

Introduction to Tuberculosis Modelling
48th World Union World Conference on Lung Health, Guadalajara
11th October 2017

Practical 1: Setting up a model of *M tuberculosis*
transmission: exploring factors influencing tuberculosis
incidence

Overview and Objectives

This practical exercise is the first of two practicals which are designed to give you practice in thinking about key inputs in mathematical models for tuberculosis (TB), how the models are set up and the uses of these models.

This first practical is made up of two parts. The first part consists of a group work pen and paper exercise, which will involve thinking and discussing possible model structures and model input parameters. The second part consists of a computer exercise using a mathematical model for tuberculosis.

By the end of this practical, you should:

- Know the key input parameters which go into a TB model;
- Know some of the key factors which influence the observed TB disease incidence and the risk of *M tuberculosis* infection.

PART I: Groupwork exercise

Context

Suppose that you belong to some country (Z), in which the overall TB incidence is high (100-200/100,000 people per year) and has been stable over the last 10 years. In this simplified example, two key factors are thought to be the cause of this high incidence, namely:

- a) Most of the population of 100,000 people is poor and lives in very crowded conditions.
- b) The time from disease onset until cases are detected and started on TB treatment is thought to be very long but has not been estimated.

HIV and other determinants of TB are very rare in this setting and are not thought to be significant causes of tuberculosis incidence.

Your government is thinking of implementing different interventions in order to reduce tuberculosis incidence, and is particularly interested in ones which can increase the rate at which TB cases are detected and start treatment.

However, before doing so, it would like to understand how the above factors may influence tuberculosis incidence.

In small groups, you are asked to discuss the questions below for the next 20 minutes and summarize your answers on the flipcharts provided.

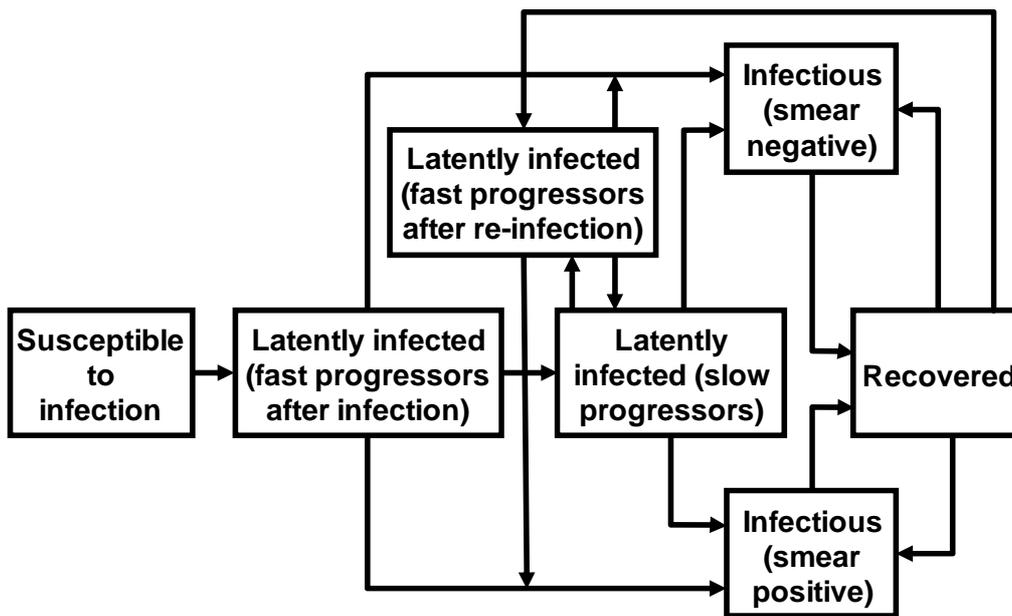
1. What would be a simple initial structure for the model which you might use to explore how the factors above influence tuberculosis incidence?
2. What are the limitations of your selected model structure?
3. What input parameters might your model require?
4. What data sources might you use to acquire these input parameters?

PART II: Computer exercise

In this part of the practical, we will use a model which has been set up for you to explore how the amount of contact between individuals and the duration of infectiousness affect tuberculosis incidence.

Introduction to the model

The general structure of the model is shown below, adapted from the model of Lin et al. So that we just focus on the key parameters in the model and how they influence the TB incidence, the model is not stratified by age.



The model has been set up for you in Berkeley Madonna.

1. Start up Berkeley Madonna and open the file **TB_epimodel.mmd** by clicking on the File option in the main menu, selecting the Open option and selecting the file from the appropriate location. This file was in the zip file that you downloaded before the course.

You will see a window, which is split into three panels. You can resize the width of each panel by clicking on the grey panel divider with your left mouse button and dragging to the left or right.

The left-hand panel contains the model equations, the second contains an empty graph (which we will refer to as the graph panel) and the third contains the actual parameters in the model. If necessary, you can hide any of these panels by selecting the Window option from the main menu and deselecting the corresponding panel from the drop-down list.

2. Spend a few minutes familiarizing yourself with the content of the equations panel. This has several sections:

- a) The model diagram, which contains a diagram of the general structure of the model.
- b) Equations for each compartment in the model. Do not worry about these for now – you can revisit them later in the practical if you wish.

- c) The numbers of individuals in each compartment at the beginning.
- d) The input parameters in the model, which are subdivided into those relating to infection, disease, case-finding/treatment, time-keeping, and demography. The parameters are based on the ones used by Lin et al.
- e) Useful summary statistics in the model, which include the prevalence of infectious people, annual tuberculosis incidence, the proportion of disease attributable to recent transmission and the annual risk of infection.

3. Run the model by clicking on the  button on the toolbar of Graph 1.

You should now see a figure with the annual TB incidence per 100,000, the annual TB mortality rate per 100,000 and the annual risk of infection for this setting (197 per 100,000 per year, about 51 per 100,000 per year and 2.4%/year respectively).

Factors influencing tuberculosis incidence – contact between individuals

We will now explore how the amount of contact between individuals influences the epidemiology of TB in the population. The model includes a parameter, called the effective contact rate (referred to as “ecr_pyr” in the model) for the number of individuals effectively contacted by each smear-positive case per year. An effective contact is defined as one that is sufficient to lead to infection if it occurs between an uninfected and a smear-positive case. It is currently set to be 15 per year.

Q1. What do you think will happen to the following statistics if you reduce the number of individuals effectively contacted by each smear-positive case:

- a) *The tuberculosis incidence?*
- b) *The tuberculosis mortality?*
- c) *The risk of infection?*

1. Check your hypotheses by changing the value for the effective contact rate from 15 to

13.5 per year, by moving the slider  for ecr_pyr below graph 1 to the left (until the indicator on the right of the slider says “ecr_pyr=13.5”) and recording your findings for these statistics in the table below. *Approximate answers are sufficient for this exercise. If you are interested, see Appendix A, technical notes 1 and 2 for other ways of getting these values accurately.*

Effective contact rate per year	Annual TB incidence per 100,000	Annual TB mortality rate per 100,000	Annual risk of infection (%)
15	197	51	2.4
13.5			
12	106	28	1.0
10.5			

2. The following table summarizes the impact of reducing the effective contact rate on the above statistics. Using a calculator or a spreadsheet, use the results from the previous step to complete the table by calculating the percentage reduction in each output statistic for each reduction in the effective contact rate.

Assumed reduction in the effective contact rate	% reduction:			
	Annual incidence 100,000	TB per	Annual TB mortality rate per 100,000	Annual risk of infection (%)
10% (15→13.5 per year)				
20% (15→12 per year)	46%		45%	58%
30% (15→10.5 per year)				

Q2. Based on your results, is the effective contact rate an important factor influencing tuberculosis incidence, mortality rate and the risk of infection?

Q3. (optional) What other infection or disease-related parameters in the model will also influence these statistics? N.B. Please do not change these statistics yet – you can do this later in the session if there is time.

Q4. (optional) According to the Styblo “thumb” rule, an annual risk of infection of 1% per year corresponds to an incidence of smear-positive disease of 50 per 100,000 per year.

a) Given that 70% of cases in the population are smear-positive, check to see if the values in the above table are consistent with this rule.

b) Suggest factors which might influence whether this relationship holds.

Factors influencing tuberculosis incidence – duration of infectiousness

We will now explore the effect of the duration of infectiousness on tuberculosis incidence. Cases in the model can die, recover naturally or be diagnosed and successfully treated. How quickly these things happen determines the length of time an individual remains infectious.

In the model, the average duration of infectiousness is represented as a single parameter for smear-negative and smear-positive cases (`ave_wks_infectiousness_smneg` and `ave_wks_infectiousness_smpos` respectively). Further details of how this average value is calculated are provided in Appendix B – you can explore this section later if you wish.

We will explore the factors which determine the duration of infectiousness, in particular the diagnostic and treatment pathway, in more detail in this afternoon’s practical.

Before continuing, change the effective contact rate to be back to its default value (15 per year) by moving the slider for the `ecr_pyr` accordingly.

Q5. What do you think will happen to the following statistics if you reduce the average duration of infectiousness for each smear-positive case:

a) The tuberculosis incidence?

- b) *The tuberculosis mortality?*
c) *The annual risk of infection?*

1. Check your hypotheses by changing the value for the average duration of infectiousness from 51 to 46 weeks, by moving the slider for ave_wks_infectiousness_smpos at the bottom of the graph panel and recording your findings for these statistics in the table below. Repeat this step for a duration of infectiousness of 36 weeks:

Average duration of infectiousness in smear-positive cases (weeks)	Annual TB incidence per 100,000	Annual TB mortality rate per 100,000	Annual risk of infection (%)
51	197	51	2.4
46			
41	122	28	1.3
36			

2. Using a calculator or a spreadsheet, use the results from the previous step to complete the table by calculating the percentage reduction in each output statistic for each reduction in the average duration of infectiousness for smear-positive cases:

Reduction in the average duration of infectiousness in smear-positive cases	% reduction:		
	Annual TB incidence per 100,000	Annual TB mortality rate per 100,000	Annual risk of infection (%)
10% (51→46 weeks)			
20% (51→41 weeks)	38%	45%	46%
30% (51→36 weeks)			

Q6. *Based on this table, is the average duration of infectiousness of smear-positive cases an important factor influencing tuberculosis incidence, mortality rate and the risk of infection?*

Q7. *How do you think your findings might have changed if you had considered smear-negative cases instead of smear-positive cases? If you have time, check your hypothesis by trying steps 3 and 4 below.*

3. **(optional)** Reset ave_wks_infectiousness_smpos to 51 weeks. Repeat step 1 above, but considering smear-negative cases, by moving the slider for ave_wks_infectiousness_smneg. Fill in the table below accordingly.

Average duration of infectiousness in smear-negative cases (weeks)	Annual TB incidence per 100,000	Annual TB mortality rate per 100,000	Annual risk of infection (%)
95	197	51	2.4
86			
76	184	44	2.2
67			

4. **(optional)** Complete the following table which summarizes the impact of reducing the average duration of infectiousness for smear-negative cases on the above statistics:

Reduction in the average duration of infectiousness in smear-negative cases	% reduction:		
	Annual TB incidence per 100,000	Annual TB mortality rate per 100,000	Annual risk of infection (%)
10% (95→86 weeks)			
20% (95→76 weeks)	7%	14%	8%
30% (95→67 weeks)			

Q8. **(optional)** *Based on this table, is it better to reduce the average duration of infectiousness of smear-negative cases or to reduce the average duration of infectiousness of smear-positive cases?*

Conclusions (for general discussion)

Q9. *For the setting considered in the model, is it better to reduce the average duration of infectiousness of smear-positive cases by 10% or to reduce the number of individuals effectively contacted by each smear-positive case per year (effective contact rate) by 10%?*

Q10. *What do you think are the key steps in the diagnostic and treatment pathway that determine the average duration of infectiousness?*

We have now completed the first computer exercise. If you have finished early, you can try the supplementary questions below.

In practical 2, we will explore the effect of improving diagnostics in further detail and look at how we might implement an intervention in the model.



Supplementary questions (optional)

Exploring factors influencing the proportion of disease attributable to recent infection, reactivation and reinfection

If you have time, try the steps below to explore how the proportion of disease attributable to recent infection, reactivation and reinfection depends on the risk of infection. Before continuing, reset the parameters in the model to be at the default values (ecr=15 per year, ave_wks_infectiousness_smpos=51 weeks, ave_wks_infectiousness_smneg=95 weeks)

S1. Why might it be useful to know if a large proportion of disease is attributable to recent transmission in a given population?

As you saw earlier in the practical, the setting considered in the model has an annual risk of infection of about 2% per year.

S2. Do you think that a high proportion of disease is due to recent infection (prop_dis_recent_infn) or reinfection (prop_dis_recent_reinfn) in the model population? Check your hypotheses by looking at Graph 2 in the graph panel, after running the model.

S3. How might your findings change if you increase the annual risk of infection in your population? You can check your hypotheses by changing the value for the effective contact rate in the model using the slider.

S4. How does the overall proportion of disease attributable to recent transmission (i.e. recent infection+recent reinfection) differ between the high and reduced infection risk settings? Why does this occur?

The model includes a parameter for the protection provided by previous infection against disease following reinfection. To date, the size of this parameter is not well understood, although several modelling studies and reviews of historical data have attempted to estimate it. Lin et al have assumed a value of 65%.

S5. How do you think the tuberculosis incidence would change if the size of this protection were to be reduced or increased? Return to graph 1 and check your hypotheses by changing the size of the protection (prot_init_infn_vs_reinfn) to 40% and 80%.

The model of Lin et al uses values for the rate at which individuals experience disease following recent infection, reinfection and through reactivation which have been estimated for Western populations. However, it is plausible that these rates differ for poor settings.

S6. How do you think the tuberculosis incidence and proportion of disease attributable to recent transmission would change if the rate at which individuals experience reactivation is 5-fold higher than has been assumed here? Check your hypotheses by changing the rate at which individuals reactivate a previous infection (dn_pyr) in the parameter window (see Appendix A, Technical note 1) and running the model.

S7. How would a low proportion of disease attributable to recent transmission influence the impact of an intervention?

Appendix A

Technical Note 1: Varying parameters

Parameters should be varied by one of two ways in Berkeley Madonna:

1. Moving the head of the slider  for the parameter of interest (located immediately below the graph) until the value displayed to the right of the slider equals the value intended. The following shows the slider for the `ecr_pyr`.



2. Clicking on the parameter of interest in the parameters panel, typing its new value in the box next to the reset button and then clicking on the run button. Note that clicking on the reset button just resets the parameter to its initial value.

Note that changing the value in the equation panel will not usually result in the value being updated when the model is run next.

Technical Note 2: Reading off values

The most accurate way of obtaining the value for a given variable in a plot in the graph panel is to select the table button  on the graph toolbar and then scrolling down to the time point that you're interested in. To return to the graph view, you just need to click on the table button  again. Note that for graph 1, the year is not displayed by default. You can include it in the display by clicking on the button for "year" at the bottom of the graph panel.

Appendix B

Method for calculating the duration of infectiousness in the model

In a mathematical model of the kind used here (deterministic, compartmental), the average length of time that an individual remains in a compartment depends on the speed or rate at which they leave the compartment and is given by the formula:

$$\text{Average length of time in compartment} = 1/(\text{rate of movement out of compartment})$$

In the TB model used in the practical, the duration of infectiousness is the length of time individuals remain in the infectious compartments. Individuals leave the infectious compartment if they die, naturally recover or are detected and start treatment (individuals are considered to be no longer infectious very soon after starting TB treatment). The duration of infectiousness is therefore given by the following formula:

$$\text{Duration of infectiousness} = 1/(\text{rate at which cases die} + \text{rate at which cases naturally recover} + \text{rate at which cases are detected and start treatment})$$

Using the values for the rates assumed in the model, we can show that the duration of infectiousness is 51 weeks (0.9 years) in smear-positive cases and 95 weeks (1.8 years) in smear-negative cases in the model.

For example, tuberculosis cases experience an increased risk of death compared to non-TB cases. The death rate depends on smear status and, following Lin et al, it is taken to be 0.19 per year in smear-negative cases and 0.22 per year in smear-positive cases. The background mortality rate is 0.021 per year.

Natural cure is assumed to occur at the same rate in both smear-negative and smear-positive cases and is equal to 0.20 per year based on Lin et al.

The rates at which cases are detected and start treatment depend on the diagnostic and treatment pathway and are described in more detail below. Smear-positive cases start treatment at a rate of 0.011 per week (or 0.572 per year). Smear-negative cases start treatment at a rate of 0.0026 per week (or 0.135 per year).

These values result in an average duration of infectiousness of 51 weeks (0.9 years) in smear-positive cases and 95 weeks (1.8 years) in smear-negative cases.

Varying the duration of infectiousness

In the version of the model used in the practical, the duration of infectiousness is captured as a single parameter which you can vary to explore its impact on the epidemiology of tuberculosis. We assume that changes in the duration of infectiousness result from changes in the rate at which cases are detected and start treatment (and that the mortality and natural cure rates remain unchanged). To be consistent with this assumption, you will see that when varying the duration of infectiousness in the model, we are changing the rate at which cases are detected and treated, according to the following equation:

rate at which cases are detected and treated = (1/duration of infectiousness) – rate at which cases die – rate at which cases naturally recover

This equation is obtained after rearranging the equation described in the section above.

Calculating the rate at which smear-positive or smear-negative cases are detected and start treatment

Cases can be detected when they visit the health services, where their sputum is examined by smear microscopy. If they are smear-negative, a proportion (30%) receive a chest X-ray and a proportion of these (depending on the sensitivity of X-ray, which is assumed to equal 80%) are referred for treatment. A proportion (0.15) of those referred for treatment default from treatment, which is assumed to be identical for both smear-positive and smear-negative cases, and of those who start TB treatment, only 75% complete treatment.

Smear-positive cases

The rate at which smear-positive cases are detected and start treatment is calculated using the following formula:

$$\begin{aligned} & \text{Rate at which cases come to the health centre and start treatment (=1/(average time until} \\ & \quad \text{they visit a health centre and start treatment))} \\ & \quad \times \\ & \text{Proportion of cases who do not default from treatment (=0.85)} \\ & \quad \times \\ & \text{Proportion of cases for whom treatment is successful (=0.75)} \end{aligned}$$

The values for the last two terms in this formula are taken as the mid-point of values provided in Table S2 of Lin et al.

The average time until cases visit a health centre and start treatment depends on several factors, namely the average time until they visit a health centre, the time delay until a sputum sample is taken for examination by smear, the time until the smear result is returned and the time delay until they start treatment.

Lin et al do not provide values for the average time until smear-positive or smear-negative cases visit a health centre; instead they provide the range of values for HIV-negative and HIV-positive individuals (1-25 and 1-7 months respectively). For consistency with Lin et al, we assume that the time for smear-negative and smear-positive cases is identical and equals 12 months or 52 weeks. The average time until a sputum sample is provided after the initial health centre visit is said to be 30 days, with a further 3 days until the result returns and a further 5 days until treatment is initiated.

Using these values, we see that the average time until cases come to the health centre and treatment is initiated is approximately $52 + 30/7 + 3/7 + 5/7$ weeks ≈ 57 weeks

The rate at which smear-positive cases are detected and start treatment is therefore given by:



$$\frac{1}{57} \times 0.85 \times 0.75 \approx 0.011 \text{ per week}$$

Smear-negative cases

The rate at which smear-negative cases are detected and start treatment is calculated in a similar way to that for smear-positive cases, except for the fact that it needs to account for the assumed proportion of smear-negative cases who are examined with a chest X-ray (0.3) and the sensitivity of the chest X-ray (80%), as given by the following formula:

$$\begin{aligned} & \text{Rate at which cases come to the health centre and start treatment if X-rayed (=1/(average} \\ & \quad \text{time until they visit a health centre and start treatment if X-rayed))} \\ & \quad \times \\ & \text{Proportion of smear-negative cases who have access to a chest X-ray (=0.3)} \\ & \quad \times \\ & \text{Sensitivity of chest X-ray for identifying smear-negative TB (=0.8)} \\ & \quad \times \\ & \text{Proportion of cases who do not default from treatment (=0.85)} \\ & \quad \times \\ & \text{Proportion of cases for whom treatment is successful (=0.75)} \end{aligned}$$

The rate at which smear-negative cases are detected and start treatment is therefore given by:

$$\frac{1}{57} \times 0.3 \times 0.8 \times 0.85 \times 0.75 \approx 0.0026 \text{ per week}$$