BIO445: Quantitative Life Sciences: from Infectious Diseases to Ecosystems *Tom Loosli*
06 December 2022 tom.loosli@usz.ch

Modelling emerging epidemics and outbreaks
Exercise sheet

In this module we will learn about the epidemiological applicability of simple mathematical models, that are often described as a set of ordinary differential equations (ODEs). This tool can assist in answering pertinent questions concerning the dynamics of the spread of an epidemic and the effectiveness of various control measures. This is especially useful at the initial stages of the outbreak when real data is still scarce and fast decisions regarding control measures and resource allocation have to be taken in real time and under high uncertainty. We demonstrate this approach using the 2014 Ebola outbreak in West Africa as described by Althaus (2014).

First, have a look at the paper (see `Althaus_PLoSCurrOutbreaks 2014.pdf`).

1 Model set-up

a. What is the difference between basic (R0) and effective (Re) reproduction numbers?

b. Draw a diagram of the model that shows the flow of individuals.

- The compartments/variables of the model are (use capital letters): _____.
- The parameters of the model are (use lower-case letters): _____.

c. Are all the parameters fixed or do some change with time? If yes, in what manner and why?

d. After how many months will the transmission rate drop to 50% of its initial value in Guinea and Sierra Leone? (use parameters from Table 2, recall the "half-life" equation from previous lecture). Guinea _____, Sierra Leone _____.
What can we learn about Liberia in this context?

e. Let's assume that due to the panic caused by the epidemic there is a rapid emigration of uninfected individuals with a constant rate of 1.5% per day. How would you modify the ODE system to reflect that?

f. What are the main simplifying assumptions of the model? Discuss with your partner.

2 Estimating R0 using linear regression, running and exploring the Ebola model

a. Read the Ebola data (from `Ebola_outbreak_West_Africa_data.csv`) into R and plot the cumulative number of cases versus time for each of the three countries. Add the number of deaths as well. (Hint: Use *as.Date()* function to convert date to right format).

b. It can be shown that one simple way to get a crude estimate of R0 is to fit a linear regression to the cumulative number of cases vs. calendar time at the initial stages of the epidemic. This stems from the following intuition: If we assume exponential growth of the number of infectious individuals:

$$I(t) = I(0) \cdot e^{\delta \cdot t}$$

Then we can take the natural logarithm of both sides, and we get:

$$ln(I(t)) = ln(I(0)) + \delta \cdot t$$

Which resembles the equation of the line (y=a+bx) that we know and love. In other words, at the initial stages of the epidemic, there is a linear relationship between the natural log of the cumulative number of reported cases and the epidemic growth rate δ. Hence, we can estimate δ as the slope of the line ("coefficient" in R) that shows the relationship between calendar time and the log of the cumulative number of Ebola cases.

c. Restrict the data to the first month (for simplicity, assume that the epidemic started at 22 Mar 2014), and use the "lm()" function in R to get the growth rate estimate for Guinea. Next, calculate the R0 using:

$$R0 = (1 + \delta \cdot incubation\, period\, (days)) \cdot (1 + \delta \cdot infectious\, period\, (days))$$

You can use incubation and infectious periods that are stated in the paper. R0 Guinea _____. Compare your estimate to Table 2.

d. Explore the starting script named `Ebola_starting_script_8.12.17`. Using the starting script, run the full model for Guinea and estimate the parameters along with their 95 % confidence intervals.

e. Plot the model results versus the real data with cases and deaths in the same plot(compare it with figure 1 of the paper). Using the starting script, run the full model for Guinea and estimate the parameters with least squares, next:

f. Plot the change of the effective reproduction number (β(t)/γ) with time (compare it with figure 1 of the paper).

g. Now, instead of plotting the cumulative number of cases, plot the number of new cases (incidence) for each day. Hint: use the R diff() function. When did the incidence reached the peak? Also add also the number of incident deaths to this plot.

h. Calculate how many cases and deaths were prevented with the control measures in Guinea, until August 20? Hint: Run a scenario without any control measures (by setting k to zero) and subtract the difference from the total observed number of cases (in the real data). Cases prevented _____, Deaths prevented _____.

3 Exploring scenarios and re-estimating parameters

a. Assume that due to lack of funding the control measures started not at day zero but with a delay of 30 days. Incorporate this change and re-run the model. Hint: which parameters govern the Beta(t) function? How many more cases and deaths could have been attributed to this one-month delayed response?
Additional Cases _____, Additional Deaths _____.

b. Until now we assumed that all the cases were diagnosed. Although Ebola is characterized by distinct symptoms (bleeding from the eyes), 100% diagnosis is still not realistic, especially in rural areas. Modify the ODE system in a way that a certain fraction of the infections is never diagnosed (and their faith remains unknown). Estimate (fit) this diagnosis rate parameter. Experiment with different starting guesses. Plot the "true" (diagnosed + undiagnosed) number of cases with time.

c. ∗For a given set of parameters, how does decreasing the diagnosis rate affects our R0 estimates (try 25% and 50%)? Why?

d. ∗The original model assumes exponential decay of transmission rate (Beta). Examine alternative scenarios (for example: a linear decay.). Estimate the parameters for the modified model. Does the model still fit the data well (compare the SSE)? Next, compare the Re change for both scenarios (two decay lines on the same plot). Is there a biologically relevant difference? What would you expect for Sierra Leone in this context (see Figure 2)?

4 Adding a vaccine

Let's assume that in addition to other interventions, a vaccine was developed that conveys 100% protection. Vaccination started 50 days after the beginning of the epidemic with 5% of the susceptible population vaccinated per day. Start from the initial model and modify the system accordingly. At which date the number of incident (new) cases will drop to 0 (round to the first decimal digit))?

Date _____.

References

Althaus, Christian L. 2014. "Estimating the Reproduction Number of Ebola Virus (EBOV) During the 2014 Outbreak in West Africa." *PLoS Currents* 6 (September). https://doi.org/10.1371/currents.outbreaks.91afb5e0f279e7f29e7056095255b288.