Review of block 2: Analyses of seroprevalence data and their application to modelling control strategies

Emilia Vynnycky

Introduction to Infectious Disease Modelling and its Applications
25th June 2018

Analysis of seroprevalence data and their application to modelling control strategies

AIMS:

To illustrate
- how you might analyse seroprevalence data to obtain the force of infection, average age at infection, proportion susceptible and R_0
- an important use of modelling for determining control strategies

---

Next steps…

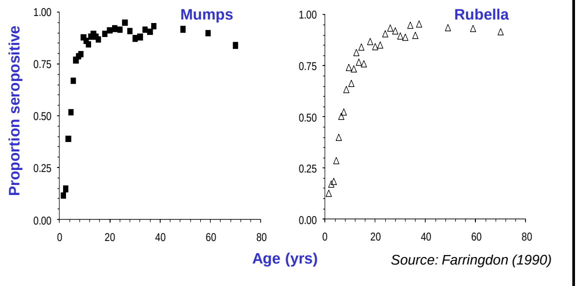
Session 8-9
Analyse infection prevalence data - estimate the average force of infection (λ) by fitting a catalytic model to data and calculate R_0 assuming random mixing

Session 10
Incorporate R_0 in age-structured model of rubella transmission, assuming random mixing and explore effect of vaccination on:
- age-specific proportion susceptible
- infection incidence

Session 11
Use λ to estimate age-dependent contact parameters
Incorporate these parameters in a model to explore effect of age-dependent contact on the impact of vaccination

Session 14
Calculate R_0 using age-dependent contact parameters

Observed proportion of individuals sero-positive to mumps and rubella antibodies in the UK (late 1980s)



Methods for analysing seroprevalence data

Seroprevalence data are typically analysed using **catalytic models** to estimate the **average force of infection,** which is then used to calculate:

- the average age at infection,
- the proportion susceptible
- R_0 and herd immunity threshold
- Infection incidence

Revision of catalytic models

For immunizing infections, the serological profile can be described using a simple catalytic model

Susceptible $s(a)$ $\xrightarrow{\lambda}$ (Ever) infected $z(a)$

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

$$\frac{dz}{da} = \lambda s(a)$$

Recall from differential equations notes/previous mathematical knowledge, these equations can be solved to give:

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

$z(a)$ = 1- proportion susceptible = $1 - e^{-\lambda a}$

What is the difference between a catalytic model and a transmission model?

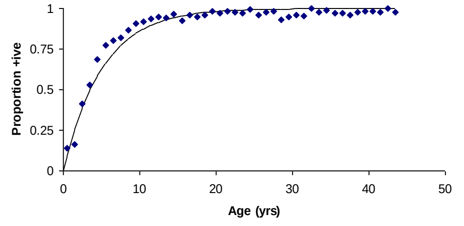
In transmission models, the force of infection is expressed in terms of the number of infectious individuals in the model, which changes over time, i.e.

$$\lambda(t) = \beta I(t)$$

Catalytic models do not explicitly describe transmission between individuals in the model. The force of infection is taken to be some value which is independent of the size of other compartments in the model

$$\boxed{\text{Susceptible } s(a)} \xrightarrow{\lambda} \boxed{\text{(Ever) infected } z(a)}$$

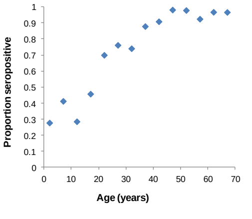
Comparison between the observed proportion of individuals found to be positive to mumps in the UK, and that predicted using a simple catalytic model



The extent to which the force of infection depends on age can be assessed using graphical methods.

Q1. The following table and figure show data on the age-specific proportion of individuals who were seropositive for hepatitis A in country Z. The life expectancy in the population was about 70 years and the age distribution was rectangular.

Age mid point	Proportion positive	Age mid point	Proportion positive
2	0.278	37	0.879
7	0.413	42	0.909
12	0.286	47	0.982
17	0.458	52	0.979
22	0.7	57	0.925
27	0.762	62	0.968
32	0.741	67	0.967



Which of the following statements is incorrect?

a) The force of infection was probably age-dependent.

b) The force of infection was probably not age-dependent.

c) The average force of infection was about 20%/year.

d) Assuming that individuals mix randomly, 20-30% of the population was probably susceptible.

e) Assuming that individuals mix randomly, we would need to attain a vaccination coverage of at least 70-80% in the population to control transmission

Which of the following statements is correct?

a) The force of infection was probably age-dependent.
b) The force of infection was probably not age-dependent.

Approach

1. Calculate -ln(observed proportion seronegative)
2. Plot -ln(observed proportion seronegative) against the age mid-point
3. If all the points fall on a straight line then the force of infection is probably not age-dependent; otherwise, it is probably age-dependent

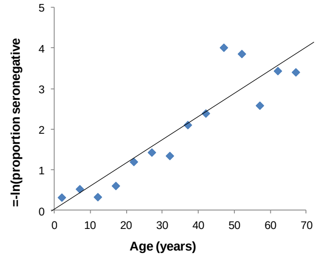
Calculating -ln(observed proportion seronegative) (1)

Age mid point	Proportion positive (p_a)	negative ($1-p_a$)	-LN(propn sero - negative) ($-\ln(1-p_a)$)	Age mid point	Proportion positive (p_a)	Negative ($1-p_a$)	-LN(propn sero-negative) ($-\ln(1-p_a)$)
2	0.278	0.722		37	0.879	0.121	
7	0.413	0.587		42	0.909	0.091	
12	0.286	0.714		47	0.982	0.018	
17	0.458	0.542		52	0.979	0.021	
22	0.7	0.3		57	0.925	0.075	
27	0.762	0.238		62	0.968	0.032	
32	0.741	0.259		67	0.967	0.033	

Calculating -ln(observed proportion seronegative) (2)

Age mid point	Proportion positive (p_a)	negative ($1-p_a$)	-LN(propn sero - negative) ($-\ln(1-p_a)$)	Age mid point	Proportion positive (p_a)	Negative ($1-p_a$)	-LN(propn sero-negative) ($-\ln(1-p_a)$)
2	0.278	0.722	0.32573	37	0.879	0.121	2.111965
7	0.413	0.587	0.53273	42	0.909	0.091	2.396896
12	0.286	0.714	0.336872	47	0.982	0.018	4.017384
17	0.458	0.542	0.612489	52	0.979	0.021	3.863233
22	0.7	0.3	1.203973	57	0.925	0.075	2.590267
27	0.762	0.238	1.435485	62	0.968	0.032	3.442019
32	0.741	0.259	1.350927	67	0.967	0.033	3.411248

Plot of -ln(observed proportion seronegative) for hepatitis in country Z)



Which of the following statements is correct?

a) The force of infection was probably age-dependent.
b) The force of infection was probably not age-dependent.

Is the following statement correct?

c) The average force of infection was about 20%/year.

Answer (APPROACH 1)

The force of infection is approximately equal to the gradient of the line passing through the points of -ln(observed proportion seronegative)

The gradient of the line ≈2/35 , i.e. the line has gone up by 2 units by age 35 years
∴ The force of infection ≈2/35 ≈0.057 or 5.7%/year

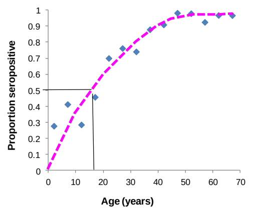
Therefore, statement c) is incorrect

Is the following statement correct?

c) The average force of infection was about 20%/year.

Answer (APPROACH 2)

The force of infection is approximately equal to 1/(average age at infection)



The median age at infection is approximately 17 years

- to see this read off the age by which 50% are seropositive, drawing a curve (if necessary) through the observed seropositive to guide your estimate

∴The force of infection ≈1/17 ≈0.06 or 6%/year

Therefore, statement c) is incorrect

BUT…this estimate is approximate since the median age at infection is an approximation to the "average" statistic that we need; it is slightly less accurate than the estimate from approach 1

Is the following statement correct?

d) Assuming that individuals mix randomly, 20-30% of the population was probably susceptible.

Answer

The force of infection is about 5.7%/year (depending on rounding)

⇒average age at infection ≈ 1/(force of infection)
≈ 1/0.057 ≈ 17.5 years

Assuming that the population has a rectangular age distribution and life expectancy of 70 years, means that the average proportion susceptible ≈ A/L ≈ 17.5/70 ≈ 0.25 or 25%

Therefore, statement d) is correct

Note: using A≈17 years (from approach 2) would have resulted in average proportion susceptible ≈ A/L ≈ 17/70 ≈ 0.24 or 24%

Is the following statement correct?

e) Assuming that individuals mix randomly, we would need to immunize at least 70-80% of the population to control transmission

Answer

To control transmission, we could need a vaccination coverage of above the herd immunity threshold = $1-1/R_0$

R_0 can be calculated from the equation R_0=1/average proportion susceptible

∴R_0 ≈ 1/0.25 =4
∴The herd immunity threshold =1-1/4 ≈ 0.75 or 75%

Therefore, statement e) is correct

Which of the following statements is incorrect?

True statement? | Correct answer

a) The force of infection was probably age-dependent.

b) The force of infection was probably not age-dependent.

c) The average force of infection was about 20%/year.

d) Assuming that individuals mix randomly, 20-30% of the population was probably susceptible.

e) Assuming that individuals mix randomly, we would need to immunize at least 70-80% of the population to control transmission

Why should a plot of $-ln\{s_a\}$ be linear if the force of infection is not age-dependent

If the force of infection is not age-dependent, then the proportion of individuals of age a who are susceptible is given by:

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

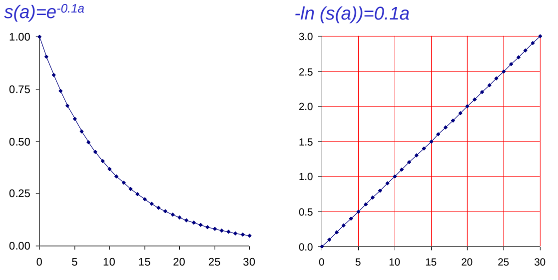
Taking natural logs of both sides of this equation, we see that

$$\ln\{s(a)\} = -\lambda a$$

or, equivalently, $-\ln\{s(a)\} = \lambda a$

i.e. the natural log of the proportion susceptible is linearly related to the force of infection

Illustration of the relationship between s(a) and $-\log_e(s(a))$



Age (yrs)

=> a graphical check applied to the data may be used to identify whether the force of infection is constant with age

Method for assessing age-dependency in the force of infection (2)

1. Estimate s(a), the proportion susceptible in an age group a as S_a/N_a
2. Calculate the values $-ln(S_a/N_a)$
3. Plot these values against the midpoints of the age groups

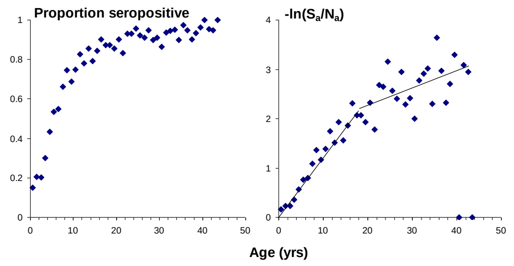
If the force of infection is constant with age,

the plot should approximate to a straight line;

the slope of the line is λ

Otherwise, the force of infection should be assumed to be age-dependent

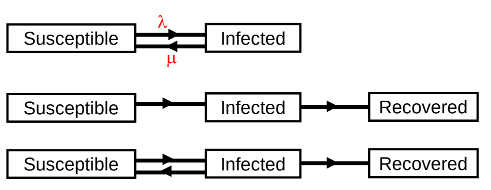
Exercise: What do the following figures suggest about the dependency of the force of infection on age?



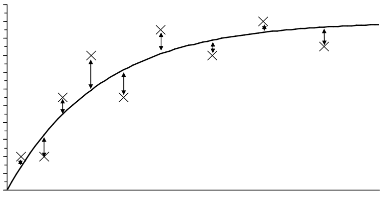
Fine-tuning catalytic models

Catalytic models can be fine-tuned to deal with:

- Maternal immunity
- Age or time-dependency in the force of infection
- Different kinds of infections:



Formal methods for estimating parameters



Minimize eg $\sum_i (O_i - E_i)^2$ Least squares

or $\sum_i \frac{(O_i - E_i)^2}{E_i}$ χ^2 statistic

Widely used: "Maximum likelihood"

Practicalities of fitting catalytic models to data

Key steps:

- Set up the expression for the age-specific prevalence of infection (if possible)
- Setting up expressions for the goodness of fit (e.g. "loglikelihood" deviance)
- Use optimization macros (eg Excel's solver) to minimize the goodness of fit and obtain unknown parameter(s)

Note:

- If it's impossible to write down an explicit expression for the proportion infected, the model can be set up using difference equations in Excel and fitted in the same way.
- Berkeley Madonna also has a curve-fitting routine...

Session 10: Modelling the impact of rubella vaccination strategies

- Vaccination => ↓ in prevalence of infectious persons (as those who would have become infected and infectious are now immune)

  => ↓ opportunity for infection

  => ↑ average age at infection

- This may have important consequences for infections for which infection late in life is associated with adverse outcomes (eg rubella and CRS)
- ↑ in average age at infection depends on setting
- The size of the increases is difficult to predict without a model...

Q2. Country Z has recently introduced rubella vaccination among very young children. If we assume that individuals mix randomly and that the vaccination coverage is below the herd immunity threshold, which of the following statements is likely to be true?

a) The average age at rubella infection is likely to increase.

b) The average age at rubella infection is likely to remain unchanged.

c) The overall proportion of the population that is susceptible may remain unchanged

d) The overall proportion of the population that is susceptible will decrease

e) The overall proportion of the population that is susceptible will increase

Q2. Country Z has recently introduced rubella vaccination among very young children. If we assume that individuals mix randomly and that the vaccination coverage is below the herd immunity threshold, which of the following statements is likely to be true?

a) The average age at rubella infection is likely to increase.

b) The average age at rubella infection is likely to remain unchanged.

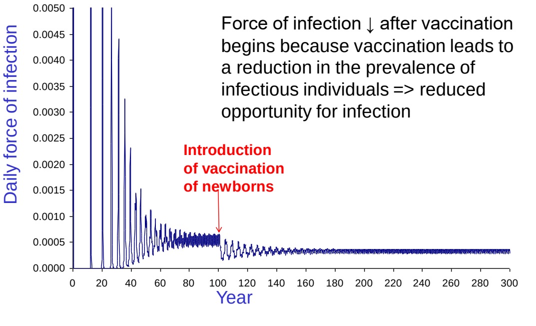
Answer

The force of infection (λ) decreases after the introduction of vaccination

Since $A \approx 1/\lambda$, if λ decreases, A increases

So a) is most likely to be correct

Model predictions of the effect of 40% vaccination coverage among newborns, introduced 100 years after the start of model simulations, on the force of infection in China



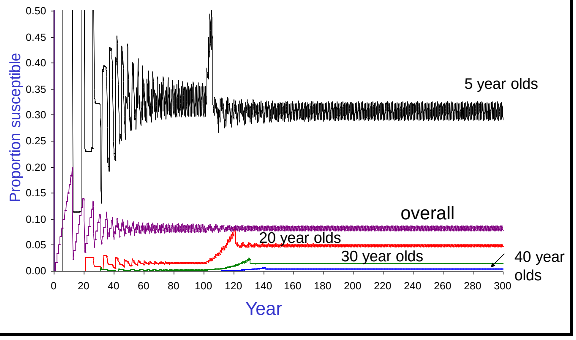
Q2. Country Z has recently introduced rubella vaccination among very young children. If we assume that individuals mix randomly and that the vaccination coverage is below the herd immunity threshold, which of the following statements is likely to be true?

c) The overall proportion of the population that is susceptible may remain unchanged

d) The overall proportion of the population that is susceptible will decrease

e) The overall proportion of the population that is susceptible will increase

Why does the overall proportion of individuals who are susceptible remain unchanged following the introduction of infant vaccination (vaccination coverage<herd immunity threshold)?



Why does the overall proportion of individuals who are susceptible remain unchanged following the introduction of infant vaccination (vaccination coverage<herd immunity threshold)?

Recall net reproduction number (R_n) is related to R_0 and proportion susceptible (s) through the equation:

$$R_n = R_0 s$$

If vaccination coverage < herd immunity threshold, the infection is still endemic, so

$$R_n = 1$$

Since $R_n = R_0 s = 1$, and R_0 is constant, this implies that

$$s = 1/R_0$$

∴ The overall proportion susceptible remains unchanged if the vaccination coverage is below the herd immunity threshold

Q2. Country Z has recently introduced rubella vaccination among very young children. If we assume that individuals mix randomly and that the vaccination coverage is below the herd immunity threshold, which of the following statements is likely to be true?

c) The overall proportion of the population that is susceptible may remain unchanged

d) The overall proportion of the population that is susceptible will decrease

e) The overall proportion of the population that is susceptible will increase

Q2. Country Z has recently introduced rubella vaccination among very young children. If we assume that individuals mix randomly and that the vaccination coverage is below the herd immunity threshold, which of the following statements is likely to be true?

a) The average age at rubella infection is likely to increase.

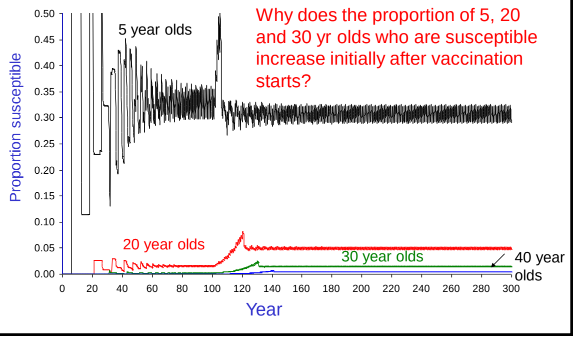
b) The average age at rubella infection is likely to remain unchanged.

c) The overall proportion of the population that is susceptible may remain unchanged

d) The overall proportion of the population that is susceptible will decrease

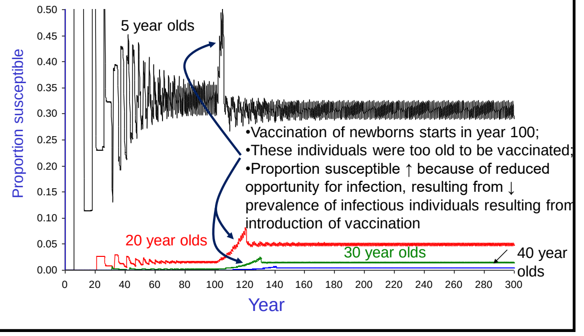
e) The overall proportion of the population that is susceptible will increase

Model predictions of the effect of vaccinating 40% of newborns annually from year 100 in China, on the proportion susceptible

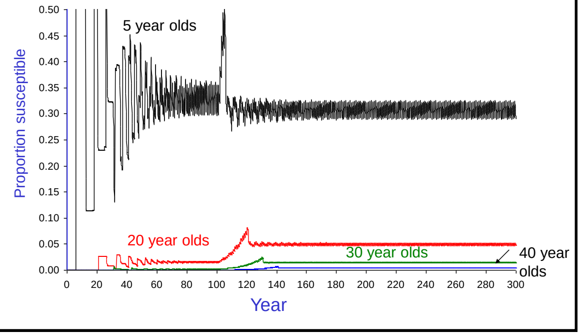


Why does the proportion of 5, 20 and 30 yr olds who are susceptible increase initially after vaccination starts?

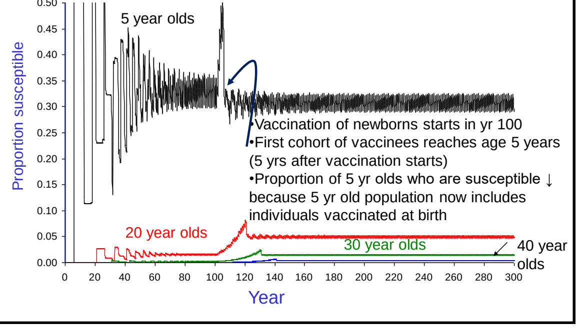
Why does the proportion of 5, 20 and 30 yr olds who are susceptible *increase* initially after starting vaccination of newborns?



Why does the proportion of 5 yr olds who are susceptible decrease 5 years following the introduction of vaccination of newborns?



Why does the proportion of 5 yr olds who are susceptible decrease 5 years following the introduction of vaccination of newborns?



Why is the proportion of adults who are susceptible to infection higher after infant vaccination is introduced than it was before vaccination was introduced? (Explanation 1)

- If the vaccination coverage is below the herd immunity threshold, the overall proportion of individuals in the population who are susceptible remains unchanged

- Note that the proportion of **children** who are susceptible **decreases**

∴The proportion of **adults** who are susceptible must **increase** for the overall proportion of individuals in the whole population who are susceptible to remain unchanged.

Why is the proportion of adults who are susceptible to infection higher following the introduction of infant vaccination than it was before vaccination was introduced? (Explanation 2)

The long-term proportion of individuals who are susceptible in any given age group depends on:

A. Direct effect of vaccination:

=> ↓ proportion susceptible

B. Indirect (herd immunity) effect of vaccination:

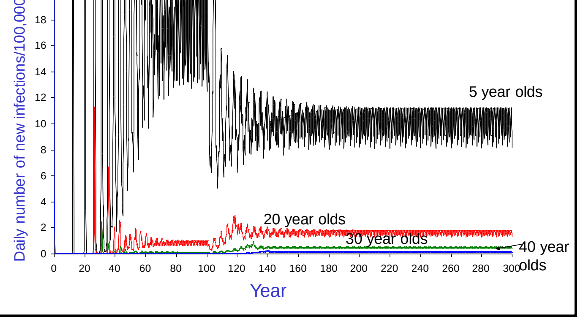
=> reduced force of infection

=> ↑ in proportion of unvaccinated individuals who are still susceptible by a given age

The relative size of these effects depends on age group

Effect of B is more obvious in the proportion of adults that are susceptible than in the proportion of children that are susceptible since adults will have experienced more years of exposure to a reduced force of infection than children

Why does the number of new infections/100,000/day in China decrease among 5 year olds after vaccination of newborns is introduced? Why does it increase for 20 year olds?



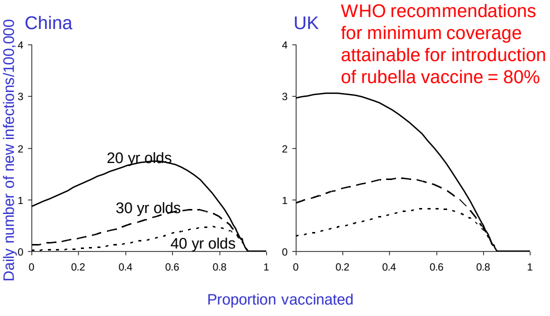
Why does the daily no. of new infections decrease among 5 year olds following the introduction of vaccination of newborns? Why does it increase for 20 year olds?

Number of new infections/unit time at age a = $\lambda(t)S(a)$

Following vaccination: ↓ ↓ 5 yr olds
↓ ↑ ≥20 yr olds

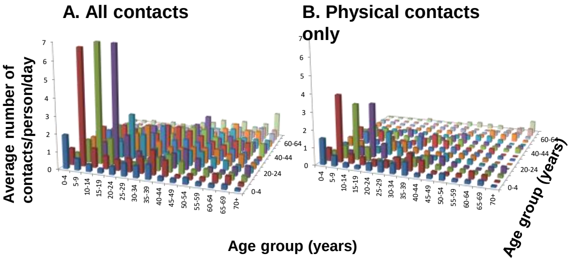
a) Daily no. of new infections **decreases** for 5 year olds since both the force of infection and the proportion susceptible go down with the introduction of vaccination
b) Daily no. of new infections **increases** for 20 year olds since the increase in proportion susceptible following the introduction of vaccination outweighs the reduction in the force of infection.
The size of the change depends on the vaccination coverage and R_0 (or λ) before vaccination starts and is difficult to predict without a model...

Model predictions of the impact of different levels of effective vaccination coverage among newborns on the long-term number of new infections/100,000/day in China and the UK



Modelling the effects of non-random mixing on the transmission dynamics and control of infections

Evidence for an age-dependency in transmission: Data on average numbers of individuals contacted per day, May 2005-Sept 2006 in the UK



Source: Mossong et al (2008)

Expression for the force of infection: random mixing

Equation for the force of infection, assuming random mixing:

$$\lambda(t) = \beta I(t)$$

To account for non-random mixing, we need separate expressions for the force of infection for each subgroup in the population

Methods for incorporating heterogeneous mixing into models

For heterogeneously mixing populations (eg in which contact patterns differ between the young and old):

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

β_{yy} = rate at which a specific young individual comes into effective contact with a specific young individual per unit time

β_{yo} = rate at which a specific young (susceptible) individual comes into effective with a specific old (infectious) individual per unit time etc

$I_y(t)$ = number of infectious young individuals in the population

$I_o(t)$ = number of infectious old individuals in the population

Expressions for the force of infection using matrices

The equations: $\overline{\lambda_y(t)} = \beta_{yy} I_y(t) + \beta_{yo} I_o(t)$

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

can be summarized using the following matrix equations:

$$\begin{pmatrix} \overline{\lambda_y(t)} \\ \overline{\lambda_o(t)} \end{pmatrix} = \begin{pmatrix} \beta_{yy} & \beta_{yo} \\ \beta_{oy} & \beta_{oo} \end{pmatrix} \begin{pmatrix} I_y(t) \\ I_o(t) \end{pmatrix}$$

Matrix of "Who Acquires Infection From Whom" ("WAIFW")

β_{yy}, β_{yo} etc can be calculated given estimates of:

$$\overline{\lambda_y(t)},\ \overline{\lambda_o(t)},\ I_y(t),\ I_o(t)$$

and assumptions about the WAIFW matrix structure

Q3. The following WAIFW matrix describes contact between individuals in urban and rural areas. (Note that the letters u and r next to the rows and above the columns reflect urban and rural areas respectively.) Assuming that β_1 is not equal to β_2, which of the statements below is incorrect? (Note that there may be more than one incorrect answer).

$$\begin{array}{cc} & \begin{array}{cc} u & r \end{array} \\ \begin{array}{c} u \\ r \end{array} & \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_2 & \beta_2 \end{pmatrix} \end{array}$$

Assuming that β_1 is not equal to β_2, which of the statements below is incorrect? (Note that there may be more than one incorrect answer).

$$\begin{array}{cc} & \begin{array}{cc} u & r \end{array} \\ \begin{array}{c} u \\ r \end{array} & \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_2 & \beta_2 \end{pmatrix} \end{array}$$

a) Individuals in urban areas effectively contact each other at a different rate from the rate at which they effectively contact individuals in rural areas.

b) The rate at which individuals from rural areas contact each other is different from the rate at which individuals from urban areas contact each other.

c) Individuals from rural areas contact individuals from urban areas at the same rate at which they contact other individuals from rural areas.

d) The rate at which individuals from rural areas contact each other is equal to the rate at which individuals from urban areas contact each other.

e) Individuals from rural areas contact individuals from urban areas at a rate which is different from the rate at which they contact other individuals from rural areas.

$$\begin{array}{cc} & \begin{array}{cc} u & r \end{array} \\ \begin{array}{c} u \\ r \end{array} & \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_2 & \beta_2 \end{pmatrix} \end{array}$$

a) Individuals in urban areas effectively contact each other at a different rate from the rate at which they effectively contact individuals in rural areas.

Answer:
Correct, because individuals in urban areas effectively contact each other at a rate β_1, which is different from the rate at which they effectively contact others from rural areas. which equals β_2

Notation for subscripts for $\beta_{yy}, \beta_{yo}, \beta_{oo}, \beta_{oy}$

$$\beta_{yo}$$

Denotes category of the infectious person

Denotes category of the susceptible person (or the recipient of the infection)

Therefore:

β_{yo} = rate at which a specific young (susceptible) individual comes into effective with a specific old (infectious) individual per unit time

Possible constraints for WAIFW matrices

1. Symmetrical contact: rate at which a child contacts and infects an adult = rate at which an adult contacts and infects a child

i.e. $\beta_{yo} = \beta_{oy}$

2. Rate at which an adult contacts and infects a child = rate at which an adult contacts and infects an adult $\begin{pmatrix} \beta_1 & \beta_2 \\ \beta_2 & \beta_2 \end{pmatrix}$

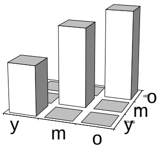
$\begin{pmatrix} \beta_1 & \beta_1 \\ \beta_1 & \beta_2 \end{pmatrix}$ $\begin{pmatrix} \beta_1 & \beta_2 \\ \beta_2 & \beta_1 \end{pmatrix}$ $\begin{pmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{pmatrix}$ Symmetrical

$\begin{pmatrix} 0 & \beta_2 \\ \beta_1 & 0 \end{pmatrix}$ $\begin{pmatrix} \beta_1 & \beta_2 \\ \beta_1 & \beta_2 \end{pmatrix}$ $\begin{pmatrix} \beta_1 & \beta_1 \\ \beta_2 & \beta_2 \end{pmatrix}$ Asymmetrical

Summary of WAIFW matrices calculated in practical

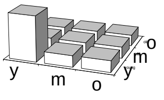
WAIFW A

$$\begin{pmatrix} 1.81\times10^{-5} & 0 & 0 \\ 0 & 2.92\times10^{-5} & 0 \\ 0 & 0 & 3.35\times10^{-5} \end{pmatrix}$$

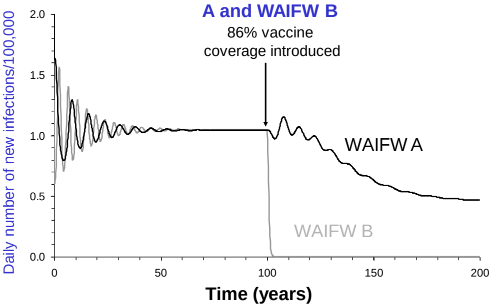


WAIFW B

$$\begin{pmatrix} 1.66\times10^{-5} & 4.16\times10^{-6} & 4.16\times10^{-6} \\ 4.16\times10^{-6} & 4.16\times10^{-6} & 4.16\times10^{-6} \\ 4.16\times10^{-6} & 4.16\times10^{-6} & 4.16\times10^{-6} \end{pmatrix}$$



Model predictions of the daily number of new infections/100,000 among old individuals in populations mixing according to WAIFW A and WAIFW B



Conclusion: R_0 for populations mixing according to WAIFW A is larger than the R_0 for populations mixing according to WAIFW B

Estimating R_0 for heterogeneously mixing populations

For heterogeneously mixing populations, the R_0 is estimated using the following steps:

- Measure the prevalence of previous infection in the population, using a serosurvey
- Estimate the forces of infection
- Choose the structure of the matrix of "Who Acquires Infection From Whom"
- Calculate the transmission coefficients β corresponding to the chosen WAIFW matrix
- Formulate the "Next Generation Matrix" (NGM)
- Calculate R_0 from the "Next Generation Matrix"

Example: writing down the Next Generation Matrix

For populations stratified into young and old individuals, the Next Generation Matrix is given by:

No. of ***young*** secondary infectious individuals by each ***young*** infectious person= $\beta_{yy}N_yD$

No. of ***young*** secondary infectious individuals generated by each ***old*** infectious person = $\beta_{yo}N_yD$

$$\begin{pmatrix} R_{yy} & R_{yo} \\ R_{oy} & R_{oo} \end{pmatrix}$$

No. of ***old*** secondary infectious individuals generated by each ***young*** infectious person = $\beta_{oy}N_oD$

No. of ***old*** secondary infectious individuals generated by each ***old*** infectious person = $\beta_{oo}N_oD$

Method for calculating the R_0

Method 1 (simulation approach)

Simulate the introduction of 1 infectious individual into a population in which individuals mix according to the given WAIFW matrix and there is an infinite supply of susceptible individuals.

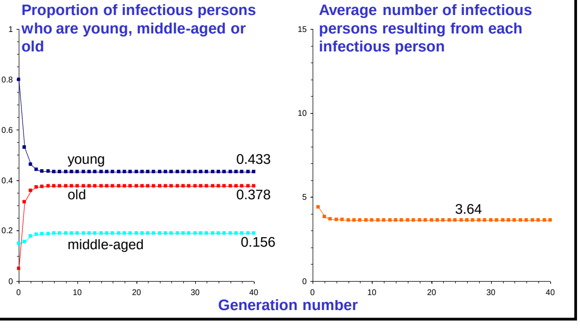
According to mathematical theory (beyond the scope of this course!)

after several generations have occurred:

- the age distribution of the infectious individuals in each generation converges to some distribution
- the number of secondary infectious individuals resulting from each infectious person equals R_0

=> Calculate R_0 as G_{k+1}/G_k, where G_k is the number of infectious individuals in generation k, once sufficient numbers of generations have occurred for this ratio to have stabilized.

The effect of simulating the introduction of one infectious person into a totally susceptible population with an unlimited supply of susceptible individuals mixing according to WAIFW B (R_0 practical)



Writing down the Next Generation Matrix to calculate R_n

For populations stratified into young and old individuals, the Next Generation Matrix is identical to that for calculating R_0, but need to substitute S_y and S_o for N_y and N_o:

No. of ***young*** secondary infectious individuals by each ***young*** infectious person= $\beta_{yy}S_yD$

No. of ***young*** secondary infectious individuals generated by each ***old*** infectious person = $\beta_{yo}S_yD$

$$\begin{pmatrix} R_{yy} & R_{yo} \\ R_{oy} & R_{oo} \end{pmatrix}$$

No. of ***old*** secondary infectious individuals generated by each ***young*** infectious person = $\beta_{oy}S_oD$

No. of ***old*** secondary infectious individuals generated by each ***old*** infectious person = $\beta_{oo}S_oD$

Methods for calculating R_0

Method 2 ("simultaneous equations" approach)

Calculate R_0 (for a population comprising 2 age groups) by solving the equations :

$$R_{yy}\,x + R_{yo}\,(1-x) = R_0 x$$
$$R_{oy}\,x + R_{oo}\,(1-x) = R_0\,(1-x)$$

The equations can also be written using the following matrix notation:

$$\begin{pmatrix} R_{yy} & R_{yo} \\ R_{oy} & R_{oo} \end{pmatrix}\begin{pmatrix} x \\ 1-x \end{pmatrix} = R_0\begin{pmatrix} x \\ 1-x \end{pmatrix}$$

NB This equation can be often be simplified to give a quadratic (i.e. of the form: ax^2+bx+c) => there may be 2 values for "R_0" which satisfy the above equation…

The formal (mathematical) definition of R_0 is that it is the "dominant eigenvalue of the Next Generation Matrix", i.e. the largest value for R_0 which satisfies the above equation

Q5. The following is the Next Generation Matrix for an infection which is transmitted from vectors to humans and from humans to vectors, but cannot be transmitted from humans to humans or from a vector to another vector.

$$\begin{array}{cc} & \begin{array}{cc} v & h \end{array} \\ \begin{array}{c} v \\ h \end{array} & \begin{pmatrix} 0 & 3 \\ 1.5 & 0 \end{pmatrix} \end{array}$$

Which of the following is incorrect?

a) The basic reproduction number is approximately 2.12.

b) One of the following statements is correct:
i) The fraction of the typical infectious "person" that is a vector is approximately 0.59. ii) The basic reproduction number is 4.5.

c) If vaccination is introduced just among humans, with two thirds of humans becoming completely protected against infection, the infection will eventually disappear among humans.

d) If vaccination is introduced just among humans, with one third of humans becoming completely protected against infection, the net reproduction of the infection will be approximately equal to 1.7.

d) If no humans are vaccinated but the vector population is sprayed with a chemical agent, so that all vectors are half as infectious as they were previously, the net reproduction will be approximately equal to 1.5.

Is the following incorrect?

a) The basic reproduction number is approximately 2.12.

$$\begin{array}{cc} & \begin{array}{cc} v & h \end{array} \\ \begin{array}{c} v \\ h \end{array} & \begin{pmatrix} 0 & 3 \\ 1.5 & 0 \end{pmatrix} \end{array}$$

Recall, for the following type of Next Generation Matrix:

$$\begin{array}{cc} & \begin{array}{cc} \text{Group 1} & \text{Group 2} \end{array} \\ \begin{array}{c} \text{Group 1} \\ \text{Group 2} \end{array} & \begin{pmatrix} 0 & R_{12} \\ R_{21} & 0 \end{pmatrix} \end{array}$$

R_0 is given by: $R_0 = \sqrt{R_{12}R_{21}}$

So R_0 is given by: $R_0 = \sqrt{R_{vh}R_{hv}}$

Substituting for R_{vh}=3 and R_{hv}=1.5 gives: $R_0 = \sqrt{3 \times 1.5} = \sqrt{4.5} \approx 2.12$

a) The basic reproduction number is approximately 2.12.

Is the following incorrect?

b) One of the following statements is correct:
i) The fraction of the typical infectious "person" that is a vector is approximately 0.59. ii) The basic reproduction number is 4.5.

Answer:
ii is incorrect as $R_0 \approx 2.12$

To find out if i is correct, calculate the fraction of the typical infectious person that is a vector, assuming $R_0 \approx 2.12$

Calculating the fraction of a typical infectious "person" (x) that is a vector (1)

Note that the fraction of a typical infectious "person" that is a vector has to satisfy the equation:

$$\begin{pmatrix} 0 & 3 \\ 1.5 & 0 \end{pmatrix}\begin{pmatrix} x \\ 1-x \end{pmatrix} = R_0\begin{pmatrix} x \\ 1-x \end{pmatrix}$$

Writing this equation in full, we obtain the following 2 equations:

$$0.x + 3(1-x) = R_0 x$$
$$1.5x + 0.(1-x) = R_0(1-x)$$

These equations simplify to the following:

$$3(1-x) = R_0 x \qquad Eq1$$
$$1.5x = R_0(1-x) \qquad Eq2$$

Rearranging equation 1, we get: $x = \dfrac{3}{R_0 + 3}$

Calculating the fraction of a typical infectious "person" (x) that is a vector (2)

Substituting for $R_0 \approx 2.12$ into $x = \dfrac{3}{R_0 + 3}$, we get:

$$x \approx \frac{3}{2.12 + 3} \approx 0.59$$

Is the following incorrect?

b) One of the following statements is correct:
i) The fraction of the typical infectious "person" that is a vector is approximately 0.59. ii) The basic reproduction number is 4.5.

Answer:
b) Is correct since only statement i) is correct

Which of the following is incorrect?

c) If vaccination is introduced just among humans, with two thirds of humans becoming completely protected against infection, the infection will eventually disappear among humans.

d) If vaccination is introduced just among humans, with one third of humans becoming completely protected against infection, the net reproduction of the infection will be approximately equal to 1.7.

d) If no humans are vaccinated but the vector population is sprayed with a chemical agent, so that all vectors are half as infectious as they were previously, the net reproduction will be approximately equal to 1.5.

Answer: Rewrite the Next Generation Matrix to account for vaccination and changes in infectiousness and calculate R_n

Which of the following is incorrect?

	True statement?	Right answe to question?
a) The basic reproduction number is approximately 2.12.	✓	✗
b) One of the following statements is correct: i) The fraction of the typical infectious "person" that is a vector is approximately 0.59. ii) The basic reproduction number is 4.5.	✓	✗
c) If vaccination is introduced just among humans, with two thirds of humans becoming completely protected against infection, the infection will eventually disappear among humans.	✗	✓
d) If vaccination is introduced just among humans, with one third of humans becoming completely protected against infection, the net reproduction number of the infection will be approximately equal to 1.7.	✓	✗
e) If no humans are vaccinated but the vector population is sprayed with a chemical agent, so that all vectors are half as infectious as they were previously, the net reproduction will be approximately equal to 1.5.	✓	✗

Q5. Which of the following is incorrect?

c) If vaccination is introduced just among humans, with two thirds of humans becoming completely protected against infection, the infection will eventually disappear among humans.

Original Next Generation Matrix:

$$\begin{array}{cc} & \begin{array}{cc} v & h \end{array} \\ \begin{array}{c} v \\ h \end{array} & \begin{pmatrix} 0 & 3 \\ 1.5 & 0 \end{pmatrix} \end{array}$$

New Next Generation Matrix – multiply the terms relating to humans by 1/3 (as only 1/3 of humans are not protected)

$$\begin{array}{cc} & \begin{array}{cc} v & h \end{array} \\ \begin{array}{c} v \\ h \end{array} & \begin{pmatrix} 0 & 3 \\ 1.5 \times \frac{1}{3} & 0 \end{pmatrix} \end{array} = \begin{pmatrix} 0 & 3 \\ 0.5 & 0 \end{pmatrix}$$

Which of the following is incorrect?

c) If vaccination is introduced just among humans, with two thirds of humans becoming completely protected against infection, the infection will eventually disappear among humans.

New Next Generation Matrix

$$\begin{array}{cc} & \begin{array}{cc} v & h \end{array} \\ \begin{array}{c} v \\ h \end{array} & \begin{pmatrix} 0 & 3 \\ 0.5 & 0 \end{pmatrix} \end{array}$$

R'_{hv} R_{vh}

Adapting the result that R_0 is related to R_{vh} and R_{hv} through the following equation:

$$R_0 = \sqrt{R_{hv} R_{vh}}$$

we obtain the result:

$$R_n = \sqrt{R'_{hv} R_{vh}} = \sqrt{0.5 \times 3} = \sqrt{1.5} \approx 1.22$$

c) Is incorrect since R_n>1, and so the infection will not eventually disappear (assuming that no other intervention is introduced)

Q5. Which of the following is incorrect?

d) If vaccination is introduced just among humans, with one third of humans becoming completely protected against infection, the net reproduction of the infection will be approximately equal to 1.7.

Original Next Generation Matrix:

$$\begin{array}{cc} & \begin{array}{cc} v & h \end{array} \\ \begin{array}{c} v \\ h \end{array} & \begin{pmatrix} 0 & 3 \\ 1.5 & 0 \end{pmatrix} \end{array}$$

New Next Generation Matrix – multiply the terms relating to humans by 2/3 (as only 2/3 of humans are not protected)

$$\begin{array}{cc} & \begin{array}{cc} v & h \end{array} \\ \begin{array}{c} v \\ h \end{array} & \begin{pmatrix} 0 & 3 \\ 1.5 \times \frac{2}{3} & 0 \end{pmatrix} \end{array} = \begin{pmatrix} 0 & 3 \\ 1.0 & 0 \end{pmatrix}$$

$$R_n = \sqrt{R'_{hv} R_{vh}} = \sqrt{1 \times 3} = \sqrt{3} \approx 1.73$$

d) Is correct

Q5. Which of the following is incorrect?

d) If no humans are vaccinated but the vector population is sprayed with a chemical agent, so that all vectors are half as infectious as they were previously, the net reproduction will be approximately equal to 1.5.

Original Next Generation Matrix:

$$\begin{array}{cc} & \begin{array}{cc} v & h \end{array} \\ \begin{array}{c} v \\ h \end{array} & \begin{pmatrix} 0 & 3 \\ 1.5 & 0 \end{pmatrix} \end{array}$$

New Next Generation Matrix – multiply the terms relating to infectious vectors by 0.5 (as vectors are half as infectious as they were previously

$$\begin{array}{cc} & \begin{array}{cc} v & h \end{array} \\ \begin{array}{c} v \\ h \end{array} & \begin{pmatrix} 0 & 3 \\ 1.5 \times 0.5 & 0 \end{pmatrix} \end{array} = \begin{pmatrix} 0 & 3 \\ 0.75 & 0 \end{pmatrix}$$

$$R_n = \sqrt{R'_{hv} R_{vh}} = \sqrt{0.75 \times 3} = \sqrt{2.25} = 1.5$$

e) Is correct

Next steps…

Session 8-9
Analyse infection prevalence data - estimate the average force of infection (λ) by fitting a catalytic model to data and calculate R_0 assuming random mixing

Session 10
Incorporate R_0 in age-structured model of rubella transmission, assuming random mixing and explore effect of vaccination on:
- age-specific proportion susceptible
- infection incidence

Session 11
Use λ to estimate age-dependent contact parameters
Incorporate these parameters in a model to explore effect of age-dependent contact on the impact of vaccination

Session 14
Calculate R_0 using age-dependent contact parameters

Summary of the key equations in block 2 (assuming random mixing)

1. $s(a)=e^{-\lambda a}$ if do not account for maternal immunity
2. $s(a)=e^{-\lambda(a-0.5)}$ if accounting for maternal immunity
3. $A \approx 1/\lambda$
4.

	Average proportion susceptible	**R_0**
Exponential	1/(1+L/A)	1+L/A
Rectangular	A/L	L/A

5. $R_0 = 1/s$, or equivalently, $s=1/R_0$

Summary of the key equations in block 2 (assuming random mixing)

1.
$$\overline{\lambda_o(t)} = \beta_{oy} I_y(t) + \beta_{oo} I_o(t)$$
$$\overline{\lambda_y(t)} = \beta_{yy} I_y(t) + \beta_{yo} I_o(t)$$

2. Next Generation Matrix to calculate:

R_0

$\beta_{yy}N_yD$ $\beta_{yo}N_yD$

$$\begin{pmatrix} R_{yy} & R_{yo} \\ R_{oy} & R_{oo} \end{pmatrix}$$

$\beta_{oy}N_oD$ $\beta_{oo}N_oD$

R_n

$\beta_{yy}S_yD$ $\beta_{yo}S_yD$

$$\begin{pmatrix} R_{yy} & R_{yo} \\ R_{oy} & R_{oo} \end{pmatrix}$$

$\beta_{oy}S_oD$ $\beta_{oo}S_oD$

Summary of the most useful equations in blocks 1 and 2 (and infectious disease epidemiology...)

1. $R_n = R_0 \times$ proportion susceptible (randomly mixing population)

2. For an endemic infection (with vaccination at below the herd immunity threshold), $R_n = 1$, so

$$R_n = R_0 \times s = 1$$

And therefore $s = 1/R_0$ if the infection is endemic.

For an epidemic to occur, $s > 1/R_0$

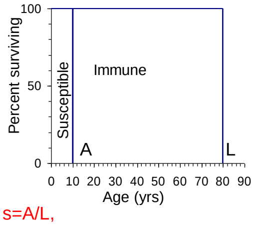
3. Herd immunity threshold = $1-1/R_0$

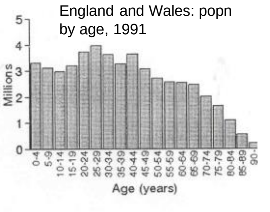
Extras…

Estimating the proportion susceptible in a population

Important for calculating the herd immunity

Rectangular populations:

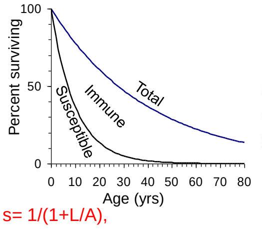


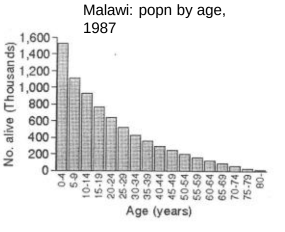


$s = A/L$,
where A= average age at infection,
L = life expectancy

Estimating the proportion susceptible in a population (2)

Exponential ("type II") age distribution:





$s = 1/(1+L/A)$,
where A= average age at infection,
L = average life expectancy

Estimating the basic reproduction number (R_0)

If the force of infection is not age-dependent, and if the infection is endemic:

$$R_0 = 1/\{\text{Proportion susceptible}\}$$

Note: $R_n = R_0 \times$ Proportion susceptible

For an endemic infection, $R_n = 1$

So,

$$R_n = R_0 \times \text{Proportion susceptible} = 1$$

Rearranging this expression gives:

$$R_0 = 1/\ \{\text{Proportion susceptible}\}$$

Rectangular age distribution: $s = A/L$ => $R_0 = L/A$

Exponential age distribution: $s = 1/(1+L/A)$ => $R_0 = 1+L/A$

Why is method 2 for calculating the R_0 equivalent to method 1?

If 1 infectious individual is introduced into a population in which individuals mix according to some WAIFW matrix and there is an "infinite" supply of susceptible individuals, then after several generations have occurred:

A. the age distribution of the infectious persons in each generation converges to some distribution

- e.g. A fraction x are young and a fraction 1-x are old

B. the number of secondary infectious persons from each infectious person equals R_0

These statements can be written using mathematical equations as:

$$R_{yy}x + R_{yo}(1-x) = R_0 x$$

$$R_{oy}x + R_{oo}(1-x) = R_0(1-x)$$

Derivation of method 2 for obtaining R_0

Note: If there are R_0 infectious individuals in a generation, then statement A=> there are R_0x young infectious persons and $R_0(1\text{-}x)$ old infectious persons in that generation.

But...using the Next Generation Matrix, the number of young infectious persons in that generation equals $R_{yy}x + R_{yo}(1\text{-}x)$

So $\quad R_{yy}x + R_{yo}(1\text{-}x) = R_0x$

Similarly, using the NGM, the number of old infectious persons in that generation equals $R_{oy}x + R_{oo}(1\text{-}x)$

So $\quad R_{oy}x + R_{oo}(1\text{-}x) = R_0(1\text{-}x)$

Summary of the matrix equation for R_0

The equations $\quad R_{yy}\,x + R_{yo}\,(1\text{-}x) = R_0x$

$R_{oy}\,x + R_{oo}\,(1\text{-}x) = R_0\,(1\text{-}x)$

can be summarized using matrix notation:

$$\begin{pmatrix} R_{yy} & R_{yo} \\ R_{oy} & R_{oo} \end{pmatrix}\begin{pmatrix} x \\ 1-x \end{pmatrix} = R_0\begin{pmatrix} x \\ 1-x \end{pmatrix}$$

NB There may be more than 1 value for "R_0" which satisfies the above equation…

The formal (mathematical) definition of R_0 is that it is the "dominant eigenvalue of the Next Generation Matrix", i.e. the largest value for R_0 which satisfies the above equation

Definitions (only for those very interested…)

An eigenvector of a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a vector $\begin{pmatrix} x \\ y \end{pmatrix}$ which, when it's multiplied by that matrix, the result is some factor (the "eigenvalue") multiplied by that vector.

e.g. $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is an eigenvector of the matrix $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ since $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector of the matrix $\begin{pmatrix} 0 & 5 \\ 5 & 0 \end{pmatrix}$ since $\begin{pmatrix} 0 & 5 \\ 5 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 5\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is an eigenvector of the matrix $\begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}$ since $\begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = 4\begin{pmatrix} 1 \\ 2 \end{pmatrix}$

2, 5 and 4 are the "eigenvalues" of these matrices

"Eigen" = "characteristic" in German…

Obtaining the result $R_0 = \sqrt{R_{12}R_{21}}$ for "disassortative" matrices

Step 1: Substitute the Next Generation Matrix into the matrix equation relating R_0 and the fraction of the typical infectious person that is in group 1 (x):

i.e. equation $\begin{pmatrix} R_{yy} & R_{yo} \\ R_{oy} & R_{oo} \end{pmatrix}\begin{pmatrix} x \\ 1-x \end{pmatrix} = R_0\begin{pmatrix} x \\ 1-x \end{pmatrix}$

becomes: $\begin{pmatrix} 0 & R_{12} \\ R_{21} & 0 \end{pmatrix}\begin{pmatrix} x \\ 1-x \end{pmatrix} = R_0\begin{pmatrix} x \\ 1-x \end{pmatrix}$

Step 2: Rewrite this matrix equation using simultaneous equations:

$$0 \times x + R_{12} \times (1-x) = R_0 x$$
$$R_{21} \times x + 0 \times (1-x) = R_0(1-x)$$

Obtaining the result $R_0 = \sqrt{R_{12}R_{21}}$ for "disassortative" matrices (2)

Step 2: The equations simplify to the following:

$$R_{12}(1-x) = R_0 x \qquad Eq1$$
$$R_{21}x = R_0(1-x) \qquad Eq2$$

Step 3: Rearrange these two equations so that you have an expression which is just in terms of R_0

Rearranging option 1 (sl laborious): Rearrange equation 1 to get an expression for x in terms of R_0 and substitute that expression into equation 2 to get an expression for R_0:

Exercise…

Obtaining the result $R_0 = \sqrt{R_{12}R_{21}}$ for "disassortative" matrices (2)

Rearranging option 2 (quick): Rearrange both equations to get an expression for x/(1-x) in terms of R_0, , R_{12} and R_{21} and set them equal to each other

Eq 1 $(R_{12}(1-x) = R_0 x)$ rearranges to: $\dfrac{R_{12}}{R_0} = \dfrac{x}{1-x}$

Eq 2 $R_{21}x = R_0(1-x)$ rearranges to $\dfrac{x}{1-x} = \dfrac{R_0}{R_{21}}$

Setting these two equations equal to each other gives: $\dfrac{R_{12}}{R_0} = \dfrac{R_0}{R_{21}}$

Multiplying both sides of the equations by R_0 gives: $R_0^{\,2} = R_{12}R_{21}$

Take the square root of this equation gives: $R_0^{\,2} = \sqrt{R_{12}R_{21}}$

$$\begin{array}{c|cc} & u & r \\ \hline u & \beta_1 & \beta_2 \\ r & \beta_2 & \beta_2 \end{array}$$

b) The rate at which individuals from rural areas contact each other is different from the rate at which individuals from urban areas contact each other.

Answer:
Correct, because individuals in rural areas effectively contact each other at a rate β_2, which is different from the rate at which individuals in urban areas effectively contact each other. which equals β_1

$$\begin{array}{c|cc} & u & r \\ \hline u & \beta_1 & \beta_2 \\ r & \beta_2 & \beta_2 \end{array}$$

c) Individuals from rural areas contact individuals from urban areas at the same rate at which they contact other individuals from rural areas.

Answer:
Correct, because individuals in rural areas effectively contact urban residents at a rate β_2, which is equal to the rate at which they contact other individuals from rural areas

$$\begin{array}{c|cc} & u & r \\ \hline u & \beta_1 & \beta_2 \\ r & \beta_2 & \beta_2 \end{array}$$

d) The rate at which individuals from rural areas contact each other is equal to the rate at which individuals from urban areas contact each other.

Answer:
Incorrect, because individuals in rural areas effectively contact each other at a rate β_2, which is different from the rate at which individuals in urban areas effectively contact each other, which equals β_1

$$\begin{array}{c|cc} & u & r \\ \hline u & \beta_1 & \beta_2 \\ r & \beta_2 & \beta_2 \end{array}$$

e) Individuals from rural areas contact individuals from urban areas at a rate which is different from the rate at which they contact other individuals from rural areas.

Answer:
Incorrect, because individuals in rural areas effectively contact urban residents at a rate β_2, which is equal to the rate at which they contact other individuals from rural areas

Assuming that β_1 is not equal to β_2, which of the statements below is incorrect?

	True statement?	Correct answer?
a) Individuals in urban areas effectively contact each other at a different rate from the rate at which they effectively contact individuals in rural areas.	✓	✗
b) The rate at which individuals from rural areas contact each other is different from the rate at which individuals from urban areas contact each other.	✓	✗
c) Individuals from rural areas contact individuals from urban areas at the same rate at which they contact other individuals from rural areas.	✓	✗
d) The rate at which individuals from rural areas contact each other is equal to the rate at which individuals from urban areas contact each other.	✗	✓
e) Individuals from rural areas contact individuals from urban areas at a rate which is different from the rate at which they contact other individuals from rural areas.	✗	✓

Q4. The following is the Next Generation Matrix relating to an infection that is transmitted between children and adults, in population Y. Children and adults are denoted by the letters c and a respectively. Which of the following statements is correct?

$$\begin{array}{c|cc} & c & a \\ \hline c & 2 & 1 \\ a & 1 & 2 \end{array}$$

a) Each adult leads to fewer infections in adults than they do in children.
b) Each adult leads to more infections in adults than they do in children.
c) The basic reproduction number is 2
d) The basic reproduction number is 1
e) The basic reproduction number 6

$$\begin{array}{cc} & \begin{array}{cc} c & a \end{array} \\ \begin{array}{c} c \\ a \end{array} & \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \end{array}$$

a) Each adult leads to fewer infections in adults than they do in children.
b) Each adult leads to more infections in adults than they do in children.

Answer
Each adult leads to 1 infection in children and to 2 infections in adults

$\therefore$b) is correct

$$\begin{array}{cc} & \begin{array}{cc} c & a \end{array} \\ \begin{array}{c} c \\ a \end{array} & \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \end{array}$$

c) The basic reproduction number is 2
d) The basic reproduction number is 1
e) The basic reproduction number 6

Answer
Each adult leads to 1 infection in children and to 2 infections in adults , i.e, 3 infections in total
Each child leads to 2 infections in children and to 1 infection in adults , i.e, 3 infections in total
$\therefore R_0=3$

$\therefore$c), d) and e) are incorrect

Q4. The following is the Next Generation Matrix relating to an infection that is transmitted between children and adults, in population Y. Children and adults are denoted by the letters c and a respectively. Which of the following statements is correct?

$$\begin{array}{cc} & \begin{array}{cc} c & a \end{array} \\ \begin{array}{c} c \\ a \end{array} & \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \end{array}$$

a) Each adult leads to fewer infections in adults than they do in children.
b) Each adult leads to more infections in adults than they do in children.
c) The basic reproduction number is 2
d) The basic reproduction number is 1
e) The basic reproduction number 6