5 \text{ years.}$$

The limitations of this estimate are as follows:

- i) The equation on which it is based assumes that individuals mix randomly, which is unrealistic (see chapter 7).
- ii) The measles vaccination coverage increased after vaccination was introduced in 1968, and therefore the inter-epidemic period would have changed over time. This equation does not account for changes in the inter-epidemic period over time.

Chapter 5 (Solutions)

Age patterns

5.1 Adapting expression 5.25 in the book, the number of new infections per person among individuals in age group a can be calculated using the expression:

$$\lambda s_a$$

where s_a is the proportion of individuals in age group a who were seronegative. The values obtained for the average number of new infections per 100 population are as follows:

	Average number of infections per year per 100 population (calculated using $\lambda s_a \times 100$)		
Age group (years)	China $\lambda=20\%/yr$	Fiji $\lambda=4\%/yr$	UK $\lambda=12\%/yr$
15-19	0.80 ($=0.2 \times 0.04 \times 100$)	1.74 ($=0.04 \times 0.435 \times 100$)	1.54 ($=0.12 \times 0.128 \times 100$)
20-29	0.86 ($=0.2 \times 0.043 \times 100$)	1.15 ($=0.04 \times 0.288 \times 100$)	1.04 ($=0.12 \times 0.087 \times 100$)
30-39	1.08 ($=0.2 \times 0.054 \times 100$)	0.77 ($=0.04 \times 0.193 \times 100$)	0.85 ($=0.12 \times 0.071 \times 100$)

We could have also used the expression $s_a \times \text{risk of infection}$, which, using the relationship $\text{risk} = 1 - e^{-\text{rate}}$ (Panel 2.2), leads to the expression $s_a(1 - e^{-\lambda})$ for the number of new infections per person. This expression leads to the following values for the average infection incidence:

	Average number of infections per year per 100 population (calculated using $s_a(1 - e^{-\lambda}) \times 100$)		
Age group (years)	China $\lambda=20\%/yr$	Fiji $\lambda=4\%/yr$	UK $\lambda=12\%/yr$
15-19	0.73	1.71	1.45
20-29	0.78	1.13	0.98
30-39	0.98	0.76	0.80

Note that the greatest discrepancy between the values obtained using the expressions λs_a and $s_a(1 - e^{-\lambda})$ occurs for the estimates for China. This follows from the facts that the force of infection is higher for China than it is for the UK and Fiji, and the difference between the value for the risk $(1 - e^{-\lambda})$ and the rate (λ) is greatest for large values of the rate (see Figure 2.5).

The estimates suggest that the highest number of new infections per 100 population would have been seen among 15-19 year olds in Fiji, followed closely by that for 15-19 year olds in the UK. Therefore, based on these estimates, we would expect the incidence of CRS to have been greater among babies born to women in these age groups in these countries, than for babies born to women in other age groups in the same countries.

However, the overall burden of CRS depends both on the infection incidence and number of livebirths among women in different age groups. Therefore, to infer the setting in which the burden of CRS is likely to be the greatest, we would need to combine the above estimates with the age-specific fertility rate.

5.2 a) Figure S5.1 plots the observed proportion seronegative. The median age at infection is the point at which the vertical dotted line in this figure crosses the x-axis, which occurs at about 6 years. This estimate suggests that the average force of infection is about $100 \times 1/6 \approx 17\%$ per year.

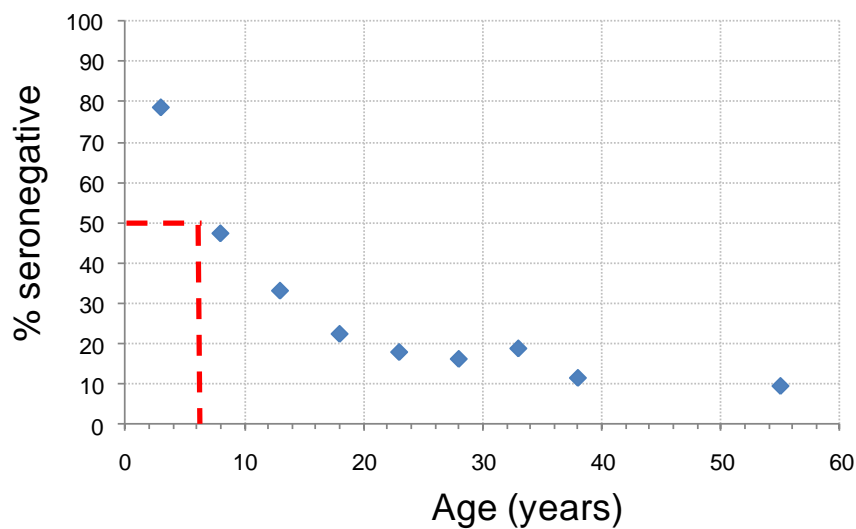


Figure S5.1: Observed proportion of individuals who did not have antibodies to rubella during 2004-5 in Bangladesh¹

b) i) The overall proportion susceptible is calculated using equation 5.17 as the sum of the proportion susceptible in each age group, weighted by the proportion of the population that is in that age group ($p_a \times S_a / N_a$). The final column in the following table gives the values for $p_a \times S_a / N_a$ in each age group.

Age group (years)	Number tested	Number negative	% negative	Proportion of the population that is in the given age range (p_a)	$p_a \times S_a / N_a$
1-5	61	48	78.7	0.1139	0.08964
6-10	61	29	47.5	0.1135	0.05391
11-15	63	21	33.3	0.1109	0.03693
16-20	62	14	22.6	0.1087	0.02457
21-25	83	15	18.1	0.104	0.01882
26-30	67	11	16.4	0.0923	0.01514
31-35	63	12	19.0	0.0775	0.01473
36-40	60	7	11.7	0.0641	0.0075
≥ 41	62	6	9.7	0.2151	0.02086

The overall proportion susceptible is therefore given by the sum of the values in the final column, i.e. 0.2821.

ii) The basic reproduction number can be estimated using the expression $1/s$ (equation 5.20), where s is the proportion of the population that is susceptible. Using the value for s obtained in part i) implies that $R_0 = 1/0.2821 \approx 3.5$.

iii) We can obtain an expression for the force of infection in terms of R_0 after rearranging either the expression $R_0 = 1 + \lambda L$ or $R_0 = \lambda L$, depending on whether the age distribution of the population is exponential or rectangular respectively. Figure S5.2, which plots the values for p_a for 5 year age groups in Bangladesh in 2005² suggests that the age distribution was closer to being exponential than to being rectangular.

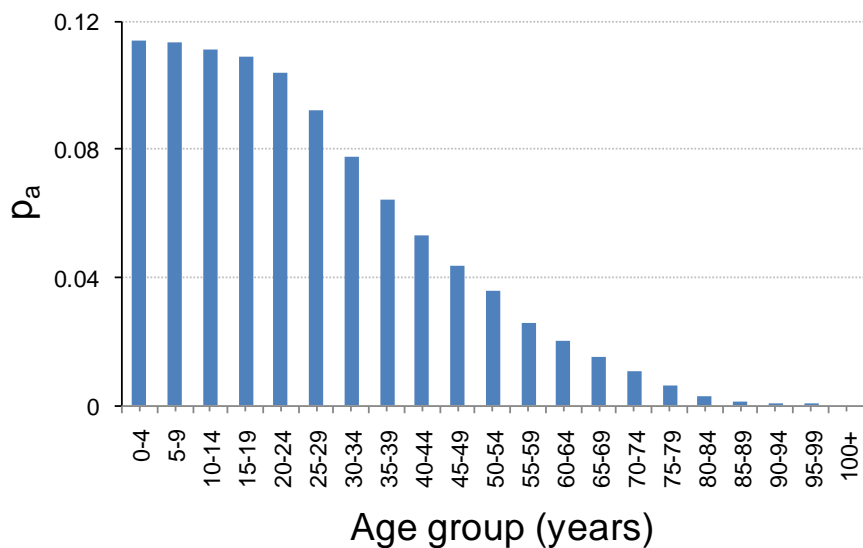


Figure S5.2: Proportion of individuals in different age groups in Bangladesh in 2005 (p_a).

Rearranging the expression $R_0 = 1 + \lambda L$ gives the following expression for λ :

$$\lambda = \frac{R_0 - 1}{L}$$

Substituting for $L=65$ years and the value for R_0 obtained in part ii) implies that the force of infection equals:

$$\lambda \approx \frac{3.5-1}{65} \approx 4\% \text{ per year}$$

iv) Using the relationship $A = \frac{1}{\lambda + 1/L}$ (equation 5.10) and substituting for $L=65$ years and the value for λ obtained in part iii) implies that the average age at infection is approximately:

$$A = \frac{1}{0.04 + 1/65} \approx 18 \text{ years}$$

Note that the values for A and λ are much smaller than those obtained in part a). This follows largely from the fact that estimates in part a) did not account for the age distribution of the population.

c) The following table summarizes the estimates for the average annual risk of infection calculated using the expression $\lambda = 1 - s_a^{1/a}$, where a was taken as the midpoint of the age group for the corresponding data point. The force of infection was calculated using the result $\text{rate} \approx -\ln(1 - \text{risk})$ (see page 111).

Age group (years)	% negative	Average annual risk of infection, calculated using $\lambda = 1 - s_a^{1/a}$	Average annual force of infection
1-5	78.7	0.0767	0.0798
6-10	47.5	0.0889	0.0931
11-15	33.3	0.0811	0.0846
16-20	22.6	0.0793	0.0826
21-25	18.1	0.0716	0.0743
26-30	16.4	0.0625	0.0646
31-35	19.0	0.0491	0.0503
36-40	11.7	0.0549	0.0565
≥ 41	9.7	0.0415	0.0424

We can also use the equation $\lambda_a = 1 - \frac{s_{a+1}}{s_a}$. However, when substituting s_a and s_{a+1}

into this equation, we obtain the risk of infection between age band a and age band $a+1$. Since each age band is of width 5 years, this infection risk is equivalent to a five year risk. We can convert this five year risk into an annual risk by adapting the logic described in section 5.2.2.1.4, which leads to the following equation for the average annual risk in age group a :

$$\lambda = 1 - (1 - \lambda_a)^{1/5}$$

The force of infection is then calculated using the result $\text{rate} \approx -\ln(1 - \text{risk})$ (see page 111).

The following table summarizes the estimates obtained using this approach:

Age group (years)	% negative	5 year risk of infection, calculated using: $\lambda_a = 1 - \frac{S_{a+1}}{S_a}$	Average annual risk of infection, calculated using: $\lambda = 1 - (1 - \lambda_a)^{1/a}$	Average annual force of infection
1-5	78.7	0.3964	0.0961	0.0798
6-10	47.5	0.2989	0.0686	0.0931
11-15	33.3	0.3213	0.0746	0.0846
16-20	22.6	0.1991	0.0434	0.0826
21-25	18.1	0.0939	0.0195	0.0743
26-30	16.4	-0.1585	-0.0299	0.0646
31-35	19.0	0.3842	0.0924	0.0503
36-40	11.7	0.1709	0.0368	0.0565
≥41	9.7	-	-	0.0424

The estimates obtained using both approaches suggest that the force of infection for adults is lower than that for children, e.g. >8% per year for those aged <20 years and <8% per year for those aged >20 years. However, the estimates for adults that are based on the equation $\lambda = 1 - \lambda_a^{1/a}$ are difficult to interpret, since the proportion of 31-35 year olds who are seronegative is smaller than that for 26-30 year olds, which leads to the (unrealistic) estimate that the risk of infection was negative between the ages 21-25 and 26-30 year olds.

d) Figure S5.3 shows a plot of $-\ln(\text{observed proportion seronegative})$ against the age midpoints for the data from Bangladesh. These figures also clearly highlight the fact that the datapoint for individuals aged 31-35 years is an outlier.

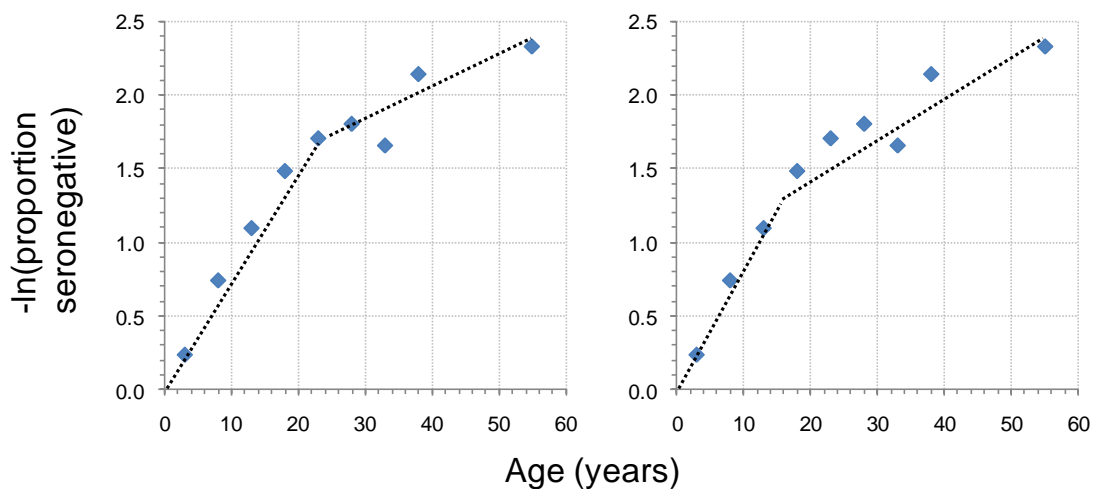


Figure S5.3: Plot $-\ln(\text{observed proportion seronegative})$ against the age midpoints for the data from Bangladesh in Nessa *et al*¹, with different lines drawn by eye through the data points for individuals aged <20 years (left-hand figure) and for those aged <15 years (right-hand figure).

As shown in the left-hand figure, the gradient of the line through the points for individuals aged <20 years is steeper than that through the points for individuals aged >20 years, suggesting that the force of infection is greater for those aged <20 years than for those aged >20 years.

However, based on these plots, we cannot conclude that the force of infection changes at age 20 years, since, as shown in the right-hand figure, we can also draw a straight line through the points for individuals aged over 15 years, which would imply that the force of infection changes at about this age. Ultimately, the age at which the force of infection is assumed to change needs to be biologically plausible, e.g. consistent with changes in behaviour, possible exposure to the infection, contact patterns etc at ages 15 or 20 years.

The gradient of the line through the points for individuals aged <15 years is about $1.6/15 \approx 0.11$ per year, suggesting that the force of infection in this age group is about 11% per year. The gradient of the line through the points for individuals aged >15 years is about $0.7/35 \approx 0.02$ per year, suggesting that the force of infection in this age group is about 2% per year.

5.3 The following figure shows a plot of $-\ln(\text{observed proportion seropositive})$ for the mumps data in section 5.2.3.2.2. The gradient of the line through the points for individuals aged <13 years is steeper than that through the points for individuals aged >13 years ($3/13 \approx 0.23$ per year vs $1/35 \approx 0.03$ per year). This suggests that the force of infection is also greater for those aged <13 years than for those aged >13 years (23% vs 3% per year respectively).

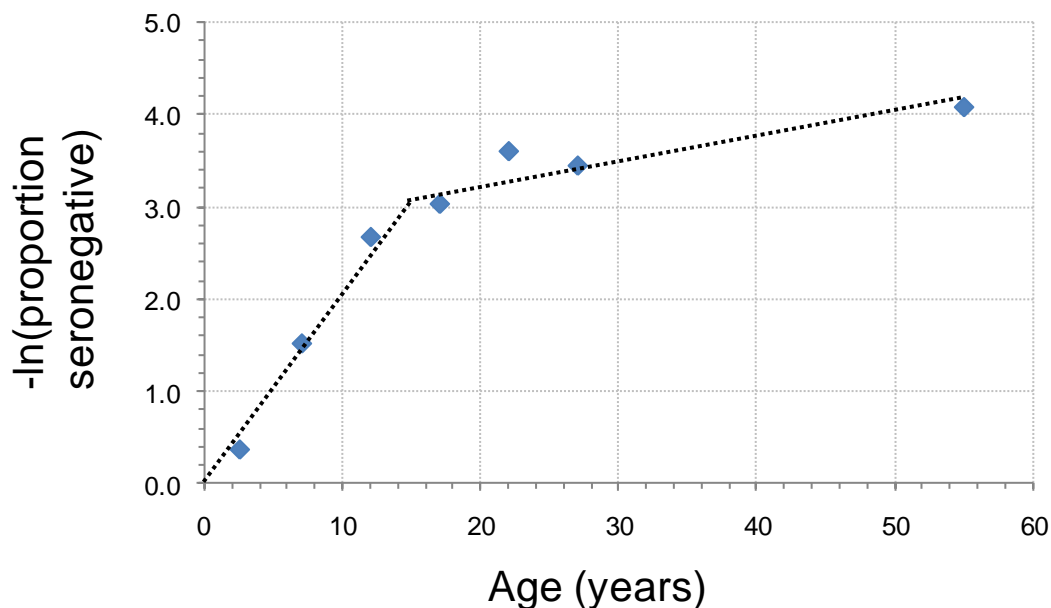


Figure S5.4: Plot of $-\ln(\text{observed proportion seronegative})$ for the mumps data in section 5.2.3.2.2.

5.4 i) One informal argument that is sometimes used to obtain this result is that, for realistic values of L , $1/L$ is small, in comparison with λ , and therefore $\frac{1}{\lambda + 1/L}$ must be approximately equal to $\frac{1}{\lambda}$.

We can also apply a formal mathematical argument, which uses the result that, for small values of x (i.e. values that are close to zero), the expression $\frac{1}{1+x}$ is approximately equal to $1-x$ (see proof at the end of the solution to this question).

This argument is as follows:

We begin by noting that the equation $A = \frac{1}{\lambda + 1/L}$ can be rewritten in the form $\frac{1}{1+x}$ as follows:

$$A = \frac{1}{\lambda} \left(\frac{1}{1 + \frac{1}{\lambda L}} \right) \quad \text{S5.1}$$

For realistic values of the average life expectancy and for large values of the force of infection, $\frac{1}{\lambda L}$ is close to zero, and so, according to the result $\frac{1}{1+x} \approx 1-x$, the term in brackets in equation S5.1 is approximately equal to $1 - \frac{1}{\lambda L}$. Substituting this approximation into equation S5.1 leads to the following:

$$A \approx \frac{1}{\lambda} \left(1 - \frac{1}{\lambda L} \right) = \frac{1}{\lambda} - \frac{1}{\lambda^2 L}$$

If both the force of infection and the life expectancy are sufficiently large, then the second term in this equation is negligible (i.e. $\frac{1}{\lambda^2 L} \approx 0$) and so $A \approx \frac{1}{\lambda}$.

ii) To show that the expression $A = \frac{1}{\lambda} \left(\frac{1 - (1 + \lambda L)e^{-\lambda L}}{1 - e^{-\lambda L}} \right)$ approximates to $1/\lambda$, we begin

by observing that, for sufficiently large values for the life expectancy and the force of infection, $e^{-\lambda L}$ is close to zero. Using the result that for small values of x (i.e. values that are close to zero), $\frac{1}{1-x} \approx 1+x$, we see that $\frac{1}{1-e^{-\lambda L}} \approx 1+e^{-\lambda L}$. Substituting this approximation into the equation for A , we obtain the following result:

$$A = \frac{1}{\lambda} \left(\frac{1 - (1 + \lambda L)e^{-\lambda L}}{1 - e^{-\lambda L}} \right) \approx \frac{(1 - (1 + \lambda L)e^{-\lambda L})(1 + e^{-\lambda L})}{\lambda}$$

This equation simplifies to the following:

$$A \approx \frac{1 - e^{-2\lambda L}(1 + \lambda L) - \lambda L e^{-\lambda L}}{\lambda}$$

If the force of infection is sufficiently large and for realistic values for the life expectancy, both $e^{-2\lambda L}(1 + \lambda L) \approx 0$ and $\lambda L e^{-\lambda L} \approx 0$, which implies that $A \approx \frac{1}{\lambda}$

Proof of the result $\frac{1}{1+x} \approx 1-x$ **for small values of x**

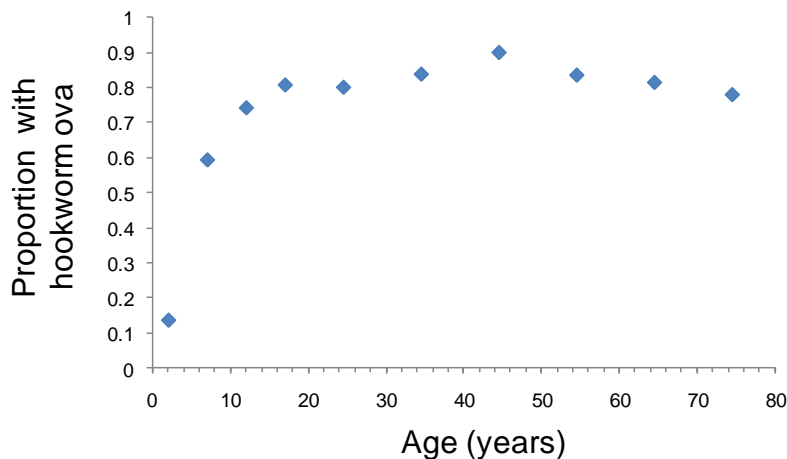
This result can be derived by using the fact that an expression of the form $\frac{1}{1+x}$ can be written using the following Binomial expansion:

$$\begin{aligned} \frac{1}{1+x} &= (1+x)^{-1} = 1 - x + \frac{(-1)(-1-1)x^2}{2!} + \frac{(-1)(-1-1)(-1-2)x^3}{3!} + \dots \\ &\quad + \frac{(-1)(-1-1)(-1-2)\dots(-1-n)x^{n-1}}{(n-1)!} + \dots \\ &= 1 - x + x^2 - x^3 + \dots + (-1)^{n-1} x^{n-1} + \dots \end{aligned}$$

For small values of x , terms in x^2 , x^3 etc (known as “higher order terms”) are small and can be ignored. Consequently $\frac{1}{1+x} \approx 1-x$.

The result that $\frac{1}{1-x} \approx 1+x$ follows after repeating the above argument but replacing $-x$ for x .

5.5 The following figure shows that the proportion of individuals who had hookworm ova in their stools (S_a/N_a) increases with age and then reaches a plateau or (plausibly) decreases with increasing age.



We might therefore use a reversible model to describe the data, which assumes that the age-specific proportion infected eventually reaches a plateau with increasing age. Alternatively, a 2-stage model or a compound catalytic model might be appropriate, since these assume that the proportion positive peaks before subsequently decreasing with increasing age. In fact, the authors of ³ used a compound model to analyse the data.

5.6 a) Assuming that maternal immunity is lost at a constant rate, μ , the rate of change in the proportion of individuals who have maternal immunity ($m(a)$) and the proportion who are susceptible ($s(a)$) are given by the following equations:

$$\begin{aligned}\frac{dm(a)}{da} &= -\mu m(a) \\ \frac{ds(a)}{da} &= \mu m(a) - \lambda s(a)\end{aligned}$$

The differential equation for $m(a)$ is of the form $\frac{dQ(t)}{dt} = -kQ(t)$ and can be solved to give the following (see section 3.5.1):

$$m(a) = m(0)e^{-\mu a}$$

where $m(0)$ is the proportion of newborns who have maternal immunity. Since all individuals are assumed to be born with maternal immunity, $m(0)=1$, and so, substituting for $m(0)=1$ into the above equation gives $m(a)=e^{-\mu a}$.

Substituting for $m(a)=e^{-\mu a}$ into the differential equation for $s(a)$, we obtain the following equation:

$$\frac{ds(a)}{da} = \mu e^{-\mu a} - \lambda s(a)$$

which can be rewritten as follows:

$$\frac{ds(a)}{da} + \lambda s(a) = \mu e^{-\mu a}$$

This equation can be solved using the technique of “integrating factors” by following the steps below:

Step 1. Multiply both sides of the equation by $e^{\lambda a}$, to obtain the following:

$$\frac{ds(a)}{da} e^{\lambda a} + \lambda s(a) e^{\lambda a} = \mu e^{-\mu a} e^{\lambda a}$$

Note that, according to the rules of differentiation (section B.5), the left-hand side of this equation is equivalent to the derivative of $s(a)e^{\lambda a}$ with respect to a , and so the equation can be rewritten as follows:

$$\frac{d}{da}(s(a)e^{\lambda a}) = \mu e^{-\mu a} e^{\lambda a} \quad \text{S5.2}$$

Step 2. We now integrate both sides of equation S5.2 between 0 and a to obtain the following:

$$\int_0^a \frac{d}{da} (s(a)e^{\lambda a}) da = \int_0^a \mu e^{-\mu a} e^{\lambda a} da \quad \text{S5.3}$$

Since integration is the converse of differentiation, the left-hand side equation S5.3 simplifies to:

$$[s(a)e^{\lambda a}]_0^a = s(a)e^{\lambda a} - s(0)$$

However, $s(0)=0$ (since no individuals are assumed to be susceptible at birth) and therefore the left-hand side of equation S5.3 simplifies to $s(a)e^{\lambda a}$.

By the rules of integration (section B.6) the right-hand side of equation S5.3 simplifies to the following:

$$\left[\frac{\mu}{\lambda - \mu} e^{(\lambda - \mu)a} \right]_0^a = \frac{\mu}{\lambda - \mu} e^{(\lambda - \mu)a} - \frac{\mu}{\lambda - \mu}$$

Step 3. Equating the expressions obtained from integrating the left-hand and right-hand sides of equation S5.3 leads to the following:

$$s(a)e^{\lambda a} = \frac{\mu}{\lambda - \mu} e^{(\lambda - \mu)a} - \frac{\mu}{\lambda - \mu}$$

Dividing both sides of this equation by $e^{\lambda a}$ leads to the intended result:

$$s(a) = \frac{\mu(e^{-\mu a} - e^{-\lambda a})}{\lambda - \mu}$$

b) Note that when $s(a)$ is at a minimum, $\frac{ds(a)}{da} = 0$. We can therefore obtain the age at which the proportion of the population that is susceptible is at a minimum by identifying the values for a for which $\frac{ds(a)}{da} = 0$.

Differentiating the equation for $s(a)$ that is discussed in part a), we obtain the following:

$$\frac{ds(a)}{da} = \frac{\mu(-\mu e^{-\mu a} + \lambda e^{-\lambda a})}{\lambda - \mu}$$

Setting this equation to zero, we see that the following must be satisfied for the proportion susceptible to be at a minimum:

$$-\mu e^{-\mu a} + \lambda e^{-\lambda a} = 0$$

Multiplying both sides of this equation by $e^{\lambda a}$ and rearranging the resulting equation, implies that the following must hold:

$$e^{(\lambda-\mu)a} = \frac{\lambda}{\mu}$$

Taking the natural logs of both sides of this equation and then dividing by $\lambda-\mu$ leads to the intended result that the minimum in the proportion susceptible occurs when

$$a = \frac{\ln(\lambda/\mu)}{\lambda-\mu}$$

5.7 a) Proof of the result that $A = \frac{1}{\lambda+m}$, or equivalently, $A = \frac{1}{\lambda+1/L}$ for populations with an exponential age distribution, where $m=1/L$ is the average mortality rate.

Suppose that N_0 individuals are born each year. Assuming a constant mortality rate of m , the number of individuals of age a is given by the equation (see section 3.5.1):

$$N(a) = N_0 e^{-ma}$$

Assuming a constant force of infection, λ , a proportion $e^{-\lambda a}$ of these individuals will be susceptible, and so the number of susceptible individuals of age a ($S(a)$) is obtained by multiplying $N(a)$ by $e^{-\lambda a}$, i.e.

$$S(a) = N(a)e^{-\lambda a} = N_0 e^{-(\lambda+m)a}$$

After substituting this expression into equation 5.9, we obtain the following equation:

$$A = \frac{\int_0^\infty a\lambda(a)S(a)da}{\int_0^\infty \lambda(a)S(a)da} = \frac{\int_0^\infty a\lambda N_0 e^{-(\lambda+m)a} da}{\int_0^\infty \lambda N_0 e^{-(\lambda+m)a} da} \quad \text{S5.4}$$

Using the techniques discussed in section B.6, the numerator of this equation simplifies to the following:

$$\int_0^\infty a\lambda N_0 e^{-(\lambda+m)a} da = \lambda N_0 \left[\frac{ae^{-(\lambda+m)a}}{-(\lambda+m)} - \frac{e^{-(\lambda+m)a}}{(\lambda+m)^2} \right]_0^\infty = \frac{\lambda N_0}{(\lambda+m)^2} \quad \text{S5.5}$$

Similarly, the denominator in equation S5.5 simplifies to the following:

$$\int_0^\infty \lambda N_0 e^{-(\lambda+m)a} da = \lambda N_0 \left[\frac{e^{-(\lambda+m)a}}{-(\lambda+m)} \right]_0^\infty = \frac{\lambda N_0}{\lambda+m} \quad \text{S5.6}$$

Substituting the right-hand sides of equations S5.5 and S5.6 into the numerator and denominator of equation S5.4, and cancelling common terms from the numerator and denominator leads to our intended expression for A :

$$A = \frac{\frac{\lambda N_0}{(\lambda+m)^2}}{\frac{\lambda N_0}{\lambda+m}} = \frac{1}{\lambda+m}$$

b) Proof of the result that, for populations with a rectangular age distribution, with a life expectancy of L , and assuming random mixing,

$$A = \frac{1}{\lambda} \left(\frac{1 - (1 + \lambda L)e^{-\lambda L}}{1 - e^{-\lambda L}} \right)$$

Suppose that N_0 individuals are born into the population each year. If the population has a rectangular age distribution in which no individuals die until age L , then the number of individuals of age a also equals N_0 .

A proportion $e^{-\lambda a}$ of these individuals will be susceptible, and so the number of susceptible individuals of age a ($S(a)$) is obtained by multiplying N_0 by $e^{-\lambda a}$, i.e.

$$S(a) = N_0 e^{-\lambda a}$$

After substituting this expression into equation 5.9, we obtain the following equation:

$$A = \frac{\int_0^\infty a\lambda(a)S(a)da}{\int_0^\infty \lambda(a)S(a)da} = \frac{\int_0^L a\lambda N_0 e^{-\lambda a} da}{\int_0^L \lambda N_0 e^{-\lambda a} da} \quad \text{S5.7}$$

Using the techniques discussed in section B.5, the numerator of this expression simplifies to the following:

$$\begin{aligned} \int_0^L a\lambda N_0 e^{-\lambda a} da &= \lambda N_0 \left[\frac{ae^{-\lambda a}}{-\lambda} - \frac{e^{-\lambda a}}{\lambda^2} \right]_0^L = \lambda N_0 \left(\frac{Le^{-\lambda L}}{-\lambda} - \frac{e^{-\lambda L}}{\lambda^2} + \frac{1}{\lambda^2} \right) \\ &= \frac{N_0(1 - (1 + \lambda L)e^{-\lambda L})}{\lambda} \end{aligned} \quad \text{S5.8}$$

Similarly, the denominator in equation 5.7 simplifies to the following:

$$\begin{aligned} \int_0^L \lambda N_0 e^{-\lambda a} da &= \lambda N_0 \left[\frac{e^{-\lambda a}}{-\lambda} \right]_0^L = \frac{\lambda N_0(1 - e^{-\lambda L})}{\lambda} \\ &= N_0(1 - e^{-\lambda L}) \end{aligned} \quad \text{S5.9}$$

Substituting the right-hand sides of equations S5.8 and S5.9 into the numerator and denominator respectively of equation S5.7, and cancelling out the common term N_0 leads to the intended result:

$$A = \frac{N_0(1 - (1 + \lambda L)e^{-\lambda L})}{\lambda N_0(1 - e^{-\lambda L})} = \frac{1 - (1 + \lambda L)e^{-\lambda L}}{\lambda(1 - e^{-\lambda L})}$$

5.8 For a reversible catalytic model, the differential equations for the rate of change in the proportion susceptible and proportion currently infected is given by the following:

$$\begin{aligned}\frac{ds(a)}{da} &= -\lambda s(a) + r_s z(a) \\ \frac{dz(a)}{da} &= \lambda s(a) - r_s z(a)\end{aligned}$$

In this model, the proportion of individuals of age a that are currently susceptible is given by 1 -proportion currently infected, i.e. $s(a) = 1 - z(a)$.

Substituting this expression for $s(a)$ into the differential equation for $z(a)$ leads to the following equation:

$$\begin{aligned}\frac{dz(a)}{da} &= \lambda(1 - z(a)) - r_s z(a) \\ &= \lambda - (\lambda + r_s)z(a)\end{aligned}\tag{S5.10}$$

At a given point on the plateau, $\frac{dz(a)}{da} = 0$. Equating equation S5.10 to zero, leads to the following result:

$$\lambda - (\lambda + r_s)z(a) = 0$$

After rearranging this equation, we obtain our intended result, i.e. $z(a) = \frac{\lambda}{\lambda + r_s}$

5.9 a i) With the information provided, we can use the equation $A' = \frac{A}{(1 - v)}$ (equation 5.34) to work out the long-term average age at infection for mumps following the introduction of vaccination. Substituting for $A=4$ years and $v=0.6$ into this equation (assuming for now, that the vaccine efficacy is 100%), the long-term average age at infection is given by $A' = \frac{4}{(1 - 0.6)} = 10$ years.

Given the debate about the efficacy of the mumps component of the MMR vaccine ⁴⁻⁶, it would be sensible to assume a vaccine efficacy of $<100\%$. Assuming a vaccine efficacy of 85%, the long-term average age at infection equals:

$$A' = \frac{4}{(1 - 0.85 \times 0.6)} = \frac{4}{(1 - 0.51)} \approx 8.2 \text{ years}$$

a ii) The long-term average force of infection λ' can be obtained after rearranging the expression $A' = \frac{1}{\lambda' + m}$ (equation 5.33) and substituting our estimate of A' that we obtained in part i) into the resulting expression.

For example, the expression $A' = \frac{1}{\lambda' + m}$ can be rearranged to give the following expression for λ' :

$$\lambda' = \frac{1}{A'} - m$$

Substituting for $A' = 10$ years (based on a vaccine efficacy of 100%), and $m=1/60$ per year into this equation leads to the following value for λ' :

$$\lambda' = \frac{1}{10} - \frac{1}{60} = \frac{5}{60} = 0.0833 \text{ per year.}$$

Substituting for $A' = 8.2$ years (based on a vaccine efficacy of 85%) leads to an estimate for the average force of infection of :

$$\lambda' = \frac{1}{8.2} - \frac{1}{60} = 0.106 \text{ per year.}$$

a iii) and a iv) The proportion susceptible and the infection incidence in the long-term can be calculated using equations 5.31 and 5.36, leading to the following values:

	100% vaccine efficacy		85% vaccine efficacy	
Age (years)	Proportion susceptible $(s(a))' = (1-v)e^{-\lambda'a}$	Average annual number of new infections per 100,000 $(\lambda' s(a))' \times 100,000$	Proportion susceptible $(s(a))' = (1-v)e^{-\lambda'a}$	Average annual number of new infections per 100,000 $(\lambda' s(a))' \times 100,000$
15	0.115	955	0.1	1060
25	0.05	415	0.035	368
35	0.022	180	0.012	128

b) As shown by the calculations below, the average annual number of mumps infections per 100,000 population among 15-35 year olds in the long-term following the introduction of vaccination is somewhat higher than that before the introduction of vaccination. You might therefore advise the government to aim to attain a coverage which is much higher than 60% (e.g. 95%) and to proceed with caution when introducing MMR vaccination if it thinks that a coverage of only 60% can be achieved.

It might want to consider having a catch-up campaign covering the birth cohorts at greatest risk, and monitor the age-specific proportion susceptible in the population through seroprevalence surveys. Most importantly, before proceeding, it should also consider the effect that 60% coverage of MMR vaccination would have on the burden of measles, rubella and Congenital Rubella Syndrome.

Calculations of the number of mumps infections per 100,000 population before the introduction of vaccination:

For these calculations, we first need to estimate the force of infection that is predicted in the absence of vaccination. Rearranging the equation $A = \frac{1}{\lambda + m}$ (equation 5.10) we

obtain the following equation for the force of infection in the absence of vaccination:

$\lambda = \frac{1}{A} - m$. Substituting for $A=4$ years (the average age at infection before the introduction of vaccination) and $m=1/60$ per year into this equation, we obtain the following for λ :

$$\lambda = \frac{1}{4} - \frac{1}{60} = 0.233 \text{ per year.}$$

Using this estimate for the force of infection, we obtain the following values for the proportion susceptible and the average annual numbers of infections per 100,000 population in different age groups before the introduction of vaccination:

Age (years)	Proportion susceptible ($s(a) = e^{-\lambda a}$)	Average annual number of new infections per 100,000 ($\lambda s(a) \times 100000$)
15	0.03	705
25	0.003	68
35	(0)	7

c) If the herd immunity effects of vaccination are not accounted for, the proportion susceptible and the number of infections per 100,000 population in the long-term following the introduction of vaccination would be given by the values calculated in part b) but multiplied by $(1-v)$, where v is the proportion of individuals that are effectively vaccinated. v is given by the expression vaccine coverage \times vaccine efficacy. The values obtained assuming that the vaccine efficacy is 100% and 85% are provided below. These show that the static model greatly underestimates the long-term numbers of mumps infections per 100,000 population in 15-35 year olds following the introduction of vaccination.

Age (years)	100% vaccine efficacy		85% vaccine efficacy	
	Proportion susceptible	Average annual number of new infections per 100,000	Proportion susceptible	Average annual number of new infections per 100,000
15	0.012	282	0.015	345
25	0.001	27	0.001	33
35	0	3	0	3

5.10 a) Multiplying both sides of equation 5.29 by $(1-v)$, we obtain the following:

$$R_0(1-v) = \lambda' L + 1$$

Subtracting both sides of this equation by 1 and dividing by L , we obtain the following equation:

$$\frac{R_0(1-v) - 1}{L} = \lambda'$$

S5.11

b) and d) Figure S5.5 compares the plot of $\lambda' = \frac{R_0(1-v)-1}{L}$ against that of $\lambda' = \lambda(1-v)$.

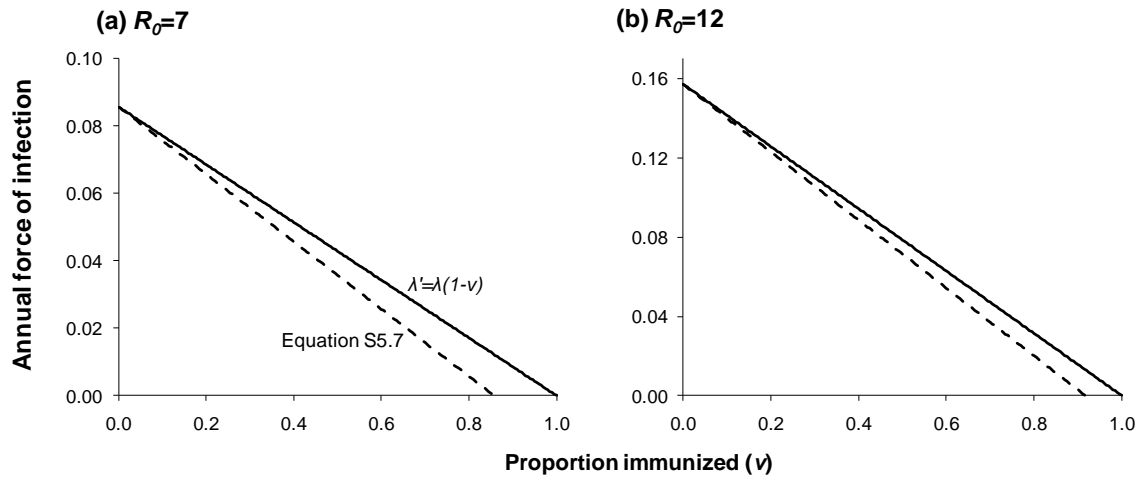


Figure S5.5: Predictions of the average annual long-term force of infection following the introduction of vaccination, calculated using $\lambda' = \frac{R_0(1-v)-1}{L}$ (dotted line) for different levels of the immunization coverage among newborns, in a A. low transmission ($R_0=7$) and B. a high transmission setting ($R_0=12$). The solid line shows the annual force of infection which would be seen if the force of infection was directly proportional to the proportion of individuals that are protected by vaccination(v), i.e. if $\lambda' = \lambda(1-v)$.

c) You might have expected the force of infection as predicted by the equation $\lambda' = \lambda(1-v)$ to decrease more slowly with increased vaccination coverage than that predicted by the line $\frac{R_0(1-v)-1}{L} = \lambda'$, since the gradient of the latter line is steeper than that for $\lambda' = \lambda(1-v)$.

For example, recall that the gradient of the line $\frac{R_0(1-v)-1}{L} = \lambda'$ is the factor by which we multiply “ v ” (i.e. the “coefficient” of v). The coefficient of v in this equation, and therefore the gradient of the line is $-R_0/L$. Substituting for $R_0=1+\lambda L$ (equation 5.21) into this equation, we see that the coefficient is equal to:

$$-\frac{R_0}{L} = -\frac{1+\lambda L}{L}$$

The expression for the gradient simplifies to the following:

$$-\left(\lambda + \frac{1}{L}\right)$$

In contrast, the gradient of the line $\lambda' = \lambda(1-v)$ is just $-\lambda$.

Since $\lambda + \frac{1}{L}$ is bigger than λ , we can conclude that the gradient of the line $\lambda' = \lambda(1 - v)$ will be steeper than that given by equation 5.11.

e) You should find that it is not possible to rearrange the equation to obtain an explicit expression for λ' in terms of R_0 , L and v , since the numerator has a term λ' and the denominator has a term $e^{-\lambda'L}$.

Instead, we need to use iterative techniques to obtain the value for λ' which results from a given value for R_0 , v and L , as follows:

We first rearrange the equation $R_0 = \frac{\lambda' L}{(1 - v)(1 - e^{-\lambda' L})}$ so that we have an expression for λ' in terms of all the other terms in the equation. For example, we could rearrange the equation to obtain the following:

$$\lambda' = \frac{R_0(1 - v)(1 - e^{-\lambda' L})}{L} \quad \text{S5.12}$$

If we substitute some value for λ' (denoted by λ'_0) into the right-hand side of this equation, then for given values for R_0 , L and v , we will obtain another value for λ' (we shall denote it by λ'_1). If we then substitute λ'_1 into the right-hand side of equation 5.12, we obtain another value for λ' (we shall denote it by λ'_2). Repeating this process several times, we eventually obtain a series of values, $\lambda'_0, \lambda'_1, \lambda'_2, \lambda'_3, \dots$, and we find that the difference between successive values of λ'_i becomes progressively smaller, until the value obtained satisfies equation S5.12 (see Table S5.1). These iterations can be set up in a spreadsheet.

Table S5.1: Illustration of how the post-vaccination force of infection at equilibrium, λ' , which satisfies the equation $R_0 = \frac{\lambda' L}{(1-v)(1-e^{-\lambda' L})}$ may be calculated iteratively from

the equation $\lambda'_{i+1} = \frac{R_0(1-v)(1-e^{-\lambda'_i L})}{L}$, assuming that $R_0=7$, $L=70$ years and $v=0.1$. In

this instance, the average force of infection which might be expected if the vaccination coverage among newborns is 10% is 0.090 per year.

Iteration number	λ'_i (per year)	Value for $\frac{R_0(1-v)(1-e^{-\lambda'_i L})}{L}$
0	$\lambda'_0 = 0.05$	$\frac{7(1-0.1)(1-e^{-0.05 \times 70})}{70} = 0.2$ per year.
1	$\lambda'_1 = 0.2$	$\frac{7(1-0.1)(1-e^{-0.2 \times 70})}{70} = 0.0899999$ per year
2	$\lambda'_2 = 0.0899999$	$\frac{7(1-0.1)(1-e^{-0.0899999 \times 70})}{70} = 0.0898347$ per year
3	$\lambda'_3 = 0.0898328$	$\frac{7(1-0.1)(1-e^{-0.0898328 \times 70})}{70} = 0.0898328$ per year

References

1. Nessa A, Islam MN, Tabassum S, Munshi SU, Ahmed M, Karim R. Seroprevalence of rubella among urban and rural Bangladeshi women emphasises the need for rubella vaccination of pre-pubertal girls. Indian J Med Microbiol 2008; 26(1):94-95.
2. World Population Prospects: The 2008 Revision and World Urbanization Prospects: The 2008 Revision. Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat . 2005.
3. Zhang YX. A compound catalytic model with both reversible and two-stage types and its applications in epidemiological study. Int J Epidemiol 1987; 16(4):619-621.
4. Kim-Farley R, Bart S, Stetler H et al. Clinical mumps vaccine efficacy. Am J Epidemiol 1985; 121(4):593-597.
5. Harling R, White JM, Ramsay ME, Macsween KF, van den BC. The effectiveness of the mumps component of the MMR vaccine: a case control study. Vaccine 2005; 23(31):4070-4074.
6. Cohen C, White JM, Savage EJ et al. Vaccine effectiveness estimates, 2004-2005 mumps outbreak, England. Emerg Infect Dis 2007; 13(1):12-17.

Chapter 7 (Solutions)

How do models deal with contact patterns?

7.1 Using a similar approach to that used in Panel 7.1, we obtain the following two expressions for the number of new infections among adults which are attributable to contact with children:

$$\beta_{oy} S_o(t) I_y(t) \quad \text{S7.1}$$

$$\lambda_{oy}(t) S_o(t) \quad \text{S7.2}$$

Equating expressions S7.1 and S7.2, we obtain the following:

$$\lambda_{oy}(t) S_o(t) = \beta_{oy} S_o(t) I_y(t)$$

Cancelling $S_o(t)$ from both sides of this equation, we obtain the following equation:

$$\lambda_{oy}(t) = \beta_{oy} I_y(t) \quad \text{S7.3}$$

Similarly, we can obtain the following two expressions for the number of new infections among adults that are attributable to contact with other adults:

$$\beta_{oo} S_o(t) I_o(t) \quad \text{S7.4}$$

$$\lambda_{oo}(t) S_o(t) \quad \text{S7.5}$$

Equating expressions S7.4 and S7.5 and cancelling $S_o(t)$ from the resulting equation, we obtain the following equation:

$$\lambda_{oo}(t) = \beta_{oo} I_o(t) \quad \text{S7.6}$$

Substituting our expressions for $\lambda_{oy}(t) = \beta_{oy} I_y(t)$ and $\lambda_{oo}(t) = \beta_{oo} I_o(t)$ into equation 7.3 in the book, we obtain our intended result, i.e.

$$\overline{\lambda_o(t)} = \beta_{oy} I_y(t) + \beta_{oo} I_o(t)$$

7.2 We will use the notation and definitions for the symbols provided on pages 183 and 184, and set $\beta_{yy}=2\times 10^{-4}$ per day, $\beta_{yo}=8\times 10^{-4}$ per day, $\beta_{oy}=3\times 10^{-4}$ per day, and $\beta_{oo}=7\times 10^{-5}$ per day. The answers to the questions are as follows:

- a) i) $\lambda_{yy}(t) = \beta_{yy} I_y(t) = 2 \times 10^{-4} \times 20 = 0.004$ per day.
 ii) $\lambda_{yo}(t) = \beta_{yo} I_o(t) = 8 \times 10^{-4} \times 50 = 0.04$ per day.
 iii) $\overline{\lambda_y(t)} = \lambda_{yy}(t) + \lambda_{yo}(t) = 0.004 + 0.04 = 0.044$ per day.
- b) i) $\lambda_{oy}(t) = \beta_{oy} I_y(t) = 3 \times 10^{-4} \times 20 = 0.006$ per day.
 ii) $\lambda_{oo}(t) = \beta_{oo} I_o(t) = 7 \times 10^{-5} \times 50 = 0.035$ per day.
 iii) $\overline{\lambda_o(t)} = \lambda_{oy}(t) + \lambda_{oo}(t) = 0.006 + 0.035 = 0.095$ per day.

7.3 Using WAIFW matrix $\begin{pmatrix} \beta_1 & 0.5\beta_2 \\ 0.5\beta_2 & \beta_2 \end{pmatrix}$, and assuming that the average force of infection among children and adults is 13% and 4% per year respectively, and that the average numbers of infectious children and adults are 18,956 and 2,859 respectively, then we would need to solve the following matrix equation to obtain values for β_1 and β_2 :

$$\begin{pmatrix} \beta_1 & 0.5\beta_2 \\ 0.5\beta_2 & \beta_2 \end{pmatrix} \begin{pmatrix} 18,956 \\ 2,859 \end{pmatrix} = \begin{pmatrix} 0.13 \\ 0.04 \end{pmatrix}$$

This equation can be written out in full as follows:

$$18,956\beta_1 + 2,859 \times 0.5\beta_2 = 0.13 \quad \text{S7.7}$$

$$18,956 \times 0.5\beta_2 + 2,859 \times \beta_2 = 0.04 \quad \text{S7.8}$$

Equation S7.8 simplifies to the following:

$$12,237\beta_2 = 0.04 \text{ per year.}$$

Dividing both side of this equation by 12,237 leads to $\beta_2 = 3.24 \times 10^{-6}$ per year.

Substituting this value for β_2 into equation S7.7 leads to the following:

$$18,956\beta_1 + 2,859 \times 0.5 \times 3.24 \times 10^{-6} = 0.13$$

This equation can be rearranged to give the following:

$$18,956\beta_1 = 0.12536 \quad \text{S7.9}$$

Dividing both sides of this equation by 18,956, we obtain $\beta_2 = 6.61 \times 10^{-6}$ per year. Dividing the values obtained for β_1 and β_2 by 365 to obtain values in units of per day leads to the following values: $\beta_1 = 1.81 \times 10^{-8}$ per day and $\beta_2 = 8.88 \times 10^{-9}$ per day.

Substituting these values for β_1 and β_2 into the matrix $\begin{pmatrix} \beta_1 & 0.5\beta_2 \\ 0.5\beta_2 & \beta_2 \end{pmatrix}$ leads to the

following matrix $\begin{pmatrix} 1.81 \times 10^{-8} & 4.44 \times 10^{-9} \\ 4.44 \times 10^{-9} & 8.88 \times 10^{-9} \end{pmatrix}$ in units of per day. This matrix is identical to matrix R2 that is presented in section 7.4.2.1.1.

7.4 a) We will use the equation $I_i = \lambda_i S_i D$, to calculate the number of infectious persons in age group i , where λ_i and S_i are the force of infection and the susceptible infectious persons in age group i respectively, and D is the duration of infectiousness ($=7$ days or $7/365$ years). The age groups 0-1, 2-4, 5-9, 10-14 and 15-74 years will be denoted using the subscripts 1, 2, 3, 4 and 5 respectively.

The following table summarizes the average numbers of infectious individuals calculated for each age group:

Age group (years)	0-1	2-4	5-9	10-14	15-74
Average prevaccination force of infection (%/year)	7.7	23.7	51.7	25.5	9.9
Average number susceptible (prevaccination)	1,062,861	1,216,541	493,355	59,269	59,182
Average number of infectious persons	$I_1=1,570$ ($=0.077$ \times 1,062,861 \times 7/365)	$I_2=5,529$ ($=0.237$ \times 1,216,541 \times 7/365)	$I_3=4,892$ ($=0.517$ \times 493,355 \times 7/365)	$I_4=290$ ($=0.255$ \times 59,269 \times 7/365)	$I_5=112$ ($=0.099$ \times 59,182 \times 7/365)

Note that when applying the equation $I_i = \lambda_i S_i D$, the units for the force of infection in age group i and the duration of infectiousness must be consistent: since we used an annual force of infection, the duration of infectiousness used in the equation is also in units of years.

b) i) After some calculations (see below) and assuming that $\alpha=1$, we obtain the following values for the β parameters for matrix 2:

$$\beta_1 = 6.20 \times 10^{-6} \text{ per year}$$

$$\beta_2 = 2.11 \times 10^{-5} \text{ per year}$$

$$\beta_3 = 7.84 \times 10^{-5} \text{ per year}$$

$$\beta_4 = 2.14 \times 10^{-5} \text{ per year}$$

$$\beta_5 = 7.99 \times 10^{-6} \text{ per year}$$

The WAIFW matrix is therefore as follows (where the parameters are in units of per year):

$$\begin{pmatrix} 6.20 \times 10^{-6} & 6.20 \times 10^{-6} & 6.20 \times 10^{-6} & 6.20 \times 10^{-6} & 7.99 \times 10^{-6} \\ 6.20 \times 10^{-6} & 2.11 \times 10^{-5} & 2.11 \times 10^{-5} & 2.11 \times 10^{-5} & 7.99 \times 10^{-6} \\ 6.20 \times 10^{-6} & 2.11 \times 10^{-5} & 7.84 \times 10^{-5} & 2.14 \times 10^{-5} & 7.99 \times 10^{-6} \\ 6.20 \times 10^{-6} & 2.11 \times 10^{-5} & 2.14 \times 10^{-5} & 7.84 \times 10^{-5} & 7.99 \times 10^{-6} \\ 7.99 \times 10^{-6} & 7.99 \times 10^{-6} & 7.99 \times 10^{-6} & 7.99 \times 10^{-6} & 7.99 \times 10^{-6} \end{pmatrix}$$

The equation that we need to solve to obtain these values for the β parameters is as follows:

$$\begin{pmatrix} \beta_1 & \beta_1 & \beta_1 & \beta_1 & \beta_5 \\ \beta_1 & \beta_2 & \beta_2 & \beta_2 & \beta_5 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 \\ \beta_1 & \beta_2 & \beta_4 & \alpha\beta_3 & \beta_5 \\ \beta_5 & \beta_5 & \beta_5 & \beta_5 & \beta_5 \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \\ I_5 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \end{pmatrix}$$

These equations can be rewritten as follows:

$$\beta_1(I_1 + I_2 + I_3 + I_4) + \beta_5 I_5 = \lambda_1 \quad \text{S7.10}$$

$$\beta_1 I_1 + \beta_2(I_2 + I_3 + I_4) + \beta_5 I_5 = \lambda_2 \quad \text{S7.11}$$

$$\beta_1 I_1 + \beta_2 I_2 + \beta_3 I_3 + \beta_4 I_4 + \beta_5 I_5 = \lambda_3 \quad \text{S7.12}$$

$$\beta_1 I_1 + \beta_2 I_2 + \beta_4 I_3 + \alpha\beta_3 I_4 + \beta_5 I_5 = \lambda_4 \quad \text{S7.13}$$

$$\beta_5(I_1 + I_2 + I_3 + I_4 + I_5) = \lambda_5 \quad \text{S7.14}$$

Equation S7.14 can be rearranged to give the following:

$$\beta_5 = \frac{\lambda_5}{I_1 + I_2 + I_3 + I_4 + I_5}$$

Substituting for $\lambda_5=0.099$ per year and for $I_1 + I_2 + I_3 + I_4 + I_5 = 12,393$ into this equation, we obtain the following value for β_5 :

$$\beta_5 = 0.099/12,393 = 7.99 \times 10^{-6} \text{ per year}$$

Equation S7.10 can be rearranged to give the following expression for β_1 :

$$\beta_1 = \frac{\lambda_1 - \beta_5 I_5}{I_1 + I_2 + I_3 + I_4}$$

Substituting for $\beta_5 = 7.99 \times 10^{-6}$ per year, $\lambda_1=0.237$ per year and for the corresponding numbers of infectious persons into this equation, we obtain $\beta_1 = 6.20 \times 10^{-6}$ per year.

Equation S7.11 can be rearranged to give the following expression for β_2 :

$$\beta_2 = \frac{\lambda_2 - \beta_1 I_1 - \beta_5 I_5}{I_2 + I_3 + I_4}$$

Substituting for $\beta_5 = 7.99 \times 10^{-6}$ per year, $\beta_1 = 6.20 \times 10^{-6}$ per year and for the corresponding numbers of infectious persons into this equation, we obtain $\beta_2 = 2.11 \times 10^{-5}$ per year.

To obtain β_3 and β_4 , we can solve equations S7.12 and S7.13 simultaneously. To simplify the notation, we will re-express these two equations as follows:

$$\beta_3 I_3 + \beta_4 I_4 = P \quad \text{S7.12'}$$

$$\alpha \beta_3 I_4 + \beta_4 I_3 = Q \quad \text{S7.13'}$$

where $P = \lambda_3 - \beta_1 I_1 - \beta_2 I_2 - \beta_5 I_5$ and $Q = \lambda_4 - \beta_1 I_1 - \beta_2 I_2 - \beta_5 I_5$. Substituting the corresponding values for the force of infection, the β values and the numbers of infectious persons into these equations, we see that $P=0.3895$ and $Q=0.1265$ (to 4 decimal places).

Multiplying equations S7.12' and S7.13' by I_3 and I_4 respectively, we obtain the following:

$$\beta_3 I_3^2 + \beta_4 I_4 I_3 = P I_3 \quad \text{S7.15}$$

$$\alpha \beta_3 I_4^2 + \beta_4 I_3 I_4 = Q I_4 \quad \text{S7.16}$$

Subtracting equation S7.16 from equation S7.15, we obtain the following equation.

$$\beta_3 (I_3^2 - \alpha I_4^2) = P I_3 - Q I_4$$

After rearranging, we obtain the following expression for β_3 :

$$\beta_3 = \frac{P I_3 - Q I_4}{I_3^2 - \alpha I_4^2}$$

Substituting for P , Q , I_3 and I_4 into this equation leads to the result that $\beta_3 = 7.84 \times 10^{-5}$ per year.

Rearranging equation S7.12', we obtain the following expression for β_4 :

$$\beta_4 = \frac{P - \beta_3 I_3}{I_4}$$

Substituting for P , I_3 , I_4 and β_3 into this equation, we obtain $\beta_4 = 2.14 \times 10^{-5}$ per year.

ii) To calculate R_0 , we first need to calculate the Next Generation Matrix, using the number of infectious individuals among those in age group i resulting from individuals in age group j , as obtained using the expression $N_i \beta_{ij} D$. Here, N_i is the number of individuals in age group i , and β_{ij} is the rate at which specific susceptible individuals in age group i come into effective contact with specific infectious individuals in age group j and $D (=7 \text{ days})$ is the duration of infectiousness.

N_i is given by the width of age group i multiplied by 650,000 (the number of individuals in each single year age category), as follows:

Age group (years)	0-1*	2-4	5-9	10-14	15-74
N_i	1,137,500 (= 1.75×650,000)	1,950,000 (= 3×650,000)	3,250,000 (= 5×650,000)	3,250,000 (= 5×650,000)	39,000,000 (= 60×650,000)

* Note that calculations for 0-1 year olds assume that individuals have maternal immunity for the first 3 months of life.

The Next Generation Matrix should resemble the following:

Age grp (yrs)	0-1	2-4	5-9	10-14	15-74
0-1	0.135	0.135	0.135	0.135	0.174
2-4	0.232	0.790	0.790	0.79	0.299
5-9	0.386	1.317	4.884	1.335	0.498
10-14	0.386	1.317	1.335	4.884	0.498
15-74	5.975	5.975	5.975	5.975	5.975

Adapting the model files provided for calculating R_0 , we obtain a value for R_0 of 9.1.

iii) The herd immunity threshold for this matrix is $1-1/R_0$ or $100 \times (1-1/9.1) \approx 89\%$

c) Increasing the size of α increases the amount of contact between 10-14 year olds and leads to an increase in the size of R_0 and the herd immunity threshold. For example, if $\alpha=2$, R_0 equals 11.4 and the herd immunity threshold is about 91%. Consequently, the greater the amount of contact between 10-14 year olds, the more difficult it is to control transmission through vaccination.

d) The following summarizes the values for R_n that are obtained for different values for α :

α	1	1.25	1.5	1.75	2
R_n	0.96	0.99	1.04	1.11	1.2

In general, the greater the value for α , the greater the value for R_n . The values obtained for R_n are generally consistent with those in Figure 7.17.

7.5 a) The Next Generation Matrix is given by the following: $\begin{pmatrix} 1.784 & 0.188 \\ 0.383 & 0.766 \end{pmatrix}$, which results in a value for R_0 of 1.85.

b) To answer this question, we calculate the reproduction number using the following Next Generation Matrix

$$\begin{pmatrix} \beta_{yy} S_y D & \beta_{yo} S_y D \\ \beta_{oy} S_o D & \beta_{oo} S_o D \end{pmatrix}$$

where S_y and S_o are the numbers of susceptible children and adults (defined as those aged <15 and ≥ 15 years respectively), calculated after incorporating the appropriate vaccination coverage for each vaccination scenario. The β parameters are as defined in the text, and D is the duration of infectiousness (2 days):

The following table summarizes the number of susceptible individuals for each vaccination scenario, the Next Generation Matrix and the values for the reproduction number:

Individuals targeted	Number of susceptible		Next Generation Matrix	Reproduction number
	Children (S_y)	Adults (S_o)		
No vaccination	2639	5361	$\begin{pmatrix} 1.784 & 0.188 \\ 0.383 & 0.766 \end{pmatrix}$	1.85 ($=R_o$)
Children only	139 (=2639-2500)	5361 (=5361-0)	$\begin{pmatrix} 0.094 & 0.010 \\ 0.383 & 0.766 \end{pmatrix}$	0.77
Adults only	2639 (=2639-0)	2861 (=5361-2500)	$\begin{pmatrix} 1.784 & 0.188 \\ 0.204 & 0.409 \end{pmatrix}$	1.81
Same proportion of children and adults*	1814 (=2639 \times (1-0.3125))	3686 (=5361 \times (1-0.3125))	$\begin{pmatrix} 1.226 & 0.130 \\ 0.263 & 0.526 \end{pmatrix}$	1.27
Equal numbers of children and adults	1389 (=2639-1250)	4111 (=5361-1250)	$\begin{pmatrix} 0.939 & 0.099 \\ 0.294 & 0.587 \end{pmatrix}$	1.01

*The proportion of children and adults that need to be targeted with this strategy equals the number of vaccine doses available \div population size = $2500/(2639+5361)=31.25\%$

The smallest value for the reproduction number is associated with the strategy of vaccinating only children, which suggests that, of the four strategies, this approach may be the best way of distributing the vaccine stocks. However, we would also need to account for the severity of influenza and the mortality rates in other age groups before making the final decision about which vaccination strategy should be adopted.

Basic maths (Solutions)

B1 a) $\log_{10} 1,000,000 = 6$

b) $\log_{10} 10,000 = 4$

c) $\log_{10} 0.001 = -3$

d) $\log_p r = q$

e) $10^5 = 100,000$

f) $q^p = r$

B2 a) $\frac{dy}{dt} = 28$

b) $\frac{dy}{dt} = 10te^{-5t} - 25t^2e^{-5t}$

c) $\frac{dy}{dt} = 3e^t$

d) $\frac{dy}{dt} = 8t^3 - 48t^7$

e) $\frac{dy}{dt} = -7t^{-8} + 2t$

B3 In the following expressions, c is some (unknown) constant.

a) $\int t^2 dt = \frac{t^3}{3} + c$

b) $\int 4t^3 dt = t^4 + c$

c) $\int \frac{1}{t^5} dt = -\frac{1}{4t^4} + c$

$$d) \int_0^{20} e^{-rt} dt = \left[-\frac{e^{-rt}}{r} \right]_0^{20} = \frac{1 - e^{-20r}}{r}$$

$$e) \int 70 dt = 70t + c$$

$$\mathbf{B4} \text{ a) i) } \begin{pmatrix} 8 & 2 \\ 2 & 13 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix} \quad \text{ii) } \begin{pmatrix} 2 & 6 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

$$b \text{ i) } \begin{aligned} 9x + y &= 1 \\ 3x + 2y &= 4 \end{aligned}$$

$$b \text{ ii) } \begin{aligned} 13x + 9y &= 6 \\ 4x + 4y &= 2 \end{aligned}$$

$$\mathbf{B5} \text{ a) The determinant is given by: } 5 \times 13 - 2 \times 10 = 65 - 20 = 45$$

$$b) \text{ The determinant is given by: } 3 \begin{vmatrix} 4 & 0 \\ 2 & 1 \end{vmatrix} + 1 \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix} + 0 \begin{vmatrix} 0 & 4 \\ 0 & 2 \end{vmatrix} = 3 \times (4 - 0) + 0 = 12$$

c) To find the conditions under which the equation $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ holds for some non-zero values of x and y , we first write out this equation using simultaneous equation notation:

$$ax + by = 0$$

$$cx + dy = 0$$

Multiplying the first equation by c and the second equation by a we obtain the following two equations:

$$acx + bcy = 0$$

$$acx + ady = 0$$

Subtracting the first equation from the second equation, we obtain the following result:

$$ady - bcy = 0$$

Factoring out y from this equation, we obtain the following equation:

$$(ad - bc)y = 0$$

Since y is non-zero, and since this equation equals zero, it follows that the term in brackets must equal zero, i.e. $ad - bc = 0$.

Since $ad - bc$ equals the determinant of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we have obtained our intended result.

B6 a) To obtain the eigenvalues of $\begin{pmatrix} 5 & 2 \\ 10 & 13 \end{pmatrix}$, we need to find the values of ρ for which the determinant of the matrix $\begin{pmatrix} 5-\rho & 2 \\ 10 & 13-\rho \end{pmatrix}$ is zero.

The determinant of this matrix is given by the expression $(5-\rho)(13-\rho) - 20$.

This expression can be rearranged as follows:

$$\begin{aligned} (5-\rho)(13-\rho) - 20 &= 65 - 18\rho + \rho^2 - 20 \\ &= \rho^2 - 18\rho + 45 \\ &= (\rho-3)(\rho-15) \end{aligned}$$

Consequently, for the determinant to equal zero, the following equation must hold:

$$(\rho-3)(\rho-15) = 0$$

This holds when either $\rho=3$ or $\rho=15$, which suggests that the eigenvalues of the matrix

$\begin{pmatrix} 5 & 2 \\ 10 & 13 \end{pmatrix}$ are 3 and 15.

b) To find the eigenvalues of $\begin{pmatrix} 3 & 1 & 0 \\ 0 & 4 & 0 \\ 0 & 2 & 1 \end{pmatrix}$, we need to find the values of ρ for which the

determinant of the matrix $\begin{pmatrix} 3-\rho & 1 & 0 \\ 0 & 4-\rho & 0 \\ 0 & 2 & 1-\rho \end{pmatrix}$ is zero, i.e. for which the following

equation holds:

$$(3-\rho) \begin{vmatrix} 4-\rho & 0 \\ 2 & 1-\rho \end{vmatrix} + 1 \begin{vmatrix} 0 & 0 \\ 1-\rho & 0 \end{vmatrix} + 0 \begin{vmatrix} 0 & 4 \\ 0 & 1 \end{vmatrix} = 0$$

This equation simplifies to the following:

$$(3-\rho)(4-\rho)(1-\rho) = 0$$

For this equation to hold, ρ (i.e. the eigenvalues) must equal 3, 4 or 1.