

## Epidemic Models Practical 2 (Lecture 7): Deterministic compartmental models in populations with different levels of susceptibility and infectivity - 1.5 hours

Accompanying R codes: *EpiTut2a.R* (Case 1), *EpiTut2b.R* (Case 2), *EpiTut2b\_init.R* (optional section) with corresponding ODEs coded in *SIR.R*, *SIRMajorSus.R*, *SIRMajorSusInf.R* and function for  $R_0$  and WAIFW-matrix in *betaAndR0fc.R*

### Notes & Tips:

1. Move all of the above R-files into your working directory and change the directory path in the R-codes for *EpiTut2a.R* and *EpiTut2b.R*
2. Skip the mathematical derivations in the 'analytical assessment' sections if you feel uncomfortable with mathematical equations

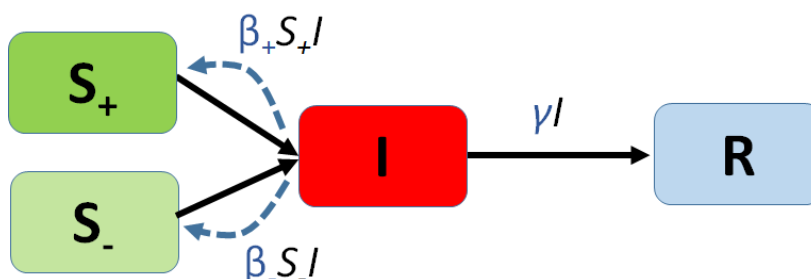
### Outline:

In this practical you will learn how to use deterministic epidemiological models for predicting epidemiological characteristics and the impact of control strategies in populations with variation in susceptibility. In particular, we explore how distinct levels of susceptibility affect epidemiological characteristics for different scenarios:

- Case 1: only difference in susceptibility
- Case 2: susceptibility confers difference in infectivity

### 1. Case 1: Difference in susceptibility only

Figure 1 shows a flow diagram for an SIR model in a closed population of constant size  $N$  for a disease for which individuals belong to one of two categories according to their susceptibility to the infection. Let  $S_+$  and  $S_-$  denote the ratio of susceptible individuals with high and low susceptibility, respectively, with corresponding transmission rates  $\beta_+ = \beta$  and  $\beta_- = \epsilon\beta$ , with  $0 \leq \epsilon \leq 1$ . Let's further assume that both groups have equal recovery rate  $\gamma$  and that the proportion of individuals with low susceptibility is  $p$ ,  $0 \leq p \leq 1$ . As in previous applications, we assume that the contact structure is homogeneous.



#### 1.1 Empirical assessment:

- a) Give some (real life) examples for which the model in Figure 1 may be a valid representation.
- b) What value does the parameter  $\epsilon$  have if the presence of a gene confers complete resistance to infection? What is the value of  $\epsilon$  in homogeneous populations?

**1.2 Analytical assessment:**

- a) Use the formula for  $R_0$  below to obtain the minimum proportion of individuals with low susceptibility required for preventing an epidemic outbreak as a function of the model parameters  $\beta, \gamma, \epsilon, p$ .
- b) Show that in the case of complete resistance this proportion is equal to  $1 - \frac{1}{\widehat{R}_0}$  where  $\widehat{R}_0$  refers to the basic reproductive ratio in a homogeneous population with transmission rate  $\beta$ .
- c) What proportion of fully resistant individuals would be required to prevent an outbreak for a disease with an estimated value of  $R_0=2$  in a population with high susceptibility?

$$R_0 = \frac{1}{\gamma} (p\epsilon\beta + (1-p)\beta)$$

**1.3 Numerical assessment of the infection dynamics:**

Open the R-script EpiTut2a.R and auxiliary file SIRMajorSus.R

The code describes the above SIR model in heterogeneous populations such as those described above with parameters

$\beta = 0.8$ ,  $\gamma = 1/4$ ,  $\epsilon=0.3$  and  $p = 0.2$  in a population of  $N=1000$  individuals, where 1 infectious individual of the type with low susceptibility is introduced into a population of 999 non-infected individuals.

Take some time to familiarize yourself with the code. Then use the code (in particular steps 1-3) to

- a) calculate  $R_0$  for the above parameters. What is the required proportion of individuals with low susceptibility to prevent a disease outbreak?
- b) generate infection profiles for the above population and compare them with those corresponding to a homogeneous population with the same average transmission rate. Does ignoring heterogeneity cause an upward or downward bias for the total proportion of infected individuals at the end of the epidemic?
- c) Explore the profiles for different parameter combinations. Does heterogeneity in susceptibility always produce an upward / downward bias?

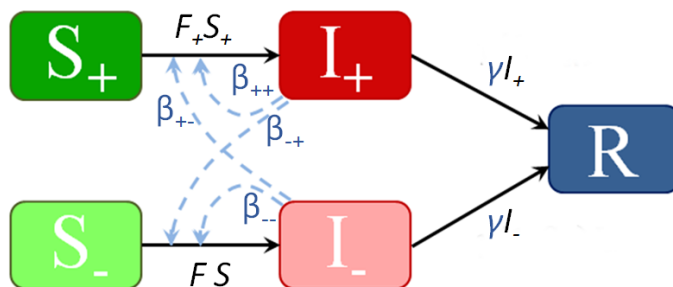
**1.4 Optional section:**

- d) Set the parameter values back to the original values (see above). Use steps 4 and 5 in the R-code to explore the impact of the effect size  $\epsilon$  on the epidemiological characteristics. Looking at the graphs you generated, how much do peak prevalence and duration of the epidemics differ between the scenarios where the maximum transmission rate  $\beta$  is reduced by a factor  $\epsilon = 0$ ,  $\epsilon = 0.5$  and  $\epsilon = 1$  for a proportion  $p=0.2$  of individuals with low susceptibility? Are 20% of completely resistant individuals sufficient to prevent an epidemic?

- e) Similarly, what impact does the proportion of individuals with low susceptibility have on the epidemiological characteristics in the case that the transmission rate for the highly susceptible group is twice as high as that of the group with low susceptibility and parameter values for  $\beta$ ,  $\gamma$  as above? Can the presence of individuals with low susceptibility prevent an outbreak?

## 2. Case 2: Difference in susceptibility and infectivity

In the model above we assumed that individuals differ in susceptibility only. Let's now relax this assumption and explore how differences in both susceptibility and infectivity affect the epidemics. For simplicity we assume that differences in susceptibility may confer differences in infectivity, and that  $S_+$  and  $I_+$  denote the category of susceptible and infectious individuals from the high susceptibility group and  $S_-$  and  $I_-$  denote the corresponding categories for the low susceptibility group. The figure below illustrates the flow chart of the corresponding epidemiological model. As above we assume that both groups have equal recovery rate  $\gamma$  and that the proportion of individuals with low susceptibility is  $p$ , with  $0 \leq p \leq 1$ .



For the analysis it is helpful to re-parameterize the above model by defining the transmission rates relative to a reference rate  $\beta_0$ . Let us set  $\beta_{++} = \beta_0$ ,  $\beta_{+-} = \epsilon_S * \beta_0$ ,  $\beta_{-+} = \epsilon_I * \beta_0$  and  $\beta_{--} = \epsilon_I * \epsilon_S * \beta_0$ . Here the scaling factor  $\epsilon_S$ ,  $0 \leq \epsilon_S \leq 1$  describes the effect of the low susceptibility type on the transmission rate and the factor  $\epsilon_I$  describes the effect of the low susceptibility type on its infectivity.

Thus, the WAIFW (beta) matrix for the above model is

$$\beta = \begin{pmatrix} \beta_{++} & \beta_{+-} \\ \beta_{-+} & \beta_{--} \end{pmatrix} = \begin{pmatrix} \beta_0 & \epsilon_I \beta_0 \\ \epsilon_S \beta_0 & \epsilon_I \epsilon_S \beta_0 \end{pmatrix}$$

### 2.1. Analytical assessment

(a) What can be said about the value of  $\epsilon_I$  and about the relative values for the transmission parameters in the case that

- (i) difference in susceptibility confers no difference in infectivity
- (ii) susceptibility and infectivity are positively correlated (i.e. high susceptibility confers high infectivity)

(iii) susceptibility and infectivity are negatively correlated (high susceptibility confers low infectivity)

Can you think of biological explanations for each of the scenarios (i) to (iii)? Which scenario is most likely in your opinion?

(b) Write down the model equations, the next generation matrix of the number of secondary cases produced through each of the four transmission routes, and the equations for the number of secondary cases produced by each of the two infectious classes ( $I_+$  and  $I_-$ ) to get expressions for the upper and lower bounds for  $R_0$ , i.e.  $R_0^+$  and  $R_0^-$ . The actual value for  $R_0$  is the eigenvalue of the next generation matrix. We will calculate this value in the R-script below.

## 2.2. Numerical assessment

Open the R-file `EpiTut2b.R` and auxiliary files `SIRMajorSusInf.R` and `c`

The code describes the above SIR model in a population comprising 1000 individuals. In these simulations, parameter values are set to  $\gamma = \frac{1}{2}$  days,  $\beta_0 = 1.5$  per day,  $\epsilon_s = 0.5$ , and  $p = 0.2$ . For  $\epsilon_i$  three values are chosen, representing different relationships between susceptibility and infectivity, i.e.  $\epsilon_i = 0.2$ ,  $\epsilon_i = 1$  and  $\epsilon_i = 5$ , respectively. The epidemics starts by introducing 1 infectious individual of the low susceptible type into a population of 999 non-infected individuals.

Take some time to familiarize yourself with the code. Then use the code (in particular steps 2 and 3) to answer the following questions.

- Check your model equations derived in 2.1b above by comparing your equations with the derivatives coded in the R-scripts (`SIRMajorSusInf.R` and `betaAndR0fc.R`). Also check your expressions for the 2 by 2 next generation matrix of the number of secondary cases and  $R_0^-$  and  $R_0^+$  derived in 2.1b (Step 2 in code – formulas coded in `betaAndR0fc.R`).
- For the parameter set above, calculate the basic reproductive ratios  $R_0$  for this model for the three scenarios  $\epsilon_i = 0.2$ ,  $\epsilon_i = 1$  and  $\epsilon_i = 5$ , respectively. Which scenario gives the highest value for  $R_0$ ? Is this expected? (Step 2 in code of `EpiTut2b.R`: call the function `betaAndR0fc` with different values of  $\epsilon_i$ ).
- How do these values compare with the values for the two thresholds  $R_0^-$  and  $R_0^+$ , and with the corresponding  $R_0$  for homogeneous populations with the same average transmission rate  $\beta = (1-p) * \beta_0 + p * \epsilon_s * \beta_0$ ? (Step 2 in code: Compare values of  $R_0$  and  $R_0\_homo$ ).
- Generate infection profiles corresponding to the four scenarios (1) homogenous population with the above average transmission rate  $\beta$  (2) heterogeneous population with difference in susceptibility only ( $\epsilon_i = 1$ ), (3) heterogeneous population where low susceptibility confers low infectivity ( $\epsilon_i = 0.2$ ), (4) heterogeneous population where low susceptibility confers high infectivity ( $\epsilon_i = 5$ ) (Step 3 in code). Inspect the plots to fill in the following table for epidemiological characteristics:

Characteristics / Scenario	Peak prevalence	Time of peak prevalence	Total proportion infected	Duration of epidemics	$R_0$
1					
2					

3					
4					

Which of the four scenarios causes the most / least severe epidemics?

- e. Now use step 4 in the code to explore the effect of the proportion of individuals with low susceptibility on the epidemiological characteristics. In particular, use the figures and  $R_0$  values to determine
- the effect (increasing or decreasing?) of increasing the proportion of individuals with low susceptibility on time and value of peak prevalence, total proportion infecteds, duration of the epidemics and on  $R_0$ .
  - in which case is a change in the proportion of individuals with low susceptibility from 50% to 75% more effective: if susceptibility and infectivity are not correlated or if they are positively correlated? [Tip: Compare the  $R_0$ -values associated with different proportions for both scenarios]
  - the proportion of individuals with low susceptibility required to achieve  $R_0 < 1$  in the case that  $\epsilon_s = 0.5$  and  $\epsilon_i = 0.2$ . Could one achieve  $R_0 < 1$  if low susceptibility confers no impact on infectivity?
- f. How may this kind of analysis help to determine the best disease eradication strategies for IPN in Atlantic salmon or other infectious diseases for which susceptibility loci have been identified? Is it ok to ignore potential differences in infectivity?

### 2.3 Optional section: Explore the impact of initial conditions on the infection dynamics

Save the code and plots of the infection dynamics generated in section 2.2 above.

Then modify this code to determine the effect of initial conditions on the infection dynamics. For example, rather than assuming that the infection is introduced by 1 individual with low susceptibility, assume that the infection is introduced by 100 individuals with low / high susceptibility type respectively or with 50 individuals of each type. How does this affect the prevalence profile? Do initial conditions affect  $R_0$ ?