Optima TB

User Guide

Optima Consortium for Decision Science

September 2018

Table of Contents

Abbreviations

ART	antiretroviral therapy
BCG	Bacillus Calmette–Guérin vaccine
CMT	comorbidity treatment
DS-TB	drug susceptible tuberculosis
DR-TB	drug resistant tuberculosis
GLC	Green Light Committee
HIV	human immunodeficiency virus
LTBI	latent tuberculosis infection
MDR-TB	multidrug-resistant tuberculosis
NTP	National Tuberculosis Program
PLHIV	people living with HIV
SN	smear positive
SP	smear negative
TB	tuberculosis
UNDP	United Nations Development Programme
WHO	World Health Organization
XDR-TB	extensively drug-resistant tuberculosis

Overview

This document is a guide to setting up and conducting an Optima TB project. The skills and familiarity gained will allow users to generate customized Optima TB models to address policy questions towards accelerating the efficient and effective response to Tuberculosis by optimizing the allocation of available funds.

Working through practical examples, some of the most common questions posed by decision makers and planners when using Optima TB to inform their TB response will be addressed.

This guide is structured into the following modules:

1. **Introduction to Optima TB**: Overview of Optima TB, including model structure
2. **Creating an Optima TB project**: preliminary steps for generating an Optima TB project and providing an overview of data inputs.
3. **Entering data and estimates**: description of the types of data and estimates required by the model and where to source these values.
4. **Calibration**: the process of fitting model parameters to historical and latest reported data and setting a baseline for scenario and optimization analyses.
5. **Programs and cost functions**: defining TB programs and cost functions which define the relationship between program spending and coverage, and population-specific outcomes.
6. **Scenarios**: setting-up and running scenario analyses.
7. **Optimizations**: setting-up and running optimization analyses.

1. Introduction to Optima TB

1.1 Optima TB model schema

Optima TB is a dynamic compartmental population-based model which reflects the current understanding of the natural history of TB. The population is divided into compartments based on health states across the TB care cascade. The TB care cascade consists of the pre-infective state, the latent infective state, and the active infection state. Within these compartments, several stages are possible depending on (e.g.) drug susceptibility of TB strain. At each point in time, people can move between health states, with transitions represented as transfer rates between compartments.

A representation of the model schema is depicted in figure 1 and 2. The square boxes represent the model compartments, with the arrows indicating the movement between compartments (with a dotted line representing treatment failure). Green ovals indicate junctions that allocate the new active TB cases based on smear status and drug resistance status. Figure 2 provides a greater focus on treatment.

Figure 1: The Optima TB model schema

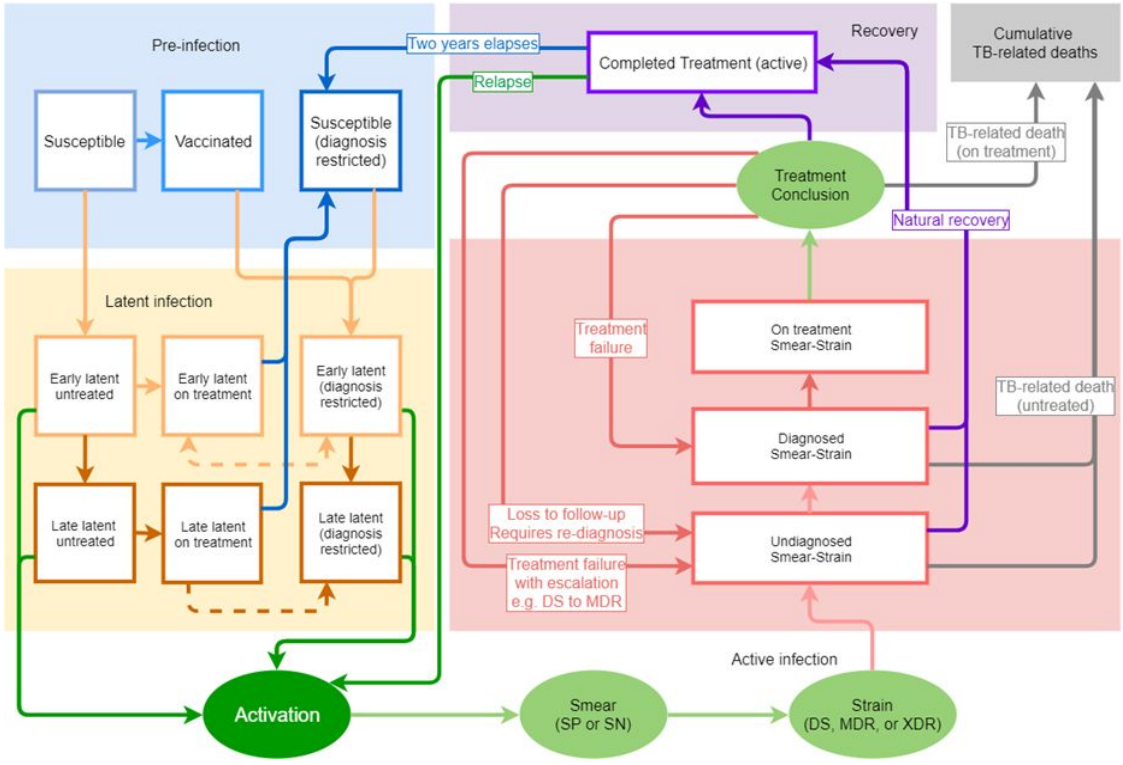
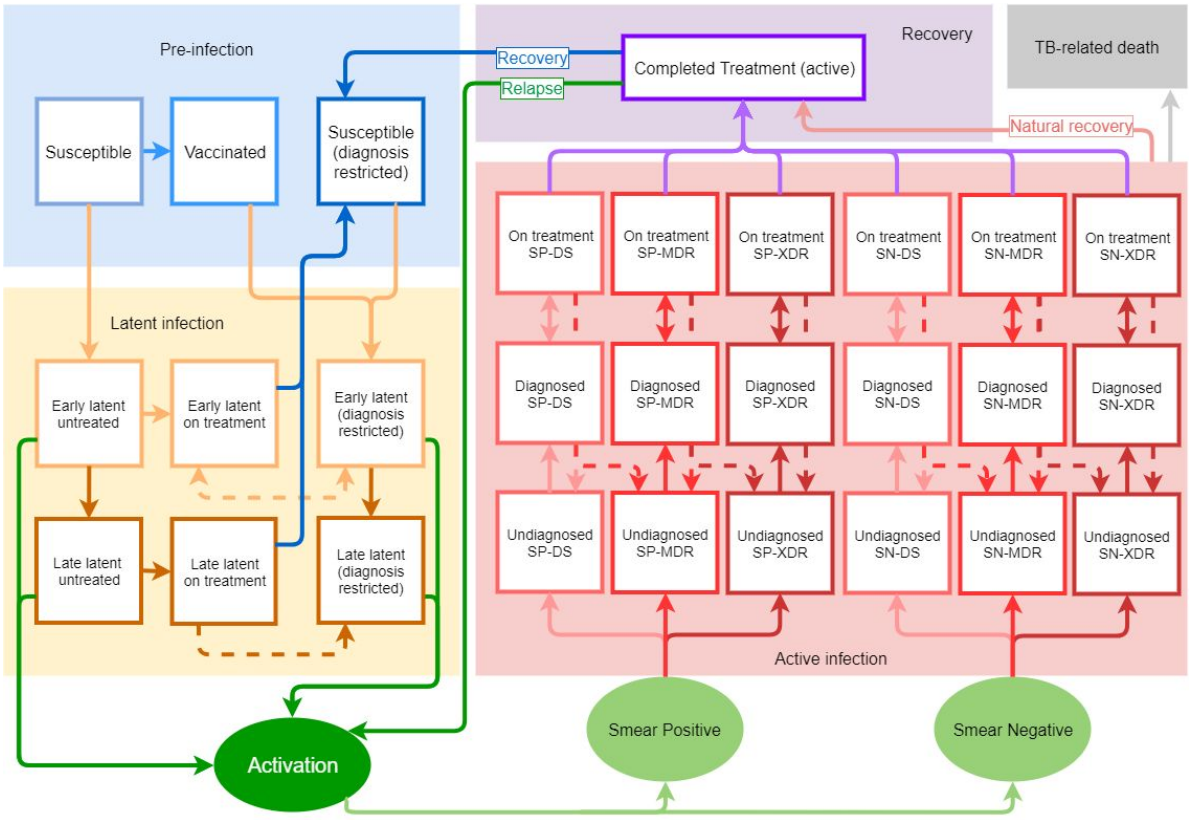
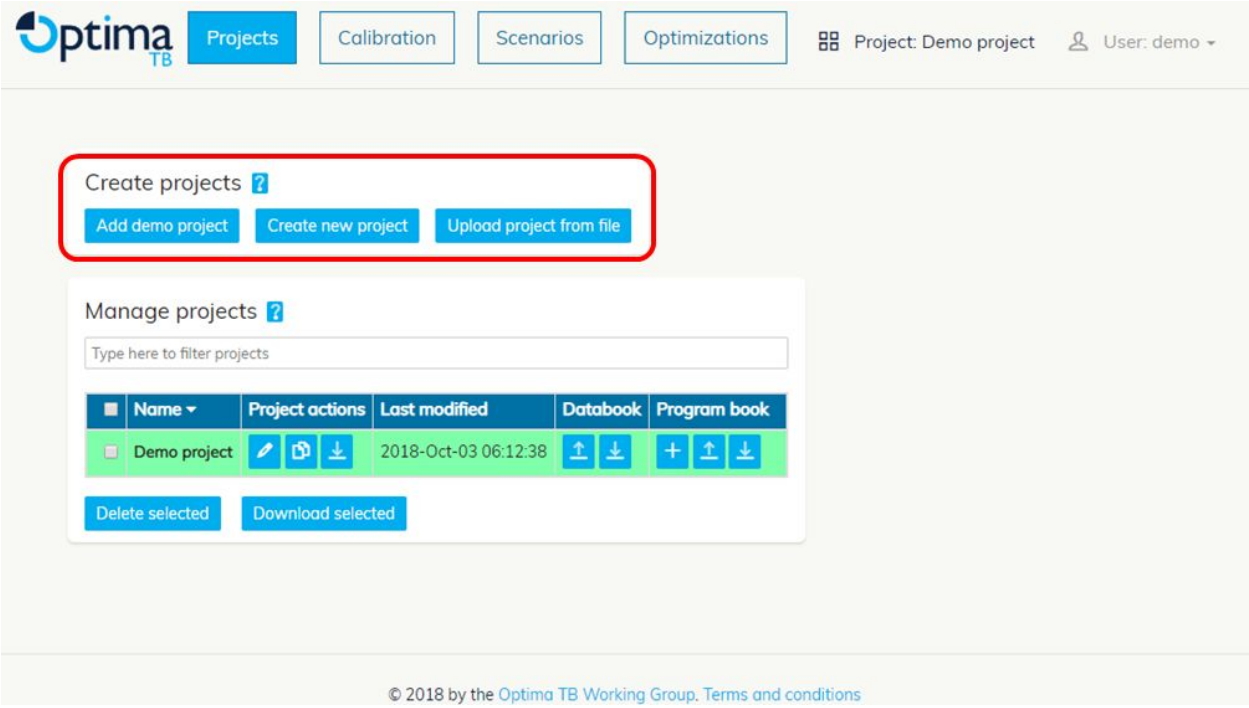


Figure 2: The Optima TB model schema with focus on treatment



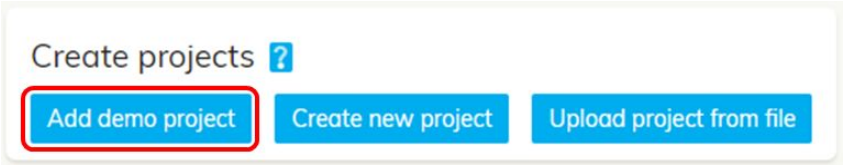
2. Creating an Optima TB project

Optima TB analyses are conducted using model **projects.** To begin an analysis, you must first create a project. To create a project within the Optima TB tool navigate to http://tb.ocds.co/, register for a new account if you have not created one before. Then login and select on the 'Projects' page one of the three options (adding demo project, creating new project, or uploading an existing project).



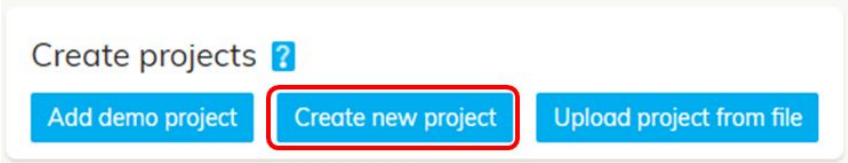
2.1. Add demo project

Optima TB comes loaded with several demonstration projects. Click "Add demo project" to select and view these. Selecting one of these demo projects means that the project will be copied from our internal database into your own personal account. This means that you can modify the project and your changes will only be made within your account, i.e., the original demonstration project from the database will remain intact. Once the project is loaded in your account, you can edit the project as outlined in section 2.4.1.



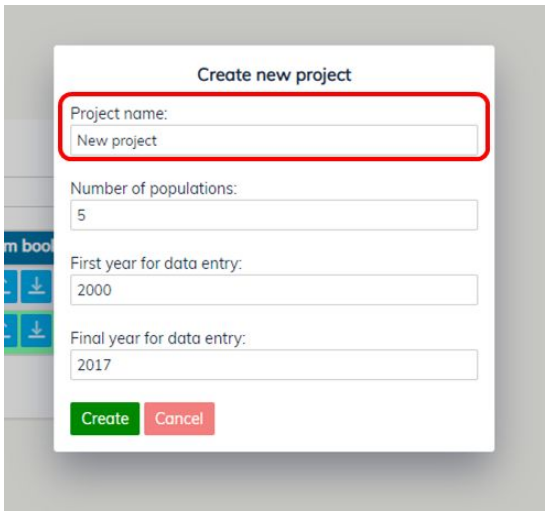
2.2. Create new project

Clicking “Create new project” will launch a pop-up box where you can enter information about the project that you would like to create.



2.2.1. Project name

The project name should be detailed enough so that you will remember what the project is for. Keep in mind that you may eventually wish to create multiple similar projects to explore different choices of populations and programs.This information will be provided separately so the project name does not need to explicitly reflect this, but you should still be as descriptive as possible. For example, if you were creating a detailed project to explore optimal allocations in Ruritania, a good project name might be “Full Ruritania model for optimal allocations” (NB: “Ruritania” would be a bad project name, unless you are quite sure that you will only have a single project for Ruritania).

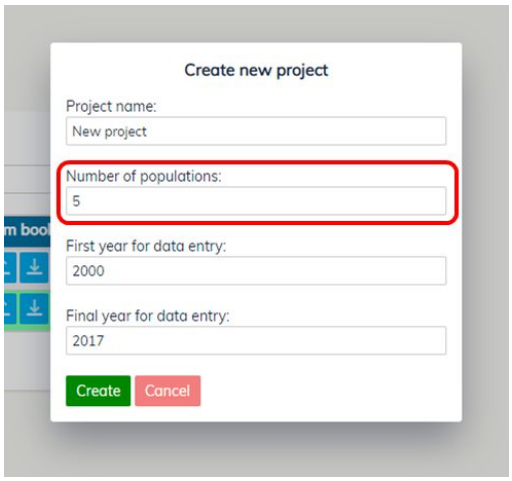


2.2.2. Determining populations

The flexibility of Optima TB allows populations to be broken down into smaller groups. A more granular population breakdown may be desirable, but is ultimately limited by data availability. A typical analysis might breakdown the general population by age, sex, and/or geography, depending on the particular characteristics of the epidemic and the response, but depends on

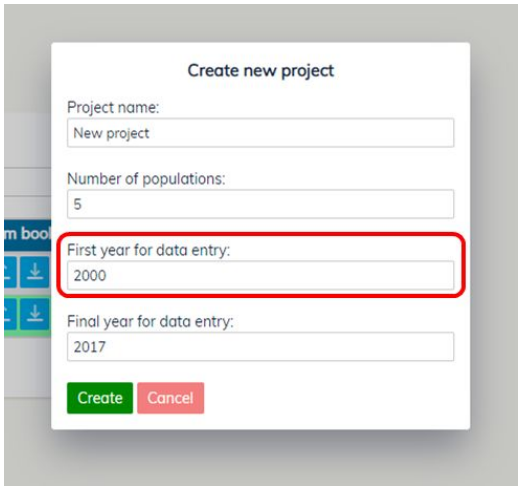
data availability. Co epidemics (such as, people living with HIV or diabetes) can be modelled as separate population groups within Optima TB to best represent the epidemic.

During the project creation stage, you need only specify the number of populations - further population specific data will be entered in the Optima TB databook.



2.2.3. First year for data entry

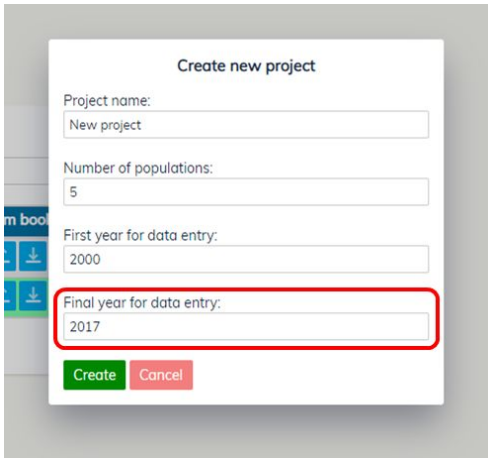
Most of the data to be entered in the spreadsheet is entered year by year. The first year of data entry determines what years will be available to enter data into.



2.2.4. Final year for data entry

The final year of data entry determine what years will be available to enter data into.
For example, if “2010” is entered as an initial year and “2018” for a final year, there will be 25 columns in which to enter data (2010 to 2018 inclusive). There is no problem with entering a year later than the last year for which data are available for. For example, you may currently

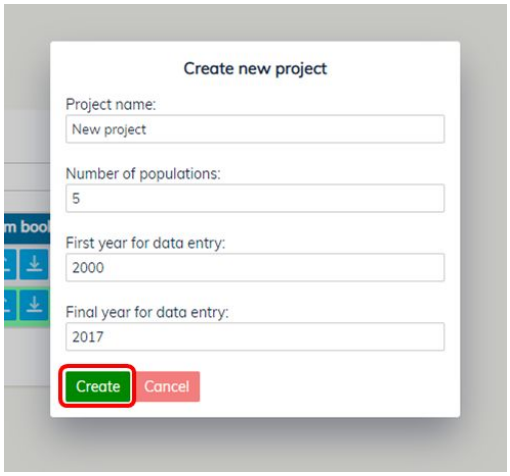
only have 2016 data, but are expecting 2017 data soon, and intend to repeat the analysis in 12 months when 2018 data become available. In this case, it would be advisable to enter (at least) “2018” as the final year, so you do not have to create a new spreadsheet to enter 2017 and 2018 data when they become available. In the meantime, these columns can be left blank. You may also wish to add demographic projects that extend beyond beyond current years, it this case it may be advisable to include .



2.2.5. Create project

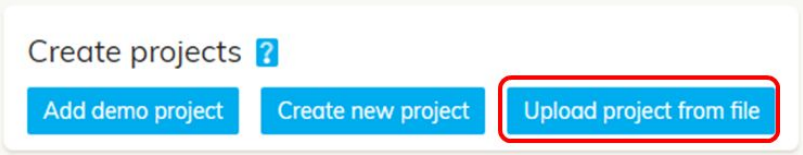
Once you have entered a project name, a number of populations and a first and final year for data entry, you will be able to click on the “Create” button. This will:

1. Add the current project to the list of available projects associated with the account.
2. Bring up a file save dialog box with the freshly generated blank input spreadsheet, known as the **databook**. This databook will be used to input data used for the Optima TB analyses.



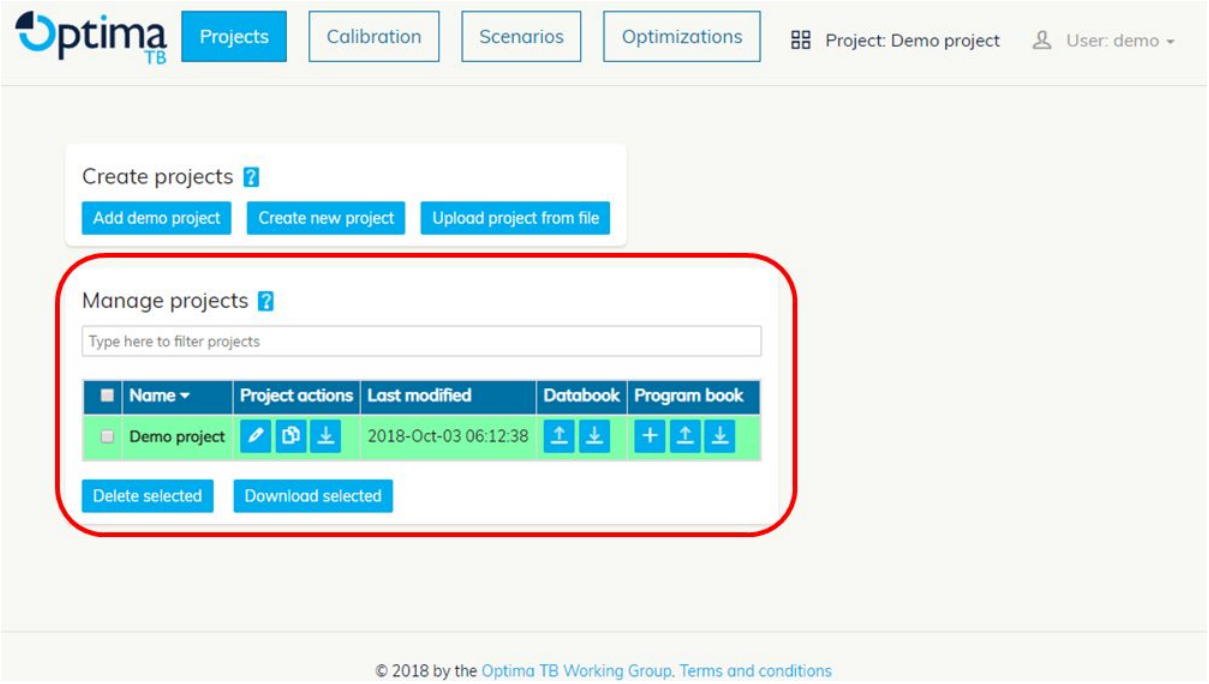
2.3. Upload project from file

Clicking this button launches a file upload box, which will allow you to navigate to where you have saved a project file. This option is useful when sharing work with collaborators. Optima TB projects have a file extension “PROJECTNAME.prj”. Find the project that you wish to load to your account and select it for upload. This will add the current project to the list of available projects associated with the account.

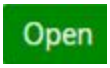


2.4. Project management options

Once you have a project loaded on your account, you will be able to see them in a table on the Projects page. Each row of the table corresponds to a single project. Here, you will find basic functions to copy, rename and delete projects.



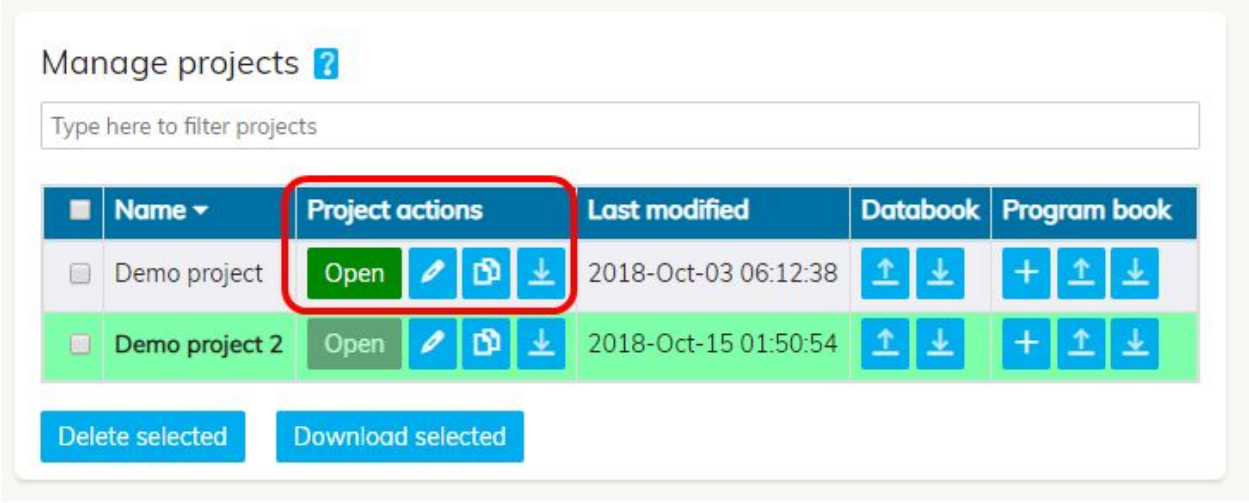
2.4.1. Project actions

Clicking the green “Open” 

button means that this is now the active project; navigating to other pages such as Calibration, Scenarios, or Optimizations will display results and information pertaining to the active project.

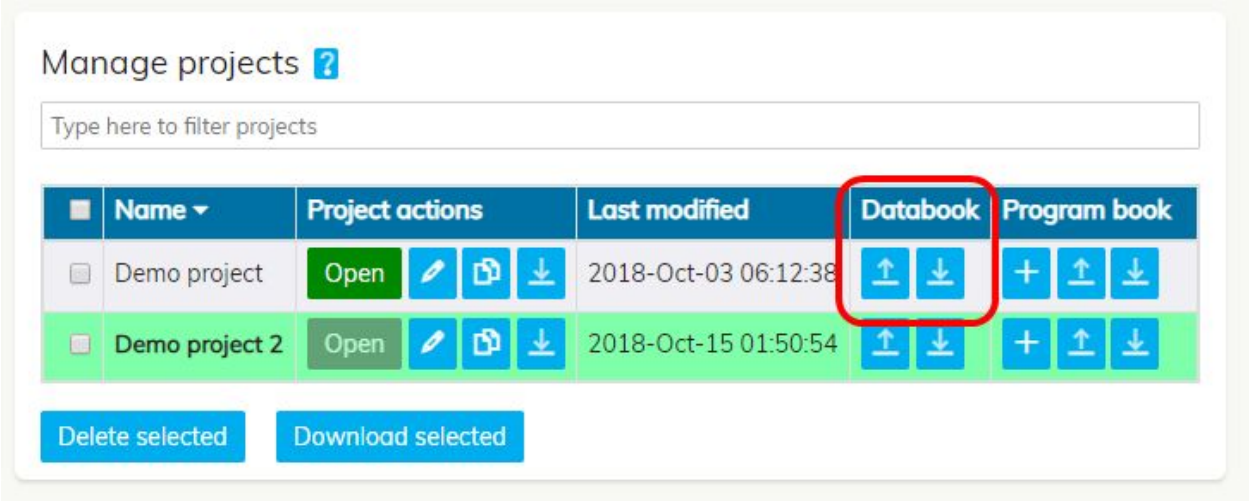
Clicking the “Rename” button will allow you to edit the name of the project, and clicking “Copy” will create a copy of your project. There is also an option to download a project. If you wish to share a project with a collaborator, the easiest way to do so is to download the project. There are two options for doing this:

1. The “Download” button will download your project as a project file, with extension .prj. This file contains all of the specifications that are used to generate results (for example, it would contain the details of any populations, parameters, programs, cost data, scenarios, and optimizations that you set up).
2. The second options for downloading projects is to select the checkbox in the left-most column of the row, then click the “Download selected” button at the bottom of the table. You can also download more than one project at a time by selecting multiple at once. Downloading multiple projects will put the projects together into a zipped folder.



2.4.2.Data entry sheets

There are two core data entry points for the Optima TB. These are divided up into two separate Excel workbooks. The first of these is referred to as the databook, and is intended for gathering data on the baseline state of the TB epidemic. The second is referred to as the program book, and is intended for gathering data on the interventions that are in place. The Databook is covered in section 3 and the program book will covered in section 4.



2.4.3. Databook management

Before an Optima TB analysis can be conducted, the databook must be downloaded, populated with data and uploaded. Details on filling out the databook spreadsheets are provided in Section 3 of this user guide. To upload a completed databook, click on “Upload” in the column “Databook” of the project table. This will bring up a file open dialog box, where you can select the file you wish to upload.

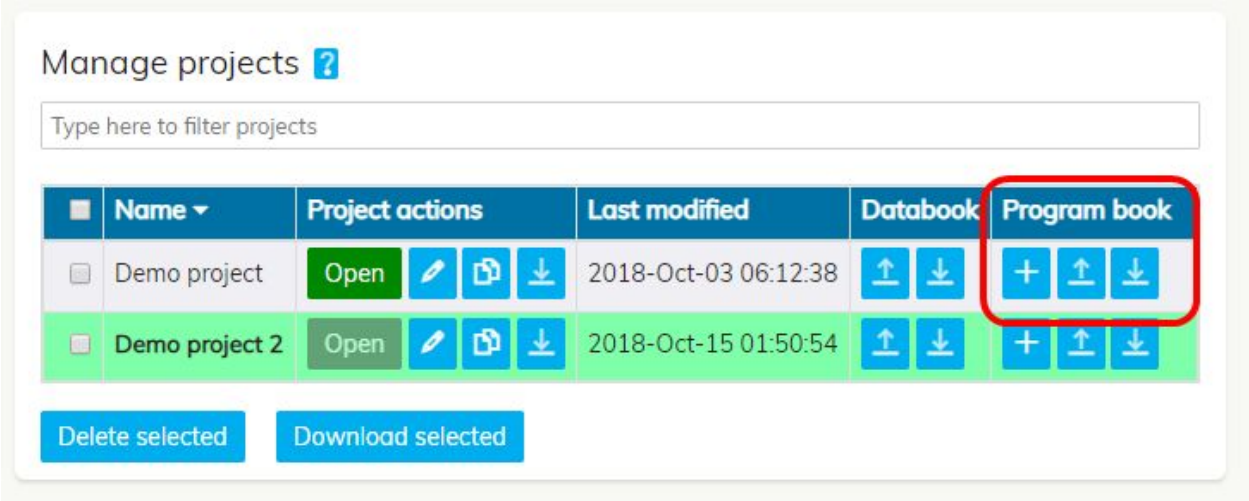
If you need to update the data inputs at any stage of the analyses, you can use this option to re-upload a modified version of the databook.

Once a databook has been uploaded, if you wish to view the databook for a given project, click on “Download” button in the column “Databook” of the project table. This will bring up a file open dialog box, where you can save the file you wish to view.

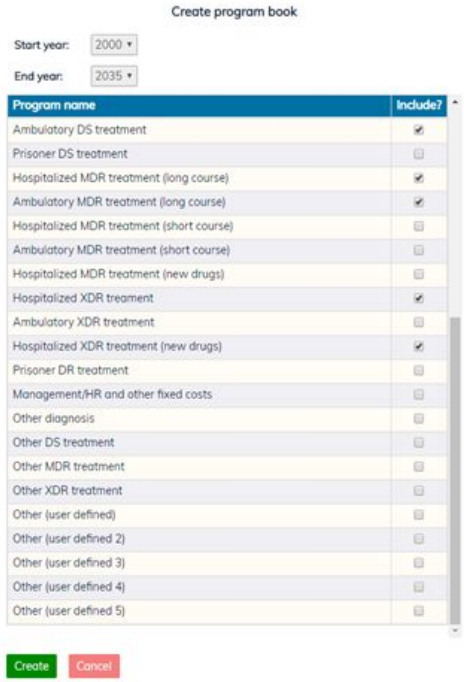
1. Upon uploading the databook to the project, values entered within data entry cells (those cells shaded in blue), parameter values, and comments that are directly associated with those cells will be saved within the project. Values and comments entered in cells other than data entry cells (shaded in blue or grey) as well as formulas used to generate the values, additional sheets will not be saved within the project.
2. If the databook is subsequently downloaded from the project, calculation cells (ie, formulas) and comments that were not associated with data entry cells will not appear.If you want to modify the databook, it is therefore better to edit the primary version (the comprehensive masterfile with all you additional working-out informations) and not a version which has been uploaded and again downloaded.

2.4.4. Program book management

The program book is essential in order to conduct any analysis of programs or interventions in Optima TB.



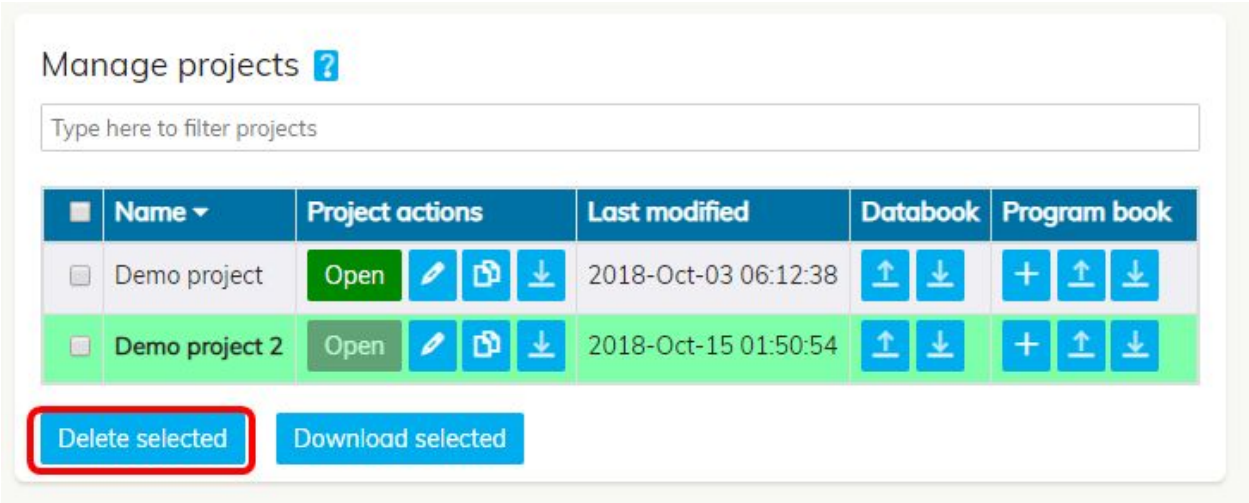
To create a program book, click on "+" (new)  in the column "Program book" of the project table, and select the programs you wish to include from the list available. Additional programs can be included in the analysis using the 'Other (user defined)' programs in the list.. Details on filling out the program book spreadsheet are provided in Section 3.



To upload a completed program book, click on “Upload” in the column “Program book” of the project table. This will bring up a file open dialog box, where you can select the file you wish to upload. If you need to update the program data at any stage of the analyses, you can use this option to re-upload. If you wish to view the program databook for a given project, click on “Download” button in the column “Program book” of the project table. This will bring up a file open dialog box, where you can save the file you wish to download.

2.4.5. Deleting a project

Select the checkbox in the left-most column of the row, then click the “Delete selected” button at the bottom of the table. You can also delete more than one project at a time by selecting multiple at once.



3. Entering data in the Databook

3.1. Introduction

The databook is intended for gathering data on the baseline state of the TB epidemic. Data is input into the databook, before being uploaded into the Optima TB tool to generate the model. In general, the data required to run the model includes;

- Population definitions
- General demographics
    - Population size
    - Birth and death rates
    - Oher population rates as necessary (such as emigration, incarceration)
- Notification data (used to calculate proportional breakdowns)
- Treatment data
- Active and Latent TB prevalence estimates
- Vaccination data
- TB death rate
    - On treatment/off treatment

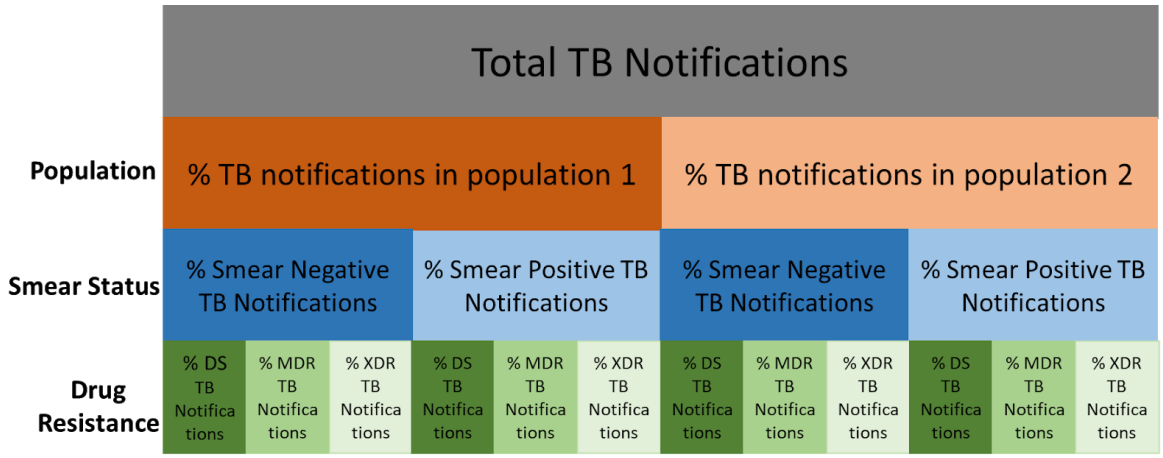
For data to be input into the model the data must be disaggregated by:

- Population,
- Smear status, and
- Drug resistance of the TB strain.

In many contexts it is very unlikely that available data will be disaggregated to this level, so estimates will have to be calculated. A small secondary analysis of TB notification data will go some way towards the required data elements.

As notification data is where there is likely the most confidence, it is used to calculate a proportional breakdown of notifications, that can be applied to other data for disaggregation. For example, a prevalence estimate for active TB cases can be used to calculate the number of cases of active TB (by multiplying by the total population size). It can then be further broken down into necessary groups by applying the proportional breakdown calculated in the notification data.

Therefore, the first step in data collection is calculating proportional breakdown of notification data. The values will be applied to data to estimate the distribution of TB cases within the population. The breakdown is highlighted in the image below.



When calculating and applying the proportions care should be taken to ensure the values seem reasonable and data is complete. Extrapolation may be required for incomplete data and care should be taken to ensure that any estimates are reasonable.

3.1.1. Documenting data sources in the databook

Documenting the source for every piece of data that is loaded into an Optima project from a databook is vital for justifying results, ensuring that a calibration matches the best quality data with priority compared with assumption data, and incorporating newly available data in future.

Every data cell should include a comment within the Excel databook, and include the following information:

- **Source**: a reference for the data .
- **Notes**: highlighting the quality of the data (such as sample size or confidence bounds), and any assumptions and/or calculations.

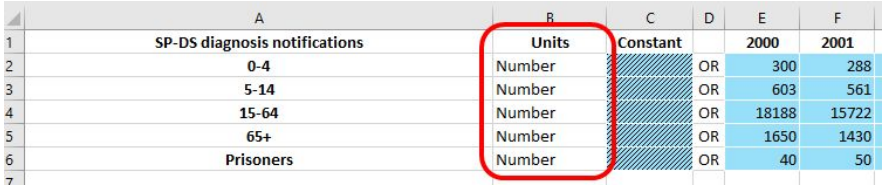
Comments can either be written on a cell by cell basis or in the first data cell of the row or sheet for which the comment applies. In cases when calculations such as programme costing are too extensive to be summarised as databook comments, it may be beneficial to use ‘working’ datasheet(s) to be stored in a subfolder where the databook is saved and referenced as part of the source, where appropriate. Examples of documenting the data sources are outlined below:

Assumption: value **Notes:** Default Optima value	**Source:** unit costs calculated using bottom-up micro-costings see …/Working files/projectname_costing.xlsx
Source: Paper, Author, Year: value A Report, Author, Year: value B **Notes**: Paper was a very small study in one town of the country, used Value B from Report as this was a national study on a large cohort.	**Source**: Report, Author, Year: value **Notes:** primary research – testing rates based on survey of XYZ people across one province. Assumed to be representative, but likely an overestimate given that this province includes the most accessible urban populations.

3.1.2 Data entry formats in the data book

The Optima TB model require data to be entered in a number of different formats, including proportions, probabilities and as whole numbers. Care should be taken to ensure these values are entered correctly to avoid any errors.

- **Proportions:** these values should be interpreted as a proportion of outcomes that are grouped together and should sum to 1, such as the proportions of people who have different treatment outcomes.
- **Probability:** this refers to an annual probability of an outcome occurring at least once during that year, such as the annual probability that latent TB progresses to active TB.
- **Number**: input values will be used directly, such as the number of people initiating treatment in a given year.



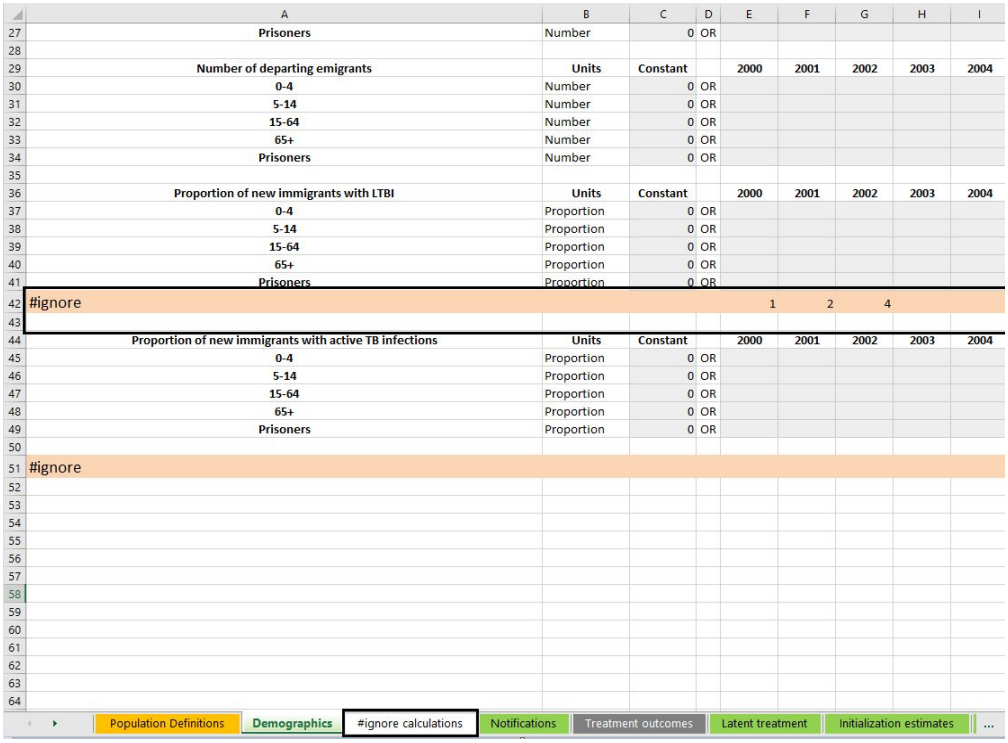
	A	B	C	D	E	F
1	SP-DS diagnosis notifications	Units	Constant		2000	2001
2	0-4	Number		OR	300	288
3	5-14	Number		OR	603	561
4	15-64	Number		OR	18188	15722
5	65+	Number		OR	1650	1430
6	Prisoners	Number		OR	40	50

3.1.2. Entering data outside of specified fields

In some cases, it may be useful to add data to the databook in cells outside of data entry cells (blue cells). This may be useful for calculations or to compare data. Data cannot be entered into the white cells unless a special tag '#ignore' is entered into the first column for a row in which additional data is entered. Data entry tables are separated by blank rows, and a row with

`#ignore` does not count as a blank row (because it is treated as non-existent) so for instance two tables cannot be separated by a `#ignore` line, there would need to be an additional blank row too.

In addition to this, to add additional tabs, you must include '#ignore' in the tab name, for example if a tab is named "#ignore calculations", Optima TB will ignore the whole tab. Without this, you will encounter an error when uploading the databook.



	A	B	C	D	E	F	G	H	I
27	Prisoners	Number	0	OR					
28									
29	Number of departing emigrants	Units	Constant		2000	2001	2002	2003	2004
30	0-4	Number	0	OR					
31	5-14	Number	0	OR					
32	15-64	Number	0	OR					
33	65+	Number	0	OR					
34	Prisoners	Number	0	OR					
35									
36	Proportion of new immigrants with LTBI	Units	Constant		2000	2001	2002	2003	2004
37	0-4	Proportion	0	OR					
38	5-14	Proportion	0	OR					
39	15-64	Proportion	0	OR					
40	65+	Proportion	0	OR					
41	Prisoners	Proportion	0	OR					
42	#ignore				1	2	4		
43									
44	Proportion of new immigrants with active TB infections	Units	Constant		2000	2001	2002	2003	2004
45	0-4	Proportion	0	OR					
46	5-14	Proportion	0	OR					
47	15-64	Proportion	0	OR					
48	65+	Proportion	0	OR					
49	Prisoners	Proportion	0	OR					
50									
51	#ignore								
52									
53									
54									
55									
56									
57									
58									
59									
60									
61									
62									
63									
64									

Population Definitions | Demographics | #ignore calculations | Notifications | Treatment outcomes | Latent treatment | Initialization estimates | ...

3.1.3. Notes for entering any data in the databook

- Data input cells are shaded as light blue for data that is generally expected to be available for a given country, such as the number of notifications, and grey for cells that contain global default values or data that was previously entered in a databook that was saved and re-downloaded, and in either case may not require modification.
- How the model handles missing values.
    a. Interpolation: if there are missing values between two existing values over time, the model will interpolate within two known values in a sequence of values
    b. Extrapolation: if there are missing value(s) in a string of values, values are assumed into the future or the past by duplicating the adjacent known value accordingly:

        For example, if data points over time were entered as:

_ _ _ 7 _ _ 10 _ 16 _
The model would interpret these values as:
7 7 7 7 8 9 10 13 16 16

c. For every row of data, there must be at least one value entered for either an annual value or as an assumption. If data is added to both the assumptions and annual values, any data added to annual values will override the assumption data.

- Formulas exist in certain cells in the databook. These formulas calculate values based on data entered as numbers. A user may add additional formulas in databook cells. Calculated values from these additional formulas will be imported into the project as numbers.

3.2. Demographics

Demographics from a certain setting (country, subnational region, or global region) within the Optima TB model are calculated for each key population using:

- initial population size,
- annual birth rates or numbers,
- annual non-TB death rates,
- transfers in and out of population groups due to aging, and
- other transfers in and out of population groups due to non-aging effects such as migration or incarceration (as specified for a given project).

Populations can be defined by certain characteristics. Transfers can occur between population groups, that are to account for:

- ageing, where a proportion of one population group is transferred to the older population group as specified by the aging matrix. Unless otherwise specified, the model assumes a uniform age distribution for general and key population groups;
- other factors including, risk-related and comorbidity related transfers:
    - incarceration and release rates for prisoner populations,
    - employment turnover for key populations defined by their occupation (eg, miners or healthcare workers),
    - incidence and treatment outcome rates for comorbidities such as HIV or diabetes.
        - Rates for these transfers are calculated using known values.

As demographic data are known with considerable confidence, model output should reflect this data with minimal modification.

3.2.1. Indicators

Indicator	Population Size
Format	Number
Source	United Nations Development Programme (UNDP) reports Country or context-specific reports
Comments	The total population size should be equal to the total population for a given context (eg, country), double-counting should be avoided.

Indicator	Number of births/Birth rate
Format	Number/Proportion
Source	United Nations Development Programme (UNDP) reports Country or context-specific reports
Comments	The model requires births to be input as a number. If a 'crude birth rate' is available as is often the case, then the number of births each year can be readily calculated from population size * crude birth rate.

Indicator	Number of new immigrants
Format	Number
Source	Country Reports
Comments	Default value is 0, unless otherwise specified This could reflect migration from other countries or regions that are not part of the study directly, or could be used in conjunction with an internal population group to reflect regular economic migration and return from a country.

Indicator	Number of departing emigrants
Format	Number
Source	Country Reports
Comments	Default value is 0, unless otherwise specified Departing emigrants are assumed to be representative of their population group in terms of TB status.

Indicator	Proportion of new immigrants with LTBI
Format	proportion
Source	
Comments	Default value is 0, unless otherwise specified

Indicator	Proportion of new immigrants with active TB infections
Format	Number
Source	
Comments	Default value is 0, unless otherwise specified

3.3. Notifications

Notification data entered in the databook is used to calculate proportional breakdown of data, rather than as direct input data. Data must be disaggregated by:

- smear positive versus smear negative cases, per population
- drug-resistant strains per smear status, per population
    - Rifampicin resistant TB is counted as part of multi-drug resistant TB.

Each notification data value should add to the total of all notifications over a year of notified cases. When calculating the breakdown of data it is useful to go with notified empirical data first (strain-specific notifications by age band and HIV status) and then estimate further breakdowns into smear status by key populations.

As the notification data is used directly by the model as the number of diagnoses each year, it may be necessary to estimate the number of diagnoses on this sheet based on the rate of reporting in settings where reporting is not comprehensive.

1. Create a new sheet labeled '#ignore Calculate Notifications'.
2. Keep all original official data here. It is always useful to be able to refer back to this number later in case estimates of under-reporting change later in the process and in order to be able to confirm how the numbers were generated.
3. Apply whatever transformations are necessary to estimate the real number of diagnoses in the country.

For example, if surveys estimate that 50% of DS-TB cases are notified because of complex chains of reporting, then the notifications sheet might have a formula to multiply all the officially reported values for DS-TB by 2.

4. Similarly, if treatment happens with non-standard and unreported interventions, those should be included in the number of people initiating treatment to the best degree possible.

2.3.1. Indicators

Diagnosis notifications

Indicator	SP-DS diagnosis notifications SP-MDR diagnosis notifications SP-XDR diagnosis notifications SN-DS diagnosis notifications SN-MDR diagnosis notifications SN-XDR diagnosis notifications
Format	Number. Disaggregated by smear status and drug resistance.
Source	Surveillance database or country reports The World bank database Smear proportions: either supplied by the user as part of notifications data, or indicated by the user as a proportion per population or per drug-resistant strain type. Drug-resistant strains: as above, as well as from values from the literature for subnational and national cases. If these data are not available, consider using proxy values from regional reports. Information surrounding when drug-resistant strains first emerged in a given setting will be useful for an Optima TB analysis.
Comments	● If notifications data are available by populations, TB drug-resistant strain type, and TB smear status, verify that the sum for all notifications is correct . ○ Complete any missing values using a combination of: i. Common sense to determine missing values, and ii. WB/WHO setting specific data disaggregated by smear status (smear positive and smear negative). ● If not, then this will require intermediate work to calculate the full disaggregation by population, smear status, and drug-resistance. Typically, there is a heterogeneity between age-populations and smear

	status (with children typically having a bias towards smear negative[1]). Correspondingly: ○ if notifications have been supplied by smear proportion across all populations, in addition to disaggregation by population and drug-resistance: check that the notifications sum following step 1a above, and then directly disaggregate each disaggregated for smear status per population, ○ if notifications have been supplied smear by population, but no data on drug-resistant strains has been provided/is unavailable: look to alternative sources for information as to the disaggregation for DS:MDR:XDR by populations. If possible, also disaggregated by smear status. Once these ratios are known, check that the notifications sum following step 1a above, and then disaggregated by drug resistance per population, then disaggregated by smear status, and ○ if notifications have only been supplied by population, with no data on smear status or drug-resistant strains: identify the ratios for each population for smear status, and for each population for drug-resistance strain. Check that the totals match those across the populations. Disaggregation can be directly calculated disaggregated by drug-resistant strain type, and then by smear status.

Treatment Initiations

Indicator	DS treatment number of initiations MDR treatment number of initiations XDR treatment number of initiations
Format	Number. Disaggregated by drug resistance status
Source	Country reporting
Comments	

[1] This may be a testing bias, as young children have difficulty in producing sputum samples; correspondingly other tests (ie, TST, bacteriological confirmation) can/are recommended. Note that generally there tends to be an underreporting of TB in child populations in the applications that have been completed to date.

3.4. Treatment outcomes

These will all have defaults, but should be entered using cohort data. There are multiple options depending on the available data, from most accurate in terms of how the model uses these values to least:

1. cohort data for people **concluding** treatment in the specified year
2. cohort data for people **initiating** treatment in the specified year
3. **Annual** numbers for each, adjusted so that the proportions add to 1
4. Total known numbers dis-aggregated according to other estimates as best as possible.

3.4.1. Indicators

Note: as data is infrequently available, it is assumed that treatment outcomes for SP and SN are identical.

Treatment average duration (days)

Indicator	DS treatment average duration of completed treatment MDR treatment average duration of completed treatment XDR treatment average duration of completed treatment
Format	Days
Source	Country Data
Comments	Default values for each treatment duration, unless otherwise specified. This should reflect the average duration of *completed* courses of treatment using the mixture of modalities in each year for which there is data. The model will internally modify this value with the assumption that people who die from TB-related causes or who are lost to follow up depart on average halfway through treatment rather than completing the full course.

Treatment proportion of loss to follow up (require re-diagnosis)

Indicator	DS treatment proportion of loss to follow up (require re-diagnosis) MDR treatment number of loss to follow up (require re-diagnosis) XDR treatment number of loss to follow up (require re-diagnosis)
Format	Proportion

Source	Country Data
Comments	This reflects people who are lost to follow up during treatment, and would need to be re-diagnosed through a diagnosis program before they could re-initiate treatment.

Treatment proportion failed (no escalation, no need to re-diagnose)

Indicator	DS treatment proportion failed (no escalation, no need to re-diagnose) MDR treatment proportion failed (no escalation, no need to re-diagnosis) XDR treatment proportion failed (no escalation, no need to re-diagnose)
Format	Proportion
Source	Country Data
Comments	This reflects people who complete the course of treatment but have not been successfully treated, and could re-initiate a new course of treatment. If uncertain about whether cases are lost to follow up, failed without escalation, or failed with escalation, this is the appropriate default category.

Treatment proportion failed (escalation, require re-diagnosis)

Indicator	DS treatment proportion failed (escalation to MDR, require re-diagnosis) MDR treatment proportion failed (escalation to XDR, require re-diagnosis)
Format	Proportion
Source	Country Data
Comments	Countries often estimate that 2-3% of DS-TB cases might escalate in terms of drug resistance, but this could be flexible and calibrated depending on what proportion of drug-resistant cases are retreated cases.

Treatment proportion of treatments completed + success

Indicator	DS treatment proportion of treatments completed + success MDR treatment proportion of treatments completed + success XDR treatment proportion of treatments completed + success
Format	Proportion
Source	Country Data
Comments	In some cases treatment outcomes are dis-aggregated between “completed” (which may not include follow-up testing to confirm success) and “success”

	which does confirm sero-conversion, but in most cases these are aggregated together in reporting, and the model treats these as being identical.

Treatment proportion of deaths

Indicator	DS treatment proportion of deaths MDR treatment proportion of deaths XDR treatment proportion of deaths
Format	Proportion
Source	Country Data
Comments	This proportion just reflects the proportion of people with active TB who die from TB-related causes while on treatment and are reported, rather than those who may die from other causes, or who die from TB-related causes while not on treatment.

3.5. Latent treatment

Latent treatment values do not have to be completed, unless latent TB treatment program exists for a country. Most countries typically have no significant programs for latent TB, except for the BCG vaccination.

3.5.1. Indicators

TB vaccinations

Indicator	Number of vaccinations administered
Format	Number
Source	UNDP/World Health Organization (WHO) reports: for general vaccination rate. Country or context-specific reports: for information as to when BCG vaccination was introduced.
Comments	This type of vaccination is typically performed among children aged 0-4 years.

Latent Treatment

Indicator	LTBI treatment initiations total
Format	Number

Source	Country Reports
Comments	Default value of 0, unless otherwise specified Includes initiations through contact tracing below. Any LTBI treatment initiations total that were not administered through contact tracing are assumed to be drawn proportionally from early and late stage latent TB (diagnosable).

Indicator	LTBI treatment initiations through contact tracing
Format	Number
Source	Country Reports
Comments	Default value of 0, unless otherwise specified LTBI treatment initiations through contact tracing are assumed to be drawn proportionally from early stage latent TB (both diagnosable and diagnosis restricted), to reflect WHO guidelines to treat all children 0-4 contacts of active TB cases.

Indicator	LTBI treatment average duration
Format	Days
Source	Country Reports
Comments	Default value of 180, unless otherwise specified It is assumed that all LTBI treatment (on diagnosis restricted, early, and late latent TB) has identical outcomes.

Indicator	LTBI treatment proportion of lost to follow up
Format	Proportion
Source	Country Reports
Comments	Default value of 0, unless otherwise specified

Indicator	LTBI treatment proportion of successful completions
Format	Proportion
Source	Country Reports
Comments	Default value of 1, unless otherwise specified

3.6. Initialization estimates

Initialization (start point) estimates are required to initiate the model. These values are input on this sheet. At least one value is mandatory for both the latent and active prevalence, but any other estimates might improve the initialization. These values are normally adjusted during calibration to ensure the model starting point is stable in subsequent years.

3.6.1. Indicators

Indicator	Initialization: population size
Format	Number
Source	Country estimates
Comments	This can typically be copied directly from the demographics tab, it is just included a second time to allow different values to potentially be entered as initialization estimates and calibrated without affecting the original data.

TB Prevalence

Indicator	Proportion of the population with active TB
Format	Proportion
Source	Country estimates
Comments	To disaggregate for each population group: use the proportions calculated from the notifications for each population group.

Indicator	Proportion of the population with latent TB

Format	Proportion
Source	Country Estimates Houben and Dodd (2016)
Comments	Few countries have completed prevalence surveys for latent TB. If this information has not been provided, use the total estimated number of latent TB cases from Houben and Dodd (2016), which provides national total estimates for 2014 as a mean and confidence bounds. For subnational analyses, take the proportional burden either: ○ using notifications (ie, number of subnational notifications, over number of national notifications, for each year of interest) ○ If national notification rates are not available, then assume population homogeneity: that the proportion of population in this region have as a likely chance of having latent TB as at a national level. To disaggregate for each population group: use the proportions calculated from the notifications for each population group.

Indicator	Proportion of latent TB cases that are early latent
Format	Proportion
Source	
Comments	Default value of 0.1, unless otherwise specified. This may vary by population e.g. children are most likely to have been infected recently and be early, while older persons are more likely to be latent, and the ratio might be different in a rising or falling epidemic.

Indicator	Proportion of latent TB cases that are on treatment
Format	Proportion
Source	
Comments	Default value of 0, unless otherwise specified.

Indicator	Proportion of active TB cases that are diagnosed
Format	Proportion

Source	
Comments	Default value of 0, unless otherwise specified.

Indicator	Proportion of diagnosed TB cases that are on treatment
Format	Proportion
Source	
Comments	Default value of 0, unless otherwise specified.

Indicator	Proportion of the population that are vaccinated
Format	Proportion
Source	UNDP/World Health Organization (WHO) reports: for general vaccination rate. Country or context-specific reports: for information as to when BCG vaccination was introduced.
Comments	Default value of 0, unless otherwise specified.

Indicator	Proportion of the population that have previously been infected with TB
Format	Proportion
Source	
Comments	Default value of 0, unless otherwise specified.

Indicator	Estimate of average time until diagnosis for new TB cases (days)
Format	Proportion
Source	
Comments	Default value of 0, unless otherwise specified.

3.7. New infections proportions

The proportions of smear/strain status for new cases are used as input data. Each time there is a new case of active TB, this new active TB case is assigned to a smear/strain compartment.

Typically, it makes sense to derive these proportions directly from notification data, but they are included as independent values are they may need to be adjusted if the smear or strain proportions are believed to be changing rapidly, such as with the emergence of more drug-resistant strains where diagnostic tools have not kept up with that emergence and DR strains are under-diagnosed compared to DS strains.

Further, notification data may well be very uneven based on e.g. specific interventions in specific years or just how many people independently seek treatment, whereas the underlying trends influencing the proportions of new infections in each smear/strain combination are likely more stable, so we should represent this in a separate set of input data.

Start from the notification data - if there were 100 SP notifications and 50 SN notifications in the most recent year, then the proportions would be 0.67 for SP proportion of new active infections, and 0.33 for SN proportion of new active infections (they should sum to 1).
Similarly, the ratio of DS:MDR:XDR with each of SP and SN can be calculated from the notification data.

Take note of any trends, review if there is believed to be under-diagnosis of any smear or strain combinations relative to others, and smooth this data out to ensure the data is consistent and reasonable. Make a note of the reasons for any major adjustments in the comments.

3.7.1. Indicators

Proportions of new infections

Indicator	SP proportion of new active infections SN proportion of new active infections DS proportion of new SP infections MDR proportion of new SP infections XDR proportion of new SP infections DS proportion of new SN infections MDR proportion of new SN infections XDR proportion of new SN infections
Format	Proportion
Source	Notification data
Comments	Proportional breakdowns by smear status and drug resistance input here. Proportions are calculated using notification data, In contexts where there is incomplete notification data, estimates here can be

	used to disaggregate notification data Some smoothing may be necessary to ensure data is consistent and reasonable. Calculating proportions: only use years for which all notifications are available and disaggregated. Note that total values can be copied into proportions as whole numbers ie, they can be calculated as the summed total for the corresponding types of notifications. ● Complete the Proportions for Smear status, per population, per year: based on the total number of smear status cases across all drug resistances; ● Calculate the Proportions for Drug-resistant strain per smear status, per population.

3.8. Optional Data

Optional, are not model inputs, rather are useful for calibration. Data inputs are used to compare known values against model outputs. Values here are point estimates e.g. the number on treatment as of January 1 each year, rather than the total over the entire year.

3.9. Comorbidities

Comorbidities that affect tuberculosis disease progression have been calculated for:

- people living with HIV (PLHIV), off antiretroviral therapy (ART),
- PLHIV, on ART,
- diabetics, on insulin
- diabetics, off insulin,
- people who consume high volumes of alcohol, referred to as high alcohol consumption.

Comorbidity values (rates, coverage values, and probabilities) are shown in the table below. For other comorbidities not listed above, these values must be derived.

3.9.1 Table of comorbidity values

	No comorbidity	**HIV**	**Diabetes**	**High alcohol consumption**
General demographics				

TB-related death rate, off CMT, TB unrelated	Informed by representatives from the setting			
Death rate, on CMT, TB unrelated	0	Informed by representatives from the setting		
CMT coverage	0	Informed by representatives from the setting		
Latent Progression Rates				
Early latency departure rate (off CMT)	0.2	0.99	0.4	0.4
Early latency departure rate (on CMT)	0.2	0.2	0.2	N/A
Late latency departure rate (off CMT)	0.0001 to 0.005 (0.00011 among adults, 0.003 is the model default). These values will be adjusted during calibration.	0.0037	0.0037	0.0037
Late latency departure rate (on CMT)	As above	As for default pop	As for default pop	Not applicable
Probability of early-active versus early-late (off CMT)	0.177-0.531 (0.177 among adults)	0.93	0.93	0.93
Probability of early-active versus early-late (On CMT)	As above	As for default pop	As for default pop	Not applicable
Active TB progression rates				
Natural recovery rates	0.03	0	0	0
Escalation rates	0	0	0	higher
Relative infectiousness				
TB infectiousness	1	0.66, could be from 0.2	As for default pop	As for default pop

		to 0.95 Peters, Julian S et al (2018)		

3.10. Reviewing data entered in the databook

Upon completion of the databook you may upload it and begin calibration of demographic and epidemiological data to model outputs. Typical reasons for errors when uploading the databook may include:

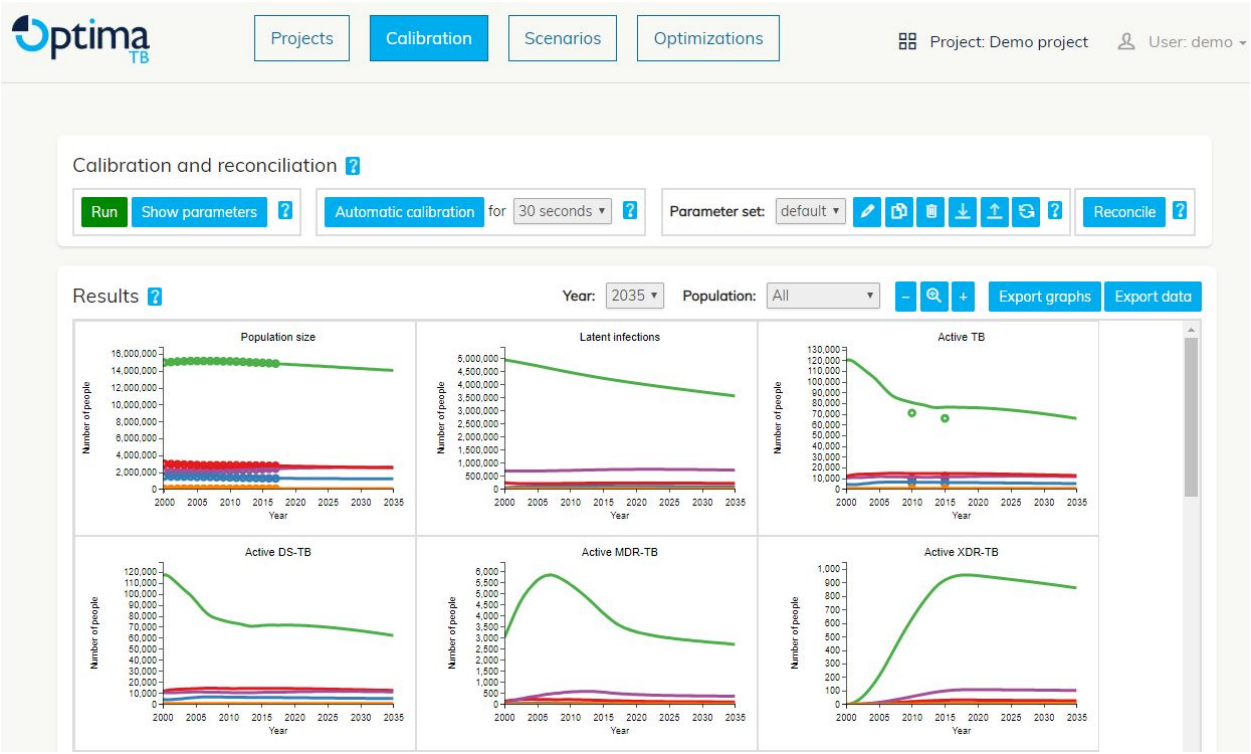
- Missing data for in required fields
- Data entered outside of data entry fields (you must include #ignore)

During calibration, issues and inconsistencies may be noted and can be updated in the data book. Rather than redownloading the databook and editing, it is advised that you maintain a 'master' spreadsheet that can be edited as necessary and re uploaded to the databook.

4. Calibration

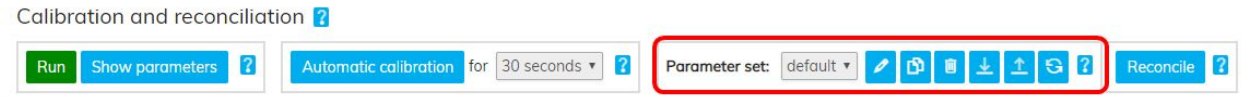
Calibration is the process of adjusting the parameters of the model to get the best possible match to all available data: behavioral, epidemiological, and biological. Initially, the model exactly matches the available behavioral and biological data. However, this may result in a poor fit to epidemiological data. In the process of calibration, behavioral and biological parameters are varied by the minimal amount required to achieve a good fit to epidemiological data.

During calibration, you will be adjusting parameters to best fit prevalence, diagnosis rates, incidence and deaths to historical data.



4.1. Parameter sets

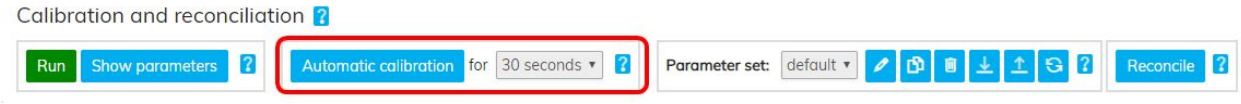
Multiple versions of calibrations can be saved as ‘parameter sets’. This functionality is particularly useful when investigating the impact of different assumptions.



- To create a new parameter set, click the ‘New’ button to the right of the parameter set selection button. This will launch an entry field for naming this new parameter set.
- To rename , copy or delete a parameter set, use the buttons to the right of the parameter set selection button.
- You can refresh the parameter set by pressing the refresh button . This is necessary if an updated databook has been uploaded, to ensure new values appear on calibration graphs.
- Parameter sets can be downloaded by pressing the download button , and uploaded by pressing the upload button .

4.2. Automatic calibration

The first step in calibrating the model is to run automatic calibration. This takes into consideration the different kinds of data available and (where available) their associated uncertainties, and calculates the optimal set of parameters given these uncertainties.



To begin automatic calibration, select a time limit, and click on ‘Calibrate’. The time required will depend on the complexity of the model being calibrated and the quality of the data entered into the model. In general, automatic calibration should be run for as long as possible. If automatic calibration is rerun, it will start from the last point found.

4.3. Manual calibration

Only if an automatic calibration has failed to produce a satisfactory calibration fit should a manual calibration be undertaken.

The parameters that are available for calibration are not parameters that are used directly by the model, but are instead ‘meta-parameters’. These are defined as parameters which affect other parameters.

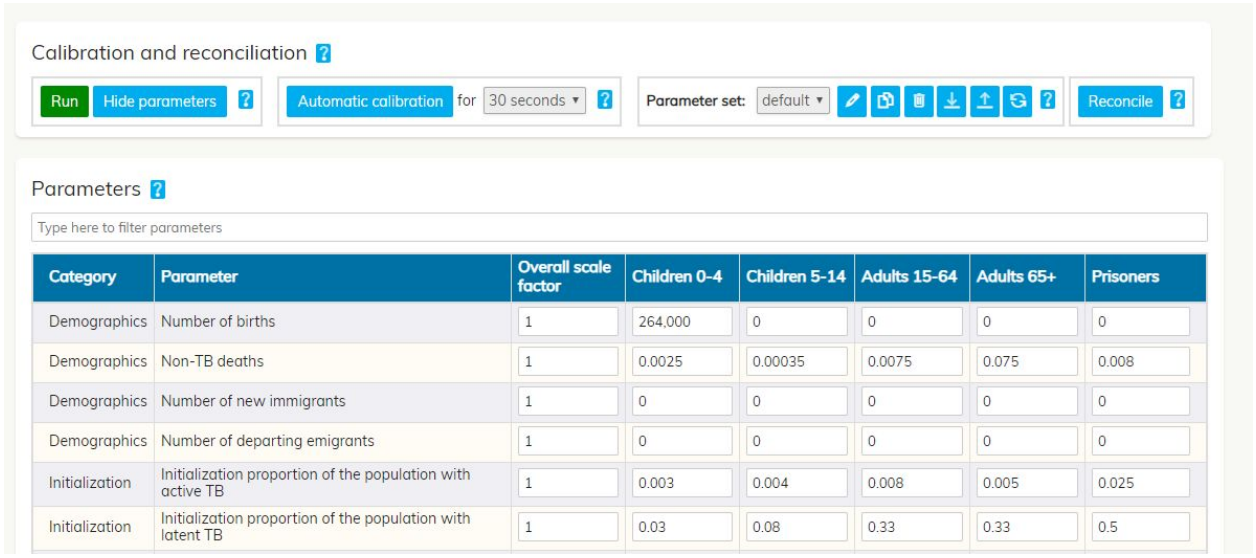
To manually modify a calibration:

1. Modify parameter values within the calibration panel. To view the calibration panel, select the ‘Show parameters’ button.
2. Click the ‘Save and run’ button.

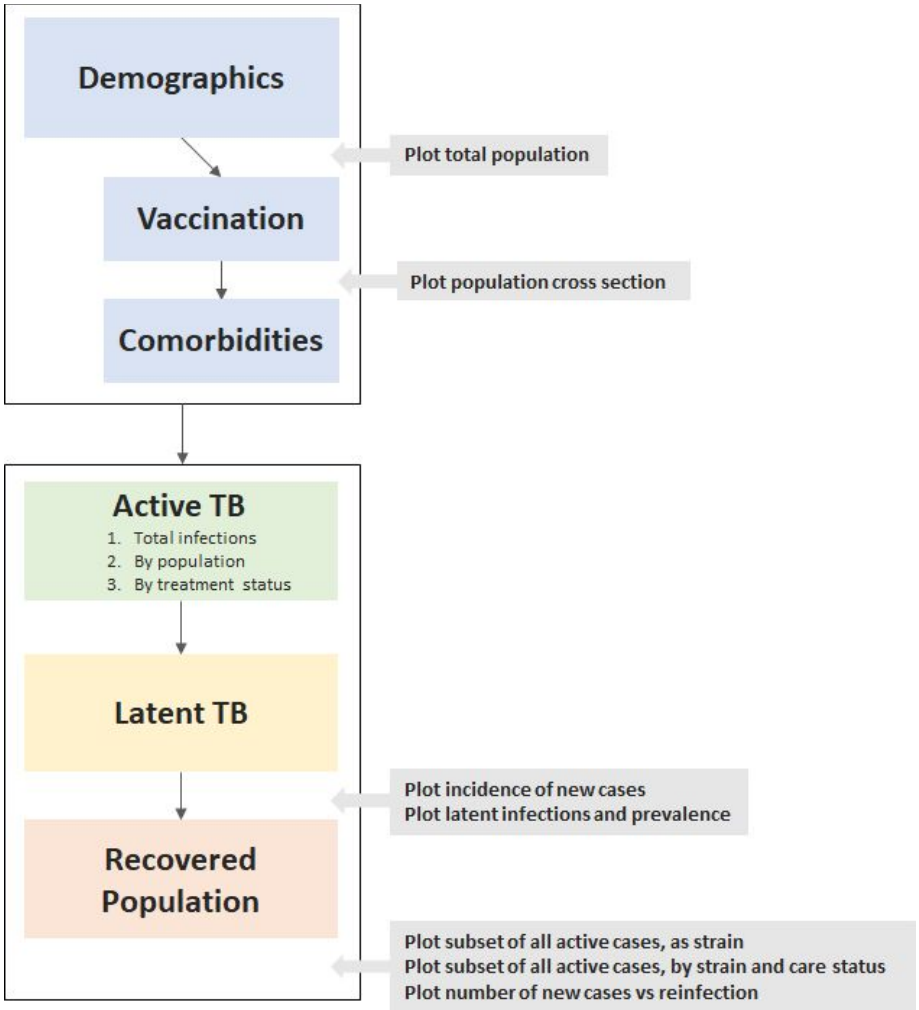
4.3.1 Adjusting parameters in the calibration panel

- Parameters can be adjusted
- Overall scale factor
- Caution for proportion
- Changing parameters does not alter values in the data book
- Creating parameter sets

Guidelines for the calibration process are outlined below.



Category	Parameter	Overall scale factor	Children 0-4	Children 5-14	Adults 15-64	Adults 65+	Prisoners
Demographics	Number of births	1	264,000	0	0	0	0
Demographics	Non-TB deaths	1	0.0025	0.00035	0.0075	0.075	0.008
Demographics	Number of new immigrants	1	0	0	0	0	0
Demographics	Number of departing emigrants	1	0	0	0	0	0
Initialization	Initialization proportion of the population with active TB	1	0.003	0.004	0.008	0.005	0.025
Initialization	Initialization proportion of the population with latent TB	1	0.03	0.08	0.33	0.33	0.5



The process for matching data points is outlined below. Each calibration is first based on demographic data. Ensure demographics data are correct by reviewing the population size

plots. Then calibrate for active TB, ensuring matching for latent TB and for the number who have recovered from TB infection.

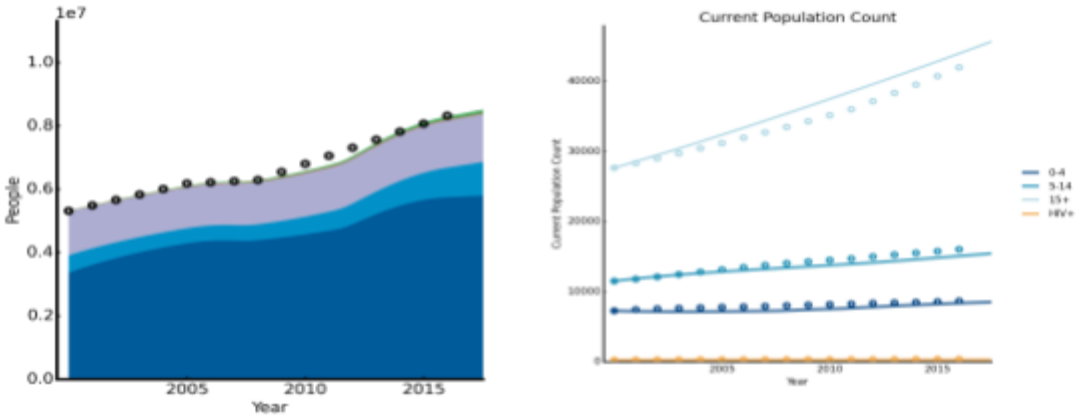
Primary parameters for calibration

1. Parameters affecting new latent TB infections
a) Infection vulnerability factor (relative population susceptibility)
b) Population contacts
2. Parameters affecting progression to active TB
a) Early latency activation rate
b) Late latency activation rate
c) Relapse/reinfection rates

Secondary parameters for calibration

3. Relative infectiousness
4. TB-related death rates
5. Natural TB recovery rates
6. TB escalation rates

Scatter points indicate observed data points matched to model outputs.



4.3.1. Verification steps in calibration

Indicators to be verified
1. Demographics

At all stages, if demographics data start to look wrong (eg, if newly corrected TB deaths have a significant enough impact such that the population sizes no longer fit, for instance), demographics data should be adjusted.

2. Stability

For each of incidence, prevalence, deaths below, as well as general demographics, if the epidemic is very unstable in early years, strongly consider adjusting the initial compartment sizes (for all compartments including diagnosed, undiagnosed, recovered, vaccinated, and each strain and smear) until it looks stable like a historical continuation rather than a jump in the initial years.

3. Incidence

If there's a mismatch that is too low, either the latent estimate is too low and/or the latency progression rates are too low, and/or relapse rate is too low.
If there's a mismatch that is too high, either the latent estimate is too high and/or the latency progression rates are too high, and/or relapse rate is too high.
Verify whether the latent prevalence estimate has wriggle room (preferable), and then consider changing the progression rates or relapse rates.

4. TB prevalence

If there's a mismatch, and active prevalence is too low, but incidence is approximately correct, then either the treatment initiation or treatment success rate is too high, the natural or on treatment death rate is too high, or the natural recovery rate is too high.
If active prevalence is too high but active incidence is about right, then the inverse is true. First verify notifications and treatment first. Treatment outcome rates may easily become too high or too low across all outcomes meaning that there will be too few or too many people accumulating in the 'on treatment' compartments, so check that the number in these compartments matches the expected duration of treatment approximately.

5. Notifications

If there's a mismatch, notification data must match for there to be confidence in model output, as notification data are typically the most reliable data. If the model output for number of diagnoses notifications is lower than the number of notifications, then likely the diagnosis rate is too low.

If the number of diagnoses is higher than the number of notifications, then likely the diagnosis rate is too high but it may be appropriate to check the data as some settings have seen under-reporting of clinical diagnoses.

If the mismatch is specifically in some strains/smears and not others, then consider if the diagnosis rates for each strain/smear combination should be differentiated, or if it would make more sense to adjust the probabilities of new active infections being assigned to those strain/smear compartments. Verify if the number of notifications is too low but the diagnosis rate is already high, then check incidence again – it is likely that incidence is too low and the latent infectiousness or latent-active progression rates need to be increased.

6. TB-related deaths per year

If there's a mismatch for TB-related deaths on treatment, verify that the number of notifications are not too low but the diagnosis rate is already high. Then verify incidence once again. Perhaps

incidence is too low and latent infectiousness or latent-active progression rates need to be increased.
7. Simulated against reported values for care cascade The model cascade plots display the number of people moving between each compartment every year, eg, the number who are first diagnosed, who initiate treatment (including re-treatment), and who have successful treatment outcomes. This can be compared to additional country data if it exists.

4.4. Additional indicators

- Infectiousness of drug susceptible TB
- Weighting of population contacts
- Reevaluation of feedback loop:
    - latent TB, active TB, recovered,
    - probability of early-active vs early-late,
    - late latency departure rate, and
    - initial population sizes for latent TB, active TB and recovered.

Population contacts ● Contact matrix entries - weighting	This represents the likelihood of contact between each pair of populations. This is weighted according to the population sizes by default. Some populations eg, a prisoner population are likely to have contact only with prisoners and not with the general population, but most general populations will have a fairly flat contact matrix. This can be adjusted along with relative infectiousness of DS as part of calibration but is typically avoided other than setting isolated populations.
Infection susceptibility ● Infection vulnerability factors	This is a key calibration parameter equivalent to the 'force of infection' for the Optima HIV model. It reflects the relative likelihood of an individual in each population becoming infected with latent TB (not the risk of infecting others, and not the probability of progression to active TB). The main factor to adjust is 'SP DS Infectiousness' ● The default value is 1, but this value can be any positive number as it

	reflects the relative likelihood, not an actual probability. ● Typical modifications from applications: ○ Young children: 5-60 ○ Older children are relatively low risk if a separate population: lower <1 ○ Prisoners: higher (5-100) ○ Patients with comorbidities (1-10)
Latency to active progression rates ● Early latent departure rate ● Late latent departure rate ● Probability of early-active versus early-late	These are the parameters that determine the likelihood of progression from latent TB to active TB. Typically should be 0.2 as an early latent departure rate (higher for comorbidities eg, 0.99 for PLHIV) Probability of early-active versus early-late should be 0.177 but may be a little higher in higher risk populations (up to 0.531 typically) or up to 0.93 in PLHIV Late latent departure rate defaults to 0.003, but could be anywhere from 0.0001 to 0.005 or even higher in extreme circumstances.
Initial population sizes ● Active infections ● Latent infections ● Recovered cases	As noted above, it is very likely that initial compartment sizes will need calibrating. Typically in the longer term the epidemic will eventually level out to a stable point depending on the parameters, but initial parameters should generally be defined to reach a relatively stable epidemic in early years assuming treatment was constant at the simulation start time.

5. Programs

The program book is where all the information related to interventions along the cascade are defined. This involves collating data on the current allocation of resources, and relationship between program spending, coverage and population-specific outcomes.

5.1. Creating a new program book

As noted in Section 2.4.4., creating a new program book is done on the “Projects” page by clicking on “New” in the column “Program book” of the project table. This will launch a pop-up box where you can specify how many programs will be included in the response. The number you select here is used only as a guide, you can add or remove programs from the program book without redownloading the program book.

When deciding which interventions to include, the following questions should be considered:

- Does the program plays an important role in the overall epidemic response?
- Is there data on program coverage?
- Is there past expenditure data? (not relevant if new program, such as new treatment regimen)
- Is there evidence to indicate the effect that the intervention has on rates of flow between model compartments?

An important consideration is to keep the number of programs manageable, for the resulting analysis to be robust. You can also add prospective or planned programs to this to be included in the analysis.

Clicking “Create” will download an Excel spreadsheet, where you will be able to enter all the information about the interventions. The program book has three sheets, which are explained in detail below.

5.2. Program targeting

The first sheet is called “Program targeting” and is where you should enter information about the names of the interventions and who they are targeted at.

Specific information about what to enter for each column in the “Program targeting” sheet is

provided in the table below.

Column(s)	What to fill in
Abbreviation	This defines the label that will be used to refer to the parameter throughout the code. It should be short and not contain any spaces or special characters. This will not appear in any model outputs.
Display name	This defines the label that will be used to refer to the parameter in any outputs. It should be descriptive (although not too long) and may contain spaces or special characters.
Targeted to (populations)	You will see a separate column for each population in your project. If a particular program is targeted to a population, you should enter 1 in the cell corresponding to that intervention and that population.
Targeted to (compartments)	You will see a separate column for each compartment in the framework that underlies your project. If a particular program is targeted to a compartment, you should enter 1 in the cell corresponding to that intervention and that population. For example, a testing program is typically targeted to the "Undiagnosed" and "Susceptible" compartment. Note that there are many instances of cascade programs that have so-called downstream effects. For example, different testing modalities may lead to different rates of treatment adherence further downstream. This information should NOT be entered here: rather, you should only consider who the target population is for the intervention. Later on you will have the opportunity to enter information about downstream effects.

5.3. Spending data

The next sheet is called "Spending data" and is where you should enter information about expenditure and coverage of interventions

For each program, you can enter information under 4 categories: total spend, capacity constraints, unit cost, and coverage.

5.3.1. Total spend

Here you should enter data on the total amount that has been spent on the intervention in any year that you have data. If you do NOT have data on this, you can enter your best estimation the "Assumption" column. All cost and spending data must be in nominal terms and in a consistent currency. To include a program that doesn't presently exist, simply enter zero spend and zero coverage.

Past expenditure data do not always give a good indication of what future investments will look like. For example the package of services may change; or past cost may only reflect an inception phase and actual unit cost could be substantially lower. You can enter information about future unit costs in the “unit cost” rows. In this section, you should only enter past spending data. This is required so that the model can learn about the relationships between past investments and past outcomes.

The historical program spending data should include all investments that would not exist, if the program was not being implemented. Therefore program specific management cost (eg. administrative costs) should be included in the program.

5.3.2. Unit cost

You will need to enter values for the unit cost of the program. You can enter a value for a single year if costs are not changing over time, and this implies a constant unit cost into the future. There are many instances where declining or increasing costs may be anticipated for certain programs due to changes in procurement and arrangements, supply shocks etc. Enter unit costs for multiple years to incorporate time-varying cost curves.

> **What if there are different program models with different unit costs?**
> If there are two substantially different cost estimates for a program, because different packages are being provided, you could split the program into two programs with different names (eg. provider-initiated testing and outreach testing).

5.3.3. Capacity

Here you have the option to enter information about any capacity constraints that the intervention may have. For example, if you are considering a vaccination program and there are only a fixed number of vaccinations available, you could indicate this here. This will impose a hard upper limit on the potential scale-up of this program.

It is NOT required to enter anything here; you are free to leave it blank if capacity constraints do not apply.

5.3.4. Saturation

The largest possible proportion of the target compartment(s) that could be covered by a program in a given year

5.3.3 Coverage

Coverage values must be in terms of people covered not goods or services provided (e.g., numbers of people covered by a testing program, not the number of tests conducted).

5.4. Program effects

Having defined the relationship between program spending and coverage, we now need to specify what happens when someone is covered by the program. The next sheet is called “Program effects” and is where you should enter information about how the programs affect the outcomes.

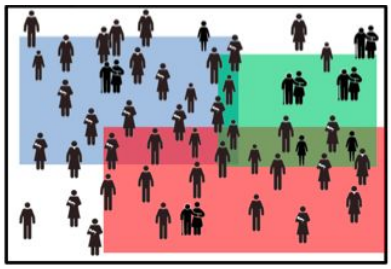
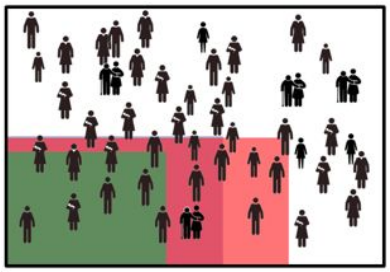
Unlike the previous two sheets, which were organized with a separate row or block of rows for each program, in this sheet there is a separate block of rows for each of the parameters that were marked can be targeted by a program in the Optima TB framework. For each parameter, you need be to enter information about the programs that target it. This will include information about the effects that the programs have in isolation as well as the way that they act in combination.

5.4.1. Baseline value

In this column, you are asked to enter a value specifying what the outcome would be in the absence of any programs targeting this parameter. These values reflect background behaviour in the absence of the program being considered and should incorporate goods and services accessed through alternative programs or private means (e.g., privately funded treatment). In some cases, it will be appropriate to enter 0 here (e.g., if there is no access to screening, testing, or treatment outside of the interventions that are being considered here), and in other cases another value would be more appropriate.

5.4.2. Coverage interaction

Optima TB provides 3 options for calculating coverage when there is more than one modality in place targeting the same outcome. Illustrations of these are provided in the table below:

RANDOM	NESTED	ADDITIVE
		

The three options work as follows:

- The *random* option assumes that any given person has a finite probability of being covered by a given program, which is independent of whether or not they are covered by any other program. For example, if community-based advertisements have coverage of 35% (shown in red in the diagram), workplace-based retention programs have a coverage of 25% (shown in blue in the diagram), and clinic-based retention programs have coverage of 15% (shown in green in the diagram), then the random option would calculate that:
    - 1.3% (=35%×25%×15%) of people will be covered by all three programs;
    - 7.4% (=35%×25%×(100%-15%)) of people will be covered by the community-based and workplace programs;
    - 3.9% (=35%×(100%-25%)×15%) of people will be covered by the community-based and clinic-based programs;
    - 2.4% (=(100%-65%)×25%×15%) of people will be covered by the workplace-based and clinic-based programs;
    - 22.3% (=35%×(100%-25%)×(100%-15%)) of people will be covered by the community program only;
    - 13.8% (=(100%-35%)×25%×(100%-15%)) of people will be covered by the workplace program only;
    - 7.3% (=(100%-35%)×(100%-25%)×15%) of people will be covered by the clinic-based program only;
    - 41.4% (=(100%-35%)×(100%-25%)×(100%-15%)) will be covered by no program.
- The *nested* assumption is that a person is only covered by a lower-coverage program if they are also covered by a higher-coverage program; for example, it could be assumed that all people are in the workplace are also exposed to community advertisements and to the clinic-based programs. In this case, 15% of people would be covered by all three

programs, 10% by both the workplace program and the community advertisements, 10% by the community advertisements only, and the remaining 65% by no program.

- Finally, the *additive* assumption is that the overlap between target populations is minimized. Let's now assume that the 3 modalities represent a school-based retention program (with coverage of 35%), a workplace-based retention program (with coverage of 25%), and a retention program aimed at homeless unemployed adults (with coverage of 15%). This would imply that 85% of people are covered by some program while the remaining 15% are not covered by any program.

Although the examples above illustrate only two modalities, they can be extended to any number of interacting programs.

5.4.3. Impact interaction

The cascade analysis tool provides 2 options for calculating the outcome for someone who is covered by more than one program.

- The *best* option assumes that the impact is equal to that of the most impactful program. For example, if there were two adherence programs, one of which lowered loss-to-follow-up by 20% and the other of which lowered it by 30%, then someone who was covered by both would have a 30% reduction in their probability of being lost to follow-up.
- The *synergistic* option allows for the possibility that modalities strengthen each other. In this case, if there were two adherence programs, one of which lowered loss-to-follow-up by 20% and the other of which lowered it by 30%, the synergistic option would assume that someone covered by both adherence interventions would experience a reduction in the probability of being lost to follow-up of 20%*(100%-30%)+30%*(100%-20%) = 38%.

5.4.4. Program Effects

In the columns that are labelled with program names, you should enter values indicating what the outcome would be for someone covered by each program in isolation.

5.5. Reconciliation

TBC

6. Scenarios

Having calibrated the model, defined your baseline epidemic profile and programmatic response, you are now ready to run analyses. There are two types of analyses within the cascade analysis tool: scenarios and optimizations. In this section, we discuss scenarios.

Scenarios are run when model parameters such as testing rates and treatment are adjusted by varying program budgets. Scenario results estimate the impact of a change in spending for one or more programs. Only parameters targeted by the program are affected.

6.1. Managing scenarios

6.1.1. Adding a scenario

To add a budget scenario, click “Add scenario”. This will launch a dialog box where you can define features of the scenario.

To start, give the scenario a name, select the parameter set that you want to use as a baseline and the year that the budget changes are going to take effect. Underneath this, you will see a table that lists the interventions that were defined within the program book, along with the latest reported spending for each one. Here you can modify any of the values to define a new budget.

6.1.2. Editing/copying/deleting a scenario

Buttons to manage these options are available in each row of the scenario table.

6.1.2. Running scenarios

After you have defined your scenarios, the next step is to run them. It is possible to define multiple scenarios and run them all at once. If you don’t want to run a particular scenario (but you also don’t want to delete it in case you want to run it later), you can deselect the “Active” checkbox in the scenarios table. After you have defined your scenarios, the next step is to run them.

Click the green “Run scenarios” button to generate results.

6.2. Viewing results of scenarios

After running scenarios, the results are displayed below. As on the 'Baseline' page, you will have options to view the cascade for different years/populations and to export graphs and data.

7. Optimization

Optimizations deploy a mathematical optimization algorithm over the defined programmatic response (cost curves and outcome functions) to determine the most effective allocation of resources to achieve user-specified objectives.
There are two basic types of optimization:

- Money optimization allows users to determine the minimum amount of money required to achieve a specified change in incidence and deaths.
- Outcome optimization allows users to determine the maximum health impact for a fixed, specified budget.

7.1. Creating an optimization

To create an optimization, click 'Add optimization'. This will launch a dialog box where you can define features of the optimization.

To start, give the optimization a name and select the parameter set that you want to use as a baseline. Underneath this, you will see other options that are described in more detail below.

7.1.1. Setting a timeline

One of the most critical aspects of performing an optimization is to set appropriate time horizons for the analysis. Typically, the start year would be the current year. The end year would be determined by the policy question that you are trying to answer.

For example, if the period 2015-2020 is chosen, the budget for 5 years (i.e., the budgets for years 2015-2016, 2016-2017, 2017-2018, 2018-2019, and 2019-2020) will be optimized.

> A timeline example: How can existing funds best be allocated in order to get as close to possible to the target of a 35% reduction in TB related deaths by 2020? For this policy question, you would set the end date to 2020.

7.1.2. Setting the budget

The budget is pre-filled with the sum of the most recent spending on each program. However, if you want to investigate a different amount, you can modify the budget total here. The baseline assumption is that the amount that you enter here will be annually available in each year over the period defined in your optimization timeline (see section 5.1.1.), and will be allocated according to the allocation between programs given in the baseline.

7.1.3. Setting objectives

This specifies how the different objectives will be weighted relative to one another. Typical "outcomes optimizations" are set up in order to figure out the optimal allocation of a fixed budget, in order to minimize:
(a) new infections,
(b) deaths
(c) a weighted combination of the two.

7.1.4. Setting constraints

Constraints define the maximum changes in funding over the program period. Often, optimization objectives do not incorporate the full set of political, social, welfare, and justice considerations that are necessary for producing a properly balanced response. For this reason, in addition to providing options for different kinds of optimization objective, there are also options for setting constraints on particular programs. Each program can be limited to increases or decreases in funding by a certain amount, either per year or over the entire program period. If "100%" is entered in any box, then the allocation for that program will not change in that direction. If both "Minimum" and "Maximum" are set to 100%, then that program will be treated as a fixed cost and will not be optimized. Note that constraints are relative, not absolute.

7.2. Viewing results of optimizations

After you have defined your optimizations, the next step is to run them. Click the green "Run" button next to each optimization in the table to generate results. The length of time required to perform an optimization depends on the complexity of the model and the number of programs being optimized. After running an optimization, the results are displayed.

9. Supplementary materials

9.1 Disease progression using the epidemiological model

Disease progression rates have already been calculated, and should not need to be determined again. However, if there is a new comorbidity, a different strain of tuberculosis or other factors that result in different disease progression rates within a country, then this step should be performed in collaboration with the Optima Consortium for Decision Science modelling team.

Progression rates for tuberculosis were determined for natural progression of the disease, and then in response to treatment for values such as disease duration and fatality (summarized in Table 1). The model was set up to replicate these experiments by considering a population that all started in the same disease state, and examining how quickly people progressed into subsequent disease states or death.

Using this process, it was generally found that there was good agreement with the values for case fatality and disease duration for HIV- population, for both the natural progression of the disease and when on treatment (Table 1; Figure 2).

Table 9.1: Disease progression values

Natural progression		Model output	Reference
Untreated, HIV-	**Disease duration:** 1-4 years 3 years **Smear positive case fatality:** Case fatality rate: 70% 5 year case fatality: 55% 10 year case fatality: 72% **Smear negative case fatality:** Case fatality rate: 20% 10 year[2] case fatality: 20%	3.25-3.5 years 46% at 5 years 66% at 10 years 15% at 10 years	 WHO Tiemersma 2011 WHO Tiemersma 2011 Tiemersma 2011 WHO Tiemersma 2011
Untreated, PLHIV	Disease duration: 0.01-0.2 years SP Case fatality rate: 83% SN Case fatality rate: 74%		WHO WHO WHO
Disease progression on treatment		Model output	Reference
Treated, HIV negative	Disease duration: 0.2-2 years	0.25-2 years	WHO
Treated, PLHIV	Disease duration: 0.01-1 year		WHO

[2] Reported as 'lifetime case fatality', although in the same study the authors refer to a maximum window of 10 years.

10. References

Andrews et al (2012) Risk of Progression to Active Tuberculosis Following Reinfection With Mycobacterium tuberculosis. Clin Infect Dis.

Glaziou, P., Sismanidis, C., Zignol, M. and Floyd, K. (2016). Methods used by WHO to estimate the global burden of TB disease. Global TB Programme. [online] Geneva, Switzerland: World Health Organisation. Available at: http://www.who.int/tb/publications/global_report/gtbr2016_online_technical_appendix_global_disease_burden_estimation.pdf [Accessed 26 Jan. 2017].

Peters, Julian S et al (2018) Advances in the understanding of Mycobacterium tuberculosis transmission in HIV-endemic settings, The Lancet Infectious Diseases (Published Online December 13, 2018 http://dx.doi.org/10.1016/S1473-3099(18)30477-8)