**Appendix**

**A.1. Selection of model structures**

Table A1 shows the 12 model structures identified in Menzies et al. [17]. A subset of 6 of these (models 1-6 in table A1) were analysed in Ragonnet et al. [16]. In this work we included 3 of these 12 structures.

Initially we selected all models giving a “good” fit to the data (see table A1). This resulted in 8 possible model structures (models 1,2,6,7,8,9,11,12).

Model 6 was excluded as in both previous studies the best estimate of the parameter “a” was zero, such that this structure is equivalent to model 3. We also excluded model 7 as it includes 3 additional parameters but does not improve the model fit. Models 8 and 9 are rarely used in the published literature. We chose not to consider models 11 and 12 as these were implemented as individual based models in Menzies et al. [17] and our focus here is differential equation models.

Model 3 was shown to provide a marginal fit to the data, however it has been used in a large proportion of published modelling studies; approximately 50% according to the literature review in [17]. We included it here to explore the implications of using an “inappropriate” structure.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **Parameters** | **% of published studies\***  | **Menzies et al. RMSE (qualitative fit)$** | **Ragonnet et al. Maximum likelihood** | **Comments** |
| 1 |  | 3 | 10.6 | 0.0447 (good) | -1254.9  | Included. |
| 2 |  | 3 | 11.2 | 0.0447 (good) | -1254.9 | Included. |
| 3 |  | 2 | 49 | 0.689 (marginal) | -1484.2 | Included for comparison of commonly used “inappropriate” structure. |
| 4 |  | 1 | 19.2 | 3.07 (poor) | -1484.2 | Excluded as provides poor fit in both studies. |
| 5 |  | 0 | 1 | 91.7 (poor) | Not reported | Excluded as provides poor fit in both studies. |
| 6 |  | 4 | 0.3 | 0.0447 (good) | -1254.9 | Excluded. In both studies `a` was estimated to be zero so collapses to model 3.  |
| 7 |  | 6 | 8.7 | 0.0448 (good) | - | Excluded. Includes extra parameters but does not improve fit. |
| 8 |  | 3 | 0.6 | 0.0447 (good) | - | Excluded as rarely used in the published literature |
| 9 |  | 3 | 0.6 | 0.0447 (good) | - | Excluded as rarely used in the published literature |
| 10 |  | 5 | 0.3 | 3.09 (poor) | - | Excluded as provides poor fit.  |
| 11 |  | 2 | 0.3 | 0.0586 (good) | - | Excluded as implemented as individual based model. |
| 12 |  | 3 | 0.3 | 0.0468 (good) | - | Excluded as implemented as individual based model. |

**Table A1.** \*% of published studies that used this structure as reported in Menzies et al [17]. $In Menzies et al [17] qualitative fit was defined as poor for RMSE >1, marginal for RMSE between 1 and 0.1 and good for RMSE <0.1.

**A.2. Full details of included models**

Tables A2 and A3. list the variable and parameter names used in the models. Equations and steady state solutions for each of the 3 included models are given below.

|  |  |
| --- | --- |
| **Variable** | **Description** |
| *S* | Susceptible. Never exposed to *M.tb.* |
| *LF* | Latently infected with *M.tb*. At risk of “fast” progression to active TB |
| *LS* | Latently infected with *M.tb*. At risk of “slow” progression to active TB |
| *I* | Active TB disease |
| *PF* | Post-preventive therapy (from *LF*) |
| *PS* | Post-preventive therapy (from *LS*) |
| *T* | Total population  |

**Table A2. Model variables**

|  |  |
| --- | --- |
| **Parameter** | **Description** |
| *a* | Proportion progressing directly to disease after infection |
| *b* | Proportion entering fast latent state after infection |
| *c* | Rate of progression to disease from slow latent state (per year) |
| *e* | Rate of movement from fast latent state to slow latent state (per year)  |
| *k* | Rate of progression to disease from fast latent state (per year) |
| *m* | Excess mortality rate due to active TB (per year) |
| *q* | Relative susceptibility to reinfection if previously infected |
| *u* | Background mortality rate (per year) |
| *β* | Effective contact rate (per year) |
| *θ* | Rate of starting preventive therapy (per year) |
| *w* | Relative rate of progression to disease from post preventive therapy state (compared to latently infected states). *w* = 0 means preventive therapy completely removes risk of disease from an existing infection. |

**Table A3. Model parameters**

**A.2.1. Model 1**

The model equations are:

At steady state, in the absence of preventive therapy (*PF*(0) = *PS*(0) = *θ* = 0), we have:

We define and . We first substitute the expression for *LF* (equation A8) into the expression for *LS* (equation A9) to get:

Then we substitute the expressions for *LF* (A8), *LS* (A11) and *S* (A7) into the expression for *I* (A10) and multiply through by the denominatiors to obtain an expression in terms of *I* only:

Equation A12 can be rearranged to obtain a qudaratic equation for *I* which can be solved to give the steady state value of *I*:

Where:

**A.2.2. Model 2**

The model equations are:

At steady state, in the absence of preventive therapy (*PF*(0) = *PS*(0) = *θ* = 0), we have:

We define . Then we substitute the expressions for *LF* (A24), *LS* (A25) and *S* (A23) into the expression for *I* (A26) and multiply through by the denominatiors to obtain an expression in terms of *I* only:

Equation A27 can be rearranged to obtain a qudaratic equation for *I* which can be solved to give the steady state value of *I*:

Where:

**A.2.3. Model 3**

The model equations are:

At steady state, in the absence of preventive therapy (*PS*(0) = *θ* = 0), we have:

We define. Then we substitute the expressions for *LS* (A37) and *S* (A36) into the expression for *I* (A38) and multiply through by the denominatiors to obtain an expression in terms of *I* only:

Equation A39 can be rearranged to obtain a qudaratic equation for *I* which can be solved to give the steady state value of *I*:

Where:

**A.3. Cumulative incidence of TB**

Figure A1 shows the cumulative incidence of TB (proportion of cohort experiencing disease) by time since infection by each model structure. Expressions for the final proportion predicted by each model are given below. These results and expressions assume no re-infection occurs.



**Figure A1. Cumulative proportion that have developed TB by time since infection.** Colours indicate model, line type indicates source of parameter estimates.

The final proportion of infected individuals who will develop disease in model *i*, *Pi* is given by:

Steady state incidence

Figure A2 shows the steady state TB incidence for each model as a function of the contact parameter, *β*.

****

**Figure A2. Steady state TB incidence as a function of the contact parameter, *β*.** Left: results using parameter set A; right: results using parameter set B. Colours indicate the different models. Dashed lines show the results when re-infection is not included in the model. Note, results for models 1 and 2 overlap.

In the full model including re-infection (solid lines in figure 3) model 3 results in the highest incidence at a given value of *β* due to the higher lifetime risk of developing TB following infection (see figure A1). At low values of *β* models 1, 2 give the same incidence. However, as the contact rate, *β*, increases the model predictions diverge with a lower incidence predicted by model 1 at a given value of *β*. This divergence occurs due to differences in the risk of reinfection. In model 1 all individuals spend some time in the “fast” latent state where they are not at risk of reinfection and therefore the incidence of disease is lower. This role of reinfection can be seen by contrasting with the results from a model with no reinfection (dashed lines in figure 3). In this case models 1 and 2 predict the same incidence for all values of *β*.

The same qualitative patterns are observed for both parameter sources but the divergence at higher incidence is smaller when using parameter set B.

**A.4. Non-monotonic relationship between NNT and steady state incidence**

Previous analysis of a simple model of preventive therapy [25] found a non-monotonic relationship between baseline incidence and NNT using model structure 1 with a minimum NNT occurring at an incidence in the region of 700/100,000. In this analysis we did not observe this non-monotonic behaviour within the range of incidence explored. The key difference between the parameterisations of model 1 here and the analysis in [25] is the duration of the fast latent state which is determined by *e*, the rate of movement from the fast latent state to the slow latent state. In [25] this was assumed to be 5 years while in the parameterisations used here it is much shorter (between 0.25 and 1.15 years).

To explore this, we compared model 1 with values of *e* of 1, 0.5, 0.3̇33, 0.25 and 0.2 which give durations of the fast latent state of 1 to 5 years. We set the rate of progression from the slow latent state *c* = 5.94x10-4 and calculated the value of *k* (the rate of progression from the fast latent state) to give a lifetime cumulative risk of TB of 0.11, the same as with the parameters estimated in Menzies et al. [17].

Figure A3 shows the predicted outputs of the simulated preventive therapy intervention for each duration of the “fast” latent state.

This shows that the non-monotonic behaviour is observed but that the “optimum” incidence (i.e. the value at which the NNT is minimised) is dependent on the duration of the fast latent state. With a duration of 5 years (yellow line) the minimum NNT occurs at an incidence of 1008/100,000. However, with a duration of 1 year (grey line) the minimum NNT occurs at a much higher incidence of 3064/100,000.

****

**Figure A3. Predicted impact of preventive therapy in model 1 with different durations of “fast” latency.** Top left: Percentage reduction in TB incidence from steady state equilibrium; top right: cumulative number given preventive therapy (assuming a constant population size of 10,000); bottom left: cumulative number of TB cases averted; bottom right: average number needed to treat with preventive therapy to avert one case of TB. Colours indicate the different durations of the “fast” latent state.

**A.5. Re-parameterisation of model 3**

The original parameterisation of model 3 results in an increased lifetime risk of TB compared to models 1 and 2 and a bigger impact of preventive therapy.

Here we explored the effect of re-parameterising model 3 to give the same cumulative probability of TB as models 1 and 2.

Rearranging equation A46 we can obtain an expression for the value of *a*, the proportion progressing directly to disease in model 3 in terms of *c* and the probability of TB (*P3*):

Setting *c* and *P3* to the values for models 1 and 2 we can calculate the value of *a* required to give the same cumulative probability of TB in all 3 models. This gives new values for *a* of 0.084 for Menzies and 0.090 for Ragonnet.

Figure A4 shows the predicted impact of the preventive therapy intervention for models 1 and 2 and the original and new parameterisation of model 3.

We find that, in contrast to our main finding the re-parameterised version of model 3 results in a lower predicted impact. This is because it is not possible to directly prevent fast progression to disease in this model structure by providing preventive therapy to the latent populations; these cases do not pass through a “fast” latent state where they can be treated with preventive therapy.



**Figure A4. Predicted impact of preventive therapy with re-parameterisation of model 3.**