Imperial College London short course on *Epidemiology & Control of Infectious Diseases*

Practical 2: Introducing Odin as a tool for modelling.

Drs Lucy Okell & Lilith Whittles

The aims of the practical are:

- To familiarise yourselves with the short course computers.
- To understand the concept, utility and limitations of numerical integration.
- To gain familiarity with the modelling software interface odin/R.
- To gain a hands-on understanding of the SEIR epidemic model.

In this hand-out, generally:

► Indicates an instruction.

► Indicates a useful tip or note.

► Indicates a question.

Part 0: Course materials

Materials for the course will be available via the short course SharePoint: https://imperiallondon.sharepoint.com/sites/DIDEShortCourseMaterials2022202-WP. This is organised as follows:

- **General information**
- **Week_1 and Week_2:** is where the lecturers will place course material for you (lecture notes, practical sheets, answers).

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

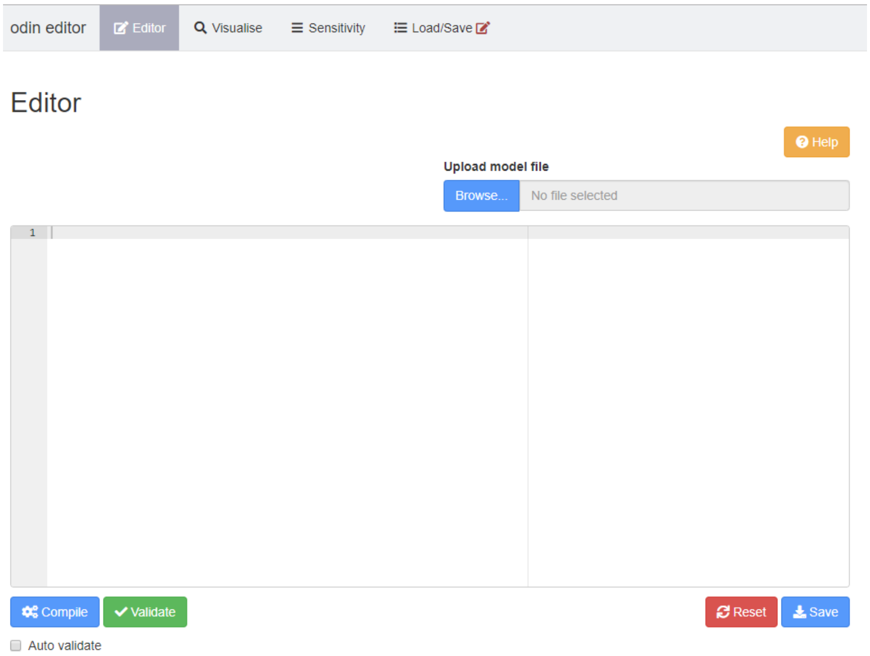
Example 1: Exponential growth

► *Navigate to the odin interface* https://shiny.dide.ic.ac.uk/infectiousdiseasemodels-2022/ *in Chrome or Safari.*

All practicals and projects for this short course will use a web-based interface. The underlying programming language is R. We will be using the "odin" R package with a Shiny interface during the next two weeks. A general help file about the odin interface is available in your folders and as a separate url. Each tab will have a Help *(help) button which will contain more useful tips.*

► *Click on "Practicals (Week 1), then "Go to App"*

The first window you see will be the **[Editor tab]**.



The first model you will solve is that of a population whose size changes at a constant per capita rate proportional to the population size. We denote *N(t)* as the population size at time *t* and a the rate of change in population size. In equation form, the model is

$$\frac{dN}{dt} = aN$$

with the initial condition $N=N_0$ when t=0.

► *Type the following into the blank Editor window exactly as follows:*

```
# variable
deriv(N) <- a * N
```

- Each variable, in this case "N" the population size, is specified by `deriv(variable)`. `deriv(N)` is R language for dN/dt and specifies that `N` is a quantity that changes over

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

time. To fully specify the model, you must also give the value of the population size at the start (in this case time 0) of the simulation (the initial value *N(0)*).

► The "#" is used to "comment out" code in the **Editor tab**. It is good practice to comment your work as much as possible so that you can understand what you did when you return to a file, for example you can note down the meaning of a parameter or piece of code. The "#" blocks out any subsequent words or symbols on the same line **# this is a comment** .

So, In this case "`# variable`" is a comment, but "`deriv(N) <- a * N`" is model code.

- o The "`<-` " is used to specify an object.

► **Caution: do not use the browser "back" button as you will lose any unsaved work.**

We recommend that you save your code regularly.

► *Type in the following:*

```
initial(N) <- N0
```

You must now give some numerical values to the parameters of the model.

► *Type:*

```
a <- user(0.5)
N0 <- user(1)
```

- The 'user' part means it will be easier to vary the parameter later.

► *Use the Validate button to check for any errors in the model code (or click the Auto validate tick-box to do this as you type). You will then see 'Validation: success' or 'Validation: error'.*

- Note that you can load code from an existing model from file (from a previous practical, for example) using Browse and edit it if you wish, rather than write code from scratch.

► *When you are happy with your model, click "Compile".*

- If there are any errors, fix them and compile again until you have no more errors.

► You can save the model code using the Save button just below the code editor. This will download a file with extension ".R" containing your model code (which is usually stored in your local 'Downloads' folder depending on your browser and computer set up. Don't forget to save your work. NB disconnection from the internet can erase the current model if unsaved).

► *In the **Visualise tab** you can visualise the model output. Notice on the left under 'Model parameters', the values you entered for N0 and a are displayed.*

► *Under "Run options" specify the "End time" by typing 20.*

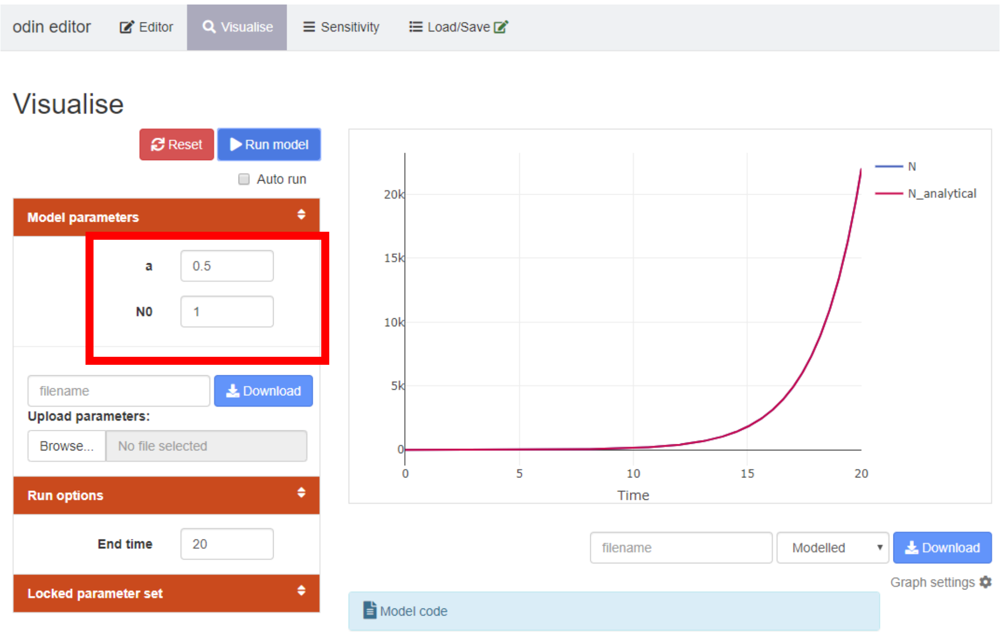
- This is how long the model will be run for (number of time steps).

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

► *Click Run model to visualise your model outputs*.

- Initially this will use default values for parameters, but you can manually change parameter values to see how they affect your model dynamics.

► *In the **Visualise tab** under "Model parameters", change the value of a from 0.5 to 0. Click Run model again to visualise your model outputs after changing your parameter values.*



a 0.5

N0 1

End time 20

- The growth rate is 0. The population size N, as seen in the Run window, is now constant over time.

► *Change the value of **a** again, this time to **-0.5** Click 'Run model**'**.*

- The population size N, as seen in the Run window, now declines exponentially. *a* is the decay rate.

► *At the top of the **Visualise tab**, click on 'Reset'.*

This resets the value to the value of the parameter 'a' to be the same as in the code **Editor tab**, i.e. 0.5. This will also "reset" the "End time" so remember to re-enter it.

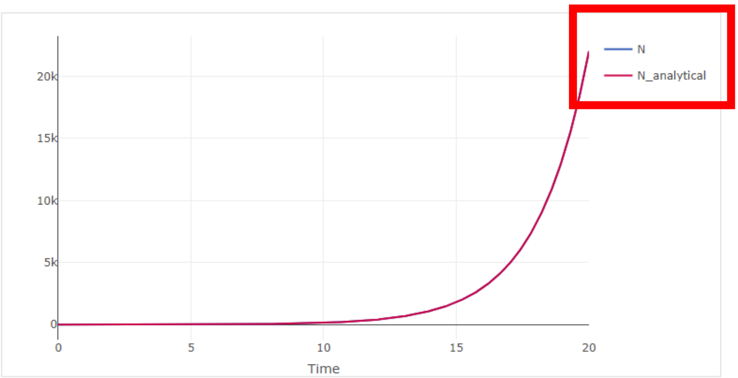
► *In the **Editor tab**, add an additional output variable. Type:*

```
output(N_analytical) <- N0 * exp(a * t)
```

- Use 'output(variable)' for variables which do not need to be calculated by a differential equation. Here "t" is implicitly time and does not have to be defined.
- Compile, then in the ***Visualise tab,*** run the model. You should see the output "N_analytical". Click on 'N_analytical' at the top right of the graph to remove it from the plot, then click again to add it back. Compare it with 'N' - what do you notice?

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

► In general, to add or remove variables on the plot, simply click the variable name in the top right of the graph.



► *See the Appendix at the end of this practical if you are interested in the derivation of N_analytical.*

► You can separately save parameter sets or tables of model outputs from the **Visualise tab** (using the Download buttons in the bottom right corner or at the bottom of the model parameters section).

► You can save plots by clicking on the 📷 icon. This will download a .png file of your plot.

► *From the 'Graph Settings" underneath the plot window* Graph settings ⚙ *, select 'log scale y-axis'*

- Now, with a log-scale axis, the population changes as a straight line. Notice however how the large numbers on the left axis now increase in multiples of 10. This is in fact the definition of a log-scale: exponential growth and decline appears as a straight line. If you want to see whether your data grows or decline exponentially, plot it on a log scale.

► The "graph settings" dialog box under the plot will also let you choose whether to plot variables on the right or left axis.

► Clicking on "Model Code" underneath the plot will expand a window that allows you to see your model code in the "**Editor tab**" in the same window as your model run.

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

Example 2. *SIR* Model - Closed

This is the basic model used yesterday to describe the dynamics of TB without births and deaths.

Step 1: the model.

The basic closed SIR model is written:

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \qquad N = S + I + R$$

$$\frac{dI}{dt} = +\frac{\beta SI}{N} - \sigma I$$

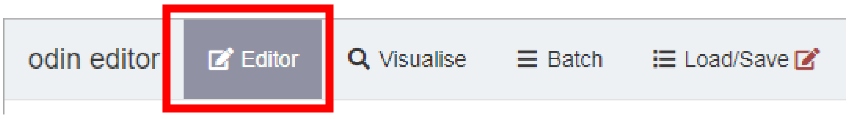
$$\frac{dR}{dt} = \sigma I$$

Recall that:

- Mean duration of infectiousness $= 1/\sigma$

► *We are now going to try and write the model from scratch. An example of the code can be found in the 'sir_closed.R' from the practicals folder of your course though we strongly encourage you to try and write your own version with the help of the demonstrators without looking at this file.*

► *Open a new* ***Editor tab*** *by right-clicking on the* ***Editor tab*** *at the top of the window, then select 'Open link in new tab'*



► *Now write out the equations for the model above. Each state variable (i.e. S, I and R) should be initialised (e.g.* `initial( S ) <- ..`*) and defined (e.g. with the* `deriv( S ) <- ..`*). The initial values add up to the total population size, for example 100,000 (you can choose a number), and should have at least one infected individual to seed an epidemic. Variables can take any name (eg.* `beta` *for* β*) just as long as there is no repetition and that they are given a value (see section below).*

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

Step 2: Evaluating the parameters.

First, we must choose time units (this is used in the definition of all the parameter values). We decide to use **weeks**, with

1 year = 52 weeks (= 364 days (!)).

- The infectious period lasts for 1/2 week, so $\frac{1}{\sigma} = 0.5$ weeks and hence σ = 2 weeks^{-1}
- For the transmission parameter β, we initially take a value of 4 people effectively contacted per person per week

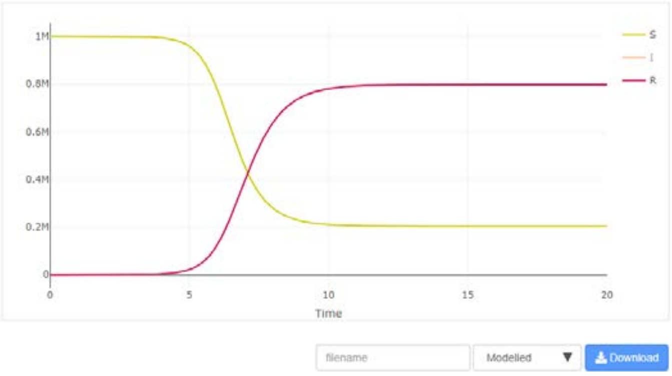
► *Define each of the parameters used in the model. Remember this is done by using `user(X)` eg.* `beta <- user(4)`.

Step 3: Compile and Run the model

► *Compile the model by clicking "Compile".*

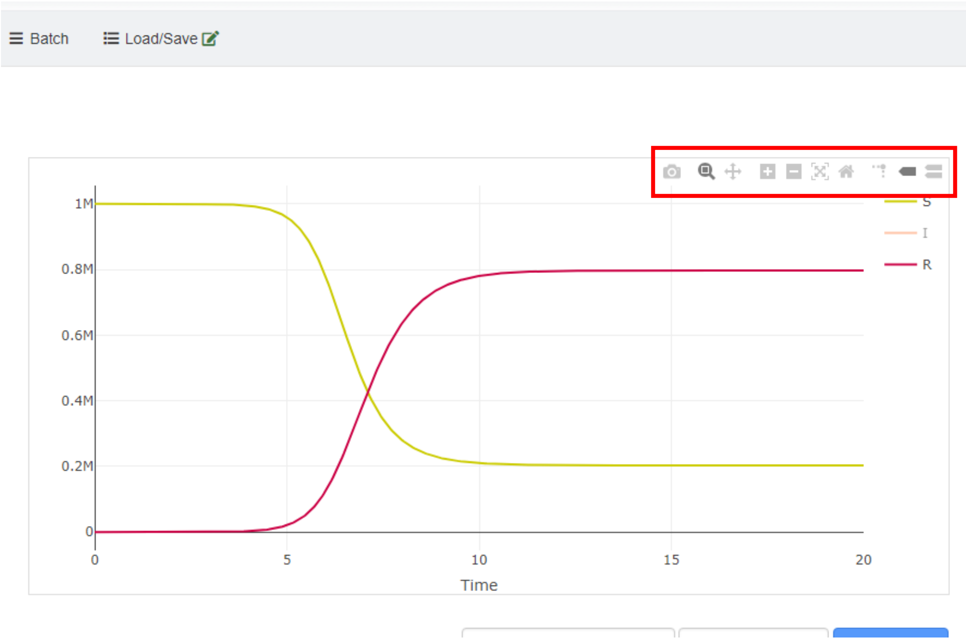
► *Then in the* ***Visualise tab****, run the model. Remember to specify an end time.*

► Remember, to add or remove variables on the plot, you can also simply click the variable name in the top right of the graph. The below example has the "I" variable toggled off.



► Plot options: hovering over the plot using the mouse will bring up several icons in the top right of the plot window.

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.



Clicking on:

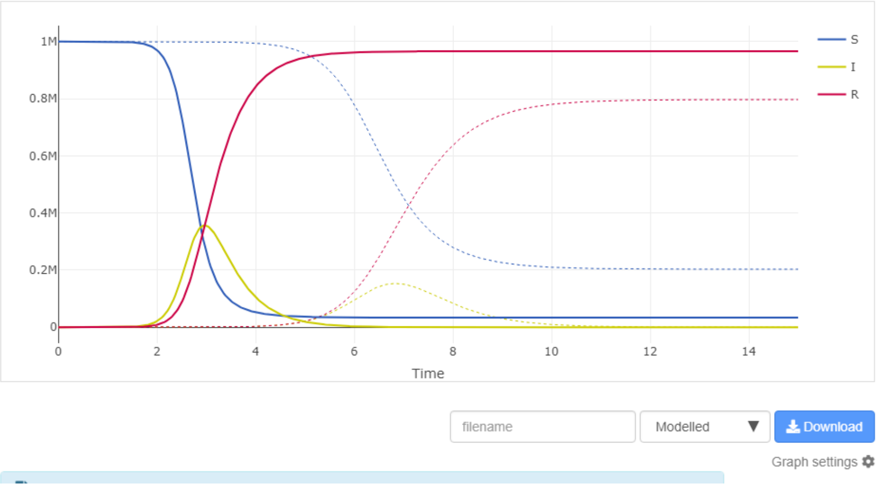
- 📷 : will download the plot as a .png.
- 🔍 : click and drag the crosshairs icon to zoom into a specific part of your plot.
- ✥ : will let you pan across the plot.
- ⊞ ⊟ : zoom in or out.
- 🏠 : will reset the axes after zooming. You can also double click anywhere on the plot to reset.
- ⁚ : will give you spike lines to read off the plot when you hover over your output.
- ◀ : will show you closest data when you hover over the plot.
- ☰ : will compare data when you hover over the plot.

The last two options are alternatives for visualising the values of the plot.

Other plot options:

- You can overlay different model runs by selecting "Lock current" under "Lock parameter set". This will store the current parameter set.
  - Then change one of your parameter values e.g. beta.
  - "Run model" again
  - This will plot the output using the stored "locked" parameter values in dashed lines and the new updated parameter values in solid lines.

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.



► You can save your modelled output (in this example, numeric values of S, I and R at each time step) and/or your parameter values by selecting “Modelled” or “Parameters” from the drop-down menu under the plot and clicking “Download”. You can then import this into Excel or any statistical software to analyze your output (or produce more visually appealing graphs!).

► *Make the overall population size and the number of initial cases parameters. For example, change your initialisers to something like,*

```
initial( S ) <- N - InfStart
initial( I ) <- InfStart
initial( R ) <-  0
```

making sure to define the new parameters,

```
N <- user(1e6)       # total population size (1e6 in scientific language means 1 x 10^6)
InfStart <- user(1)  # no of infectious cases at start of epidemic
```

The `N <- user(1e6)` *replaces the* `N = S + I + R.` *You can comment this out using “#”.*

We will now use the **Sensitivity tab** to explore the impact of different parameters and their values on the model outputs.

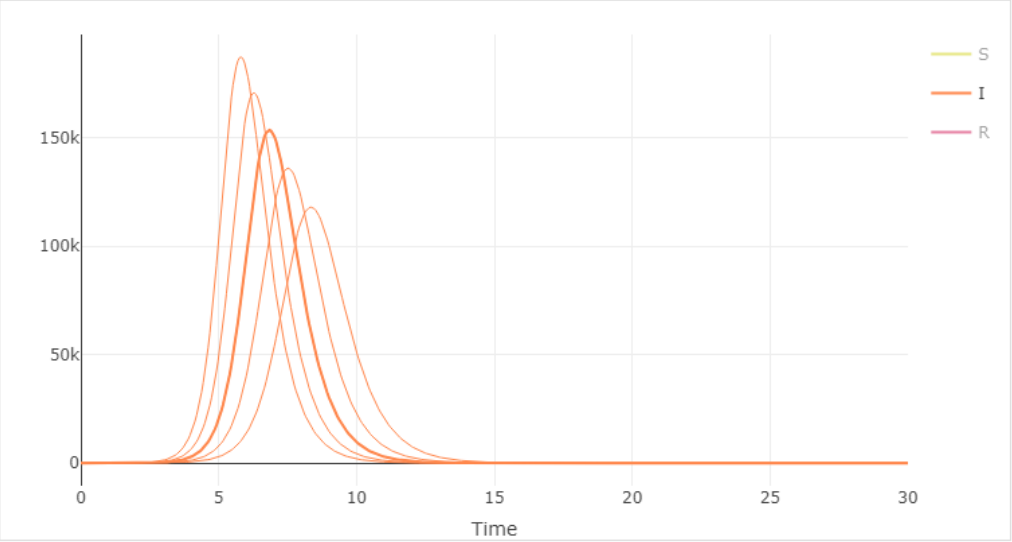
► *Conduct a sensitivity analysis by examining different parameters, and by changing the range of variation in the* ***Sensitivity tab.***

- In the **Sensitivity tab,** specify an “End time”
- In the “Vary parameter” section, select a “Parameter to vary” e.g. beta.
- In the ‘Variation type’ dropdown menu, select ‘Range’
- In the ‘From’ and ‘To’ boxes, enter the minimum and maximum values of the parameter you would like to use.
- Choose the “Number of runs” to specify how many values to test between this range.
- Select the “Trace over time” from the “Type of plot” drop-down menu.

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

  - “Trace over time” = how e.g. S, I, R, vary over time with each parameter value.

- Click “Run model”
- **The plot will look blank at this stage**. Select which variables to plot by clicking on the relevant legend e.g. I (shown below), in the top right of the plot.
- The bold line shows the model run for the value of the parameter you entered in ‘Model parameters’, and the thinner lines show model runs at other values of the parameter. Try hovering over one of the thinner lines using the mouse. This will display the x and y axis values at that point, and the parameter value used for that line (e.g. beta).



► *Conduct a sensitivity analysis by varying the parameter values one by one.*

► Which parameter(s) influence whether or not there is an epidemic (not the size or the speed of the peak)? By ‘epidemic’ here we mean any increase in the number of infected individuals after the initial introduction, even if not a very large increase.

► For each of these parameters explain in words why there appears to be a threshold value beneath which there is no epidemic.

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

► Explain in words what is happening to the number of new infectious cases over time above and below this threshold

Above:

Below:

► *Play around with the parameter values and choose 3 different values for sigma and beta that leads to an epidemic within the first year.*

Value of sigma	Minimum value of beta for an epidemic to occur

► Which parameter is larger when there is an epidemic, beta or sigma? Can you write this mathematically?

► The Basic Reproduction Number, often denoted R_0 is an expression that plays a key role in epidemiology. At the start of the epidemic, while most people are still susceptible to infection, R_0 is the average number of people one infected person newly infects. For an epidemic to occur in a population the newly introduced disease must have a $R_0 > 1$.

Can you write R_0 in terms of beta and sigma?

$$\boldsymbol{R_0} = \frac{\boxed{}}{\boxed{}}$$

R_0 is a central component of theoretical infectious disease epidemiology and you will be hearing plenty more of it in the next two weeks. Mathematical definitions of R_0 will depend on the biology of the infection and as a result the structure of the model used to represent it. We shall now go and explore the R_0 for other models in the *SIR* family.

END OF SECTION I

Please tell a demonstrator once you reach here. You may then proceed on to the next section of the practical.

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

Part 3. *SIR* Model - Open

This type of model is the same as the one you used yesterday to describe the dynamics of TB. To illustrate the importance of including births and deaths in the system we are going to parameterise the model for a short-lived organism.

Step 1: the model.

Births and natural mortality are included within the equations below,

$$\frac{dS}{dt} = B - bS - \frac{\beta SI}{N}$$

$$\frac{dI}{dt} = +\frac{\beta SI}{N} - (b + \sigma) I$$

$$\frac{dR}{dt} = \sigma I - bR$$

$$N = S + I + R$$

The simplest method of introducing basic demographic processes into the model is to assume all hosts die at a constant death rate. Though this may be an oversimplification for high income countries it has been shown to relatively accurately represent the age distribution of low-income populations.

To ensure that there is a constant population size the number of births, *B*, is adjusted so that *N* remains constant.

The meaning of the parameters:

- Life expectancy of uninfected = 1/b
- Mean duration of infectiousness = 1/sigma (conditional on survival)
- Mean duration of infectiousness = 1/(b+sigma) (including natural mortality)

Step 2: Evaluating the parameters.

We must choose time units, so again we decide to use **weeks**,

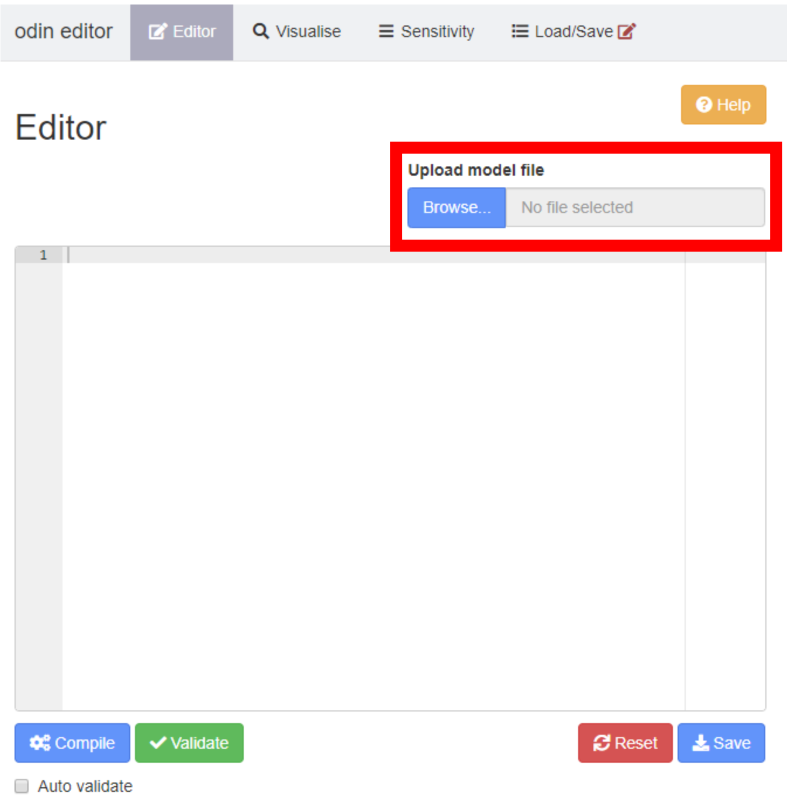
- Assume the average life-time is 1 year (52 weeks), the natural per-capita mortality rate is: $b = \frac{1}{52} \approx 0.02$ per week. (Note you cannot write `b <- user(1/52)` as the programme does not understand it, but `b <- user(0.02)` is ok.
- The population is steady at 1 million individuals (1e6), so the birth rate is:

  $Births = 10^6 * \frac{1}{52} \approx 20000$ per week.
- Use the same parameter values as the closed SIR model
- In the **Visualise tab**, set ‘End time’ to 5 years i.e. 52 weeks*5=260.

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

Step 3: The program

► *Copy the code file: 'sir_open.R' from the practicals folder for today to your own computer and open/load the code by clicking on the "Browse" button just above the code editor.*



► *Compile the model.*

► *Have a play with the basic model and see how including births and deaths influences disease dynamics.*

Using the R0 from the previous section and the parameters above can you derive the R0 for an open SIR model?

► *Programme the R0 expression into the Editor window*

As beta is defined in this model as a function of R0, so first you have to re-define beta as a standard user-defined parameter (`beta <- user(5)`), as in SIR closed.

As you want R0 as an "output" of the model now, type `output(R0) <- ..)`.

Note that once you introduce a new term into the equation window, you will need to re-compile the model. Hint: outputs and parameter values cannot have the same name.

How can you see if your prediction is correct?

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

You need to make sure that parameters are not used cyclically.

► *Investigate the dynamics of the open SIR model. How does it differ from the closed SIR?*

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

Part 4. *SEIRS*: A generic model of microparasite transmission

The SEIRS model has been described in the lecture.

Consider a measles-like virus spreading in a population of 10,000,000 (or 10^7 or 1E+7) individuals who have never been exposed to the virus (and who are thus all susceptible to infection).

The virus is readily transmitted primarily by the airborne route, and we assume that the population effectively mixes homogenously, with no subgroup being at increased or reduced risk of infection.

The duration of the latent and infectious period are both approximately 7 days (page 31 Anderson and May). To begin with we assume lifelong immunity. The life expectancy in the population is 75 years, and the case fatality rate is extremely low (almost 0).

Step 1: the model.

For this situation, the SEIR model seems an appropriate place to start. The model equations are:

$$\frac{dS}{dt} = B - bS - \frac{\beta SI}{N} + \delta R$$
$$\frac{dE}{dt} = +\frac{\beta SI}{N} - (b+\gamma)E$$
$$\frac{dI}{dt} = \gamma E - (b+\sigma+\alpha)I \qquad N = S + E + I + R$$
$$\frac{dR}{dt} = \sigma I - (b+\delta)R$$

Recall from the lectures the meaning of the parameters:

- Life expectancy of uninfected =1/*b*
- Mean duration of latency = $1/\gamma$ (*conditional on survival*)
- Mean duration of infectiousness =$1/\sigma$ (*conditional on survival*)
- Mean duration of immunity = $1/\delta$ (*conditional on survival*)
- Disease induced mortality rate *α*
- Prob. of dying from infection = $\alpha/(b+\alpha+\delta)$

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

Step 2: Evaluating the parameters.

First, we must choose time units (this is used in the definition of all the parameter values). Again, we decide to use **weeks**, with

- Since, average the average life-time is 75 years (=75*52=3900 weeks), the natural per-capita mortality rate is: $b = \frac{1}{3900} \approx 0.00026$ per week.
- Few people die from the infection, so $\alpha \approx 0$.
- Immunity is lifelong, so $\delta \approx 0 \; (1/\delta \to \infty)$.
- The population is steady at 10^7, so the birth rate is:
- $B = 10^7 * 0.00026 = 2600$ per week
- The latent period lasts 7 days, or 1 week: $\frac{1}{\gamma} = 1$ so $\gamma = 1$
- Similarly, the infectious period lasts 1 week: $\frac{1}{\sigma} = 1$ so $\sigma = 1$

The last parameter we need is *beta*, the transmission parameter: its value is far from obvious. Instead, we relate it to the basic reproduction number R0, by the formula derived in the lectures:

$$\boxed{R_0 = \frac{\gamma}{\gamma + b} \frac{\beta}{b + \alpha + \sigma}}$$

Inverting this equation relates *beta* to *R*$_0$, so R$_0$ becomes the parameter vary, which is much more comprehensible.

$$\beta = R_0 \frac{b + \gamma}{\gamma} \left(b + \alpha + \sigma \right)$$

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

Step 3: The program

► *Copy the file: 'seirs_generic.R' from the practicals folder to your computer and open/load it.*

► *Compile the model*

► *In the* ***Visualise tab****, set 'End time' to 500.*

► *In the* ***Sensitivity tab****, explore the effect of waning immunity on the long-term course of the epidemic.*

- **Hint**: this is the 'delta' parameter. Set the value as 0.05, and vary the parameter from 0 to 0.1 (remember, use the 'From' and 'To' boxes with 'Variation type' set to 'Range').

► *What is the effect of waning immunity on the long-term course of the epidemic?*

► *How many people are infected at the trough (bottom of the curve) of the inter-epidemic period? What do you think of this? Realistically, what is the* <u>*longest*</u> *duration of immunity consistent with the infection reaching a steady endemic state, in this population of 10 million?*

- **Hint:** the ≡ icon on the plot may be useful for exploring the output.

.

You may have noticed that in the equations, '*beta*' is not a parameter, but is expressed in terms of '*R0*'.

► *Explore the impact of 'R0' on the epidemic. Explore values of R0 ranging from 0 to 10.*

If you are interested in studying more systematically the effect of '*R0*' on the long-term outcome of infection, we could do it as follows:

► *In the* ***Sensitivity tab****:*

- *Set the rate of loss of immunity to a high value (e.g. 0.1), so that an endemic state is reached rapidly.*
- *Set the End time to be shorter to view the 1st epidemic in more detail (e.g. 120)*
- *Select 'R0' as the 'Parameter to vary' from the drop-down menu.*
- *Enter 'Number of runs' = 11 and explore values of R0 of 0 – 10.*
- *Run model*
- *Select just to visualise "I".*
- *You can now see lots of epidemic curves with differing values of R_0.*

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

If you are interested in studying more systematically the effect of '*R0*' on the final endemic prevalence of infection, try:

► *In the* ***Sensitivity tab****:*

- *Select 'R0' as the 'Parameter to vary' from the drop-down menu.*
- *Enter 'Number of runs' = 11 and explore values of R0 from 0 – 10.*
- *Select 'Value at a single time' as the 'Type of plot' from the drop-down menu. You can specify the 'Time to use' which by default is the last time point. Keep this default value.*
- *Run model*
- *Select just to visualise "I".*

The new window shows the number of people infected at the model End time as a function *not of time*, but of the basic reproduction number R_0.

► *Can you explain what happens when* R_0*=1? This is known as a threshold.*

TIP: There are other R files in the practicals folder ("other_examples") for the other types of models discussed: you are welcome to load them and explore their behaviour too.

END OF PRACTICAL

Other types of plots in the "Sensitivity tab"

- Value at a single time = the value of e.g. S, I, R, at a specific time point (can be specified using "Time to use". Default is the end/last time point) as the parameter value varies.
- Value at its min/max = the min/max value of e.g. S, I, R as the parameter value varies.
- Time at value's min/max = the time at which e.g. S, I, R reaches its minimum or maximum as the parameter value varies.

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.

Appendix. Population with constant per-capita growth rate model: analytic solution

In the first part of the practical, you are given a model of constant per-capita population growth. This is an example of a model which can be solved analytically (i.e. by hand using algebra) as well as numerically (e.g. using odin), unlike the SIR model. Here we show workings for solving a differential equation model via integration to derive the analytic solution shown in the lecture:

The model equation is: $\frac{dN}{dt} = aN$

where N is the number of individuals in the population, t is time, and a is the constant per-capita growth rate.

First, we rearrange to put all the terms containing N on one side of the equation (multiply both sides by dt, and divide both sides by N):

$$\frac{1}{N}dN = adt$$

Then apply the indefinite integral to both sides.

$$\int \frac{1}{N}dN = \int adt$$

Note that on the left hand side equation, we will integrate with respect to N, whilst on the right hand side, we will integrate with respect to t. $\int \frac{1}{N}dN = \ln(N)$ and $\int adt = at$.

Using these standard integrals, we obtain: $ln(N) = at + C$

Where C is the constant of integration. We exponentiate both sides of the equation (noting that $\exp(\ln(x)) = x$ to obtain:

$$N = e^{at+C}$$

By the laws of indices $x^{ab} = x^a x^b$ we can rewrite this as:

$N = e^{at}e^{C}$ or, by seeing that e^c is still a constant, more simply as $N = e^{at}C$

Now we set initial value (boundary) conditions in order to solve for the constant C. To obtain a simple result, we choose $t = 0$. Replacing $t = 0$ gives the following equation for N at time $0, N(0)$:

$$N(0) = e^{a \cdot 0}C$$

Since $e^{a \cdot 0} = e^0 = 1$,

$$C = N(0)$$

Replacing C in our above equation with N(0), we obtain our general solution:

$$N = N(0)e^{at}$$

Copyright © Imperial College London 2022. All rights reserved. Not for reproduction or distribution.