**TB Infection – TB Disease Model Overview, Based on Philippines Population**

Sarah Weber, supervised by Bob Horsburgh and Laura White

Last Updated August 13, 2024

PARAMETERS:

Static Parameters:

* Probability of progression to disease: 0.025 for new infections, 0.01 for one year infections, 0.005 for two to five year infections and 0.00075 for ‘old’ infections greater than five years old
* Life expectancy at birth: 69.7 years. The life expectancy at birth in the Philippines was 68.77 in 2001 and 70.55 in 2015. The value used in this model is the average of the 2001 and 2015 estimates.
* Annual mortality: 0.00563 or 5.63 per 1000 persons. The annual mortality was 5.47 per 1000 in 2001 and 5.797 in 2015. The value used in this model is the average of the 2001 and 2015 estimates.

Varying parameters (R Shiny widgets):

* Starting population size: defaults to one million for simplicity, but can be changed to any value up to 120 million which is the approximate population size of the Philippines in 2024.
* Start and end years: defaults to 2001 and 2016, because our static parameters are grounded within this time period. However, these may be changed to any years from 1990-2030. The duration of the model is also flexible.
* ARI: defaults to 0.037, which is the ARI we found to be closest to Tupasi (see ‘Identifying ARI with lowest MSE with Tupasi’ above). The ARI can be any value between 0.005 and 0.05.
* Starting prevalence of infection: defaults to 0.634 (63.4%) which is the prevalence found by Tupasi. This value can range from 0 to 1.
* Proportion of infections that are old for starting population: defaults to 0.61 (61%) which is the proportion of infections estimated by the expected prevalence at median age of 25 for our default ARI of 0.037 (prevalence = 1-(1-0.037)^25). This value can range from 0 to 1.
* Proportion of those with disease who are treatment cures: default is 0.47 (47%) which was derived from Bob’s estimates. From 2007 to 2016, the average TB disease incidence was 372,000 and average number cured was 175,000 (175000/372000=0.47). This value can range from 0 to 1.
* Rates for death and self-cure from prevalent disease: 6.62% of persons with disease each year in each group. We estimate that after about four years with untreated disease, half of individuals self-cure and half die. Averaging over a four year period, this proportion is calculated by ((1- tx cure rate)/2)/4

MODEL GENERATION STEPS:

Data preparation:

1. Empty matrix is created with one row per year from start year to end year specified in varying parameters.

First year of model:

1. Starting population for the first year of the model is specified (based on parameter selection, and defaults to one million).
2. Total infected set to be the population size\*starting infection prevalence.
3. New infections set to be ARI% of the population
4. Of the existing infections (63.4% total infected – 3.7% new infections), 61% of them are old or 6+ year duration. This is the default value based on expected prevalence at the median age (see parameter overview above); however, this value may be changed in R shiny widget.
5. The remaining infections needed to reach 63.4% of the population infected are equally distributed across the 1-5 year infection duration groups.
6. All individuals without an infection are considered uninfected (other 36.6% of population).
7. Infected deaths are calculated via population size\*mortality rate\*(1-(1-ARI)age of life expectancy). Using the mortality rate and life expectancy values specified above, infected deaths are population size\*0.00563\*(1-(1-0.037)69.7).
8. Uninfected deaths are calculated by subtracting infected deaths from total deaths.
9. No one progresses to disease in the first year. All disease generated from the first year’s infections are considered disease cases in year 2.

For each subsequent year…

1. The disease treatment cures and those with disease who self-cured (6.62% of prevalent disease) from the previous year are added back into the model as new infections in the current year.
2. The infected deaths from the previous year are removed from the current year’s infection groups. This removal is proportional to group size by infection time (i.e. a much higher proportion of deaths are removed from the old infections compared to new or 1-5 year infections).
3. The population at risk of infection at the beginning of the year (before replacements) is the sum of all infected and uninfected in the previous year. Therefore, the population at risk should be the population from the previous year excluding those with disease.
4. New infections are generated by applying the ARI to uninfected and current infections (1-5 year and 6+ year infections).
5. Infections of 1-5 year durations are generated by advancing those infected without a reinfection or disease progression to the next year (e.g. those with a 3-year infection in previous year who were not newly infected or progressed to disease move on to become a 4-year infection).
6. Old infections (6+ years) are a sum of old infections in previous year and those who advanced from a 5 year old infection.
7. Disease progression is computed per probability of progression to disease conditional on years since infection outlined in static parameters.
8. Infected and uninfected deaths are calculated as per steps 8 and 9 above.
9. Disease progression is computed per probability of progression to disease conditional on years since infection outlined in static parameters.
10. Group of everyone with prevalent disease is the sum of incident disease and prevalent disease from previous year, excluding those with treatment cure, death from disease and self-cure (amounts to 39.7% of disease in prior year because 47% are treatment cure, 6.62% die, and 6.62% self-cure).
11. Current year’s treatment cures are 47% of all disease calculated in step 19. This proportion may be changed in R shiny widget.
12. Death from prevalent disease and self-cures in the current year are each 6.62% of all disease calculated in step 19.
13. Uninfected population is uninfected from previous year who were not infected (1-ARI), plus all deaths from previous year (infected deaths, uninfected deaths, deaths from prevalent disease).
14. Lastly, total population is sum of infected, uninfected, and those who progress to disease the following year. While the population at risk may decrease as we generate cases of disease (esp in first few years), the total population should remain stable if replacements are balancing individuals being removed and added back into model each year.

Other summary statistics calculated after model generation…

- Annual infection incidence: proportion of the population at risk with incident infection

- Total number of people infected for each year

- Proportion of the population at risk who is infected

- Proportion of the population at risk who is uninfected

- Percent change in the number of people infected each year

- Disease incidence: proportion of the population at risk with incident disease