Optima HIV User Guide Volume II

Software Reference Manual

Optima Consortium for Decision Science

July 2024

Introduction

This document is a step by step guide to setting up, running and evaluating results for core Optima HIV analyses. The skills and familiarity gained will allow users to customize and apply the software to a variety of epidemiological and policy contexts.

Working through practical examples, we address some of the most common questions posed by decision makers and planners when using Optima HIV to inform their HIV response.

The guide is structured into five modules:

1. **Create project and data input**: preliminary steps to generate a project and provide an overview of data inputs.
2. **Calibration**: the process of fitting model parameters to historical data and setting a baseline for scenarios and optimization analyses.
3. **Programs, cost functions and outcomes**: define relationships between program spending, coverage and population-specific outcomes.
4. **Scenarios**: set up and run scenario analyses.
5. **Optimizations**: set up and run optimization analyses.
6. **Geospatial**: set up and run geospatial analyses.

This document is Volume II of the **Optima HIV User Guide**.

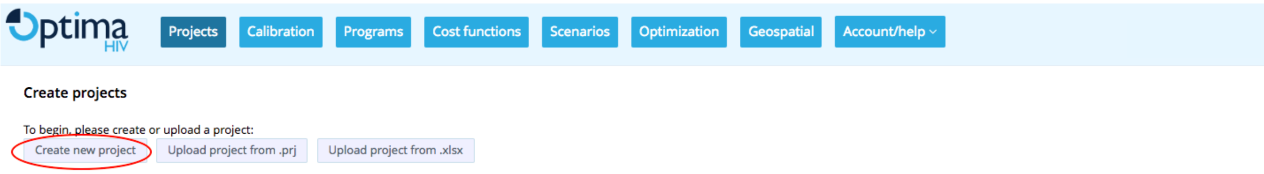
1. Create project and data input

Optima HIV analyses are grouped into **projects**. To start a set of analyses using Optima HIV, you must first create a project.

> Each project is associated with a fixed number of populations. To use different choices for the populations, you will need to create a separate project.

On the opening *Projects* page, there are three options: i) create a project from scratch ii) upload a project file or iii) upload an Excel data input sheet.

1.1. Create new project



1.1.1. 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). If you were creating a simple model to look at counterfactual scenarios in southern Ruritania, you might enter "Southern Ruritania, simplified model".

1.1.2. 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.

When entering a date range, it is advisable to start the model from at least 2000 to allow sufficient time for epidemiological dynamics to develop.

1.1.3. Final year for data entry

The final year of data entry determines what years will be available to enter data into.

For example, if "2000" is entered as an initial year and "2015" for a final year, there will be 16 columns in which to enter data (2000 to 2015 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 2013 data, but are expecting 2014 data soon, and intend to repeat the analysis in 12 months when 2015 data become available. In this case, it would be advisable to enter (at least) "2015" as the final year, so you do not have to create a new spreadsheet to enter 2014 and 2015 data when they become available. In the meantime, these columns can be left blank.

1.1.4. Manage populations

Optima includes 13 predefined populations, although you can add as many more as desired.

The flexibility of Optima allows populations to be broken down into smaller cohorts. A more granular population breakdown may be desirable but is ultimately limited by availability of data. A typical Optima analysis would breakdown the general population into i) males and females 0-14 (before being generally sexually active); ii) males/females 15-49 (at higher risk) and iii) males/females 50+. Also consider the programmatic relevance of population selection. Is the population specifically targeted by a program? A more complex model structure may make the interpretation of results more complex and distract from key findings.

Consider data availability and programmatic relevance of these populations upfront, as any later amendments will require certain steps to be redone.

To use one of the existing predefined populations, simply check the box in the "Active" column corresponding to that population. For any of the existing populations, you may also edit or copy them (if you copy them, it automatically opens the Edit dialog), since you cannot have two populations with the same short name or long name).

In the "Edit" dialog box, you can change the name and properties of a population. An important use of the "Edit" dialog box is for generalized epidemics, where age structure is important. Only basic age groups are included by default in Optima. If a full age structure is required, you can enter these by editing and/or copying the "Other males [enter age]" or Other females [enter age]" populations. Finally, in addition to editing a population, you can add a new population with the "Add a population" button. This has the same interface as the "Edit" button, except no values are preselected.

Note that the "Short name" is what will appear in plot legends and titles. The "Name" is largely for mnemonic purposes and does not appear in Optima's key outputs. Both the short and full names for a population must be unique. There is no strict restriction on how short the "Short name" must be; however, names longer than 10 characters may result in plots that have poor layouts.

1.1.4.1. Deciding which populations to include: general tips

The following criteria can guide this decision:

1. the population plays a substantial role in the country's epidemic;
2. the population can be clearly defined;
3. the population is currently or could be targeted with HIV programs;
4. there is a minimum amount of data or reliable estimates for this population, most critically on population size and HIV prevalence;
5. the population does not have substantial overlaps with other populations, for which there is insufficient data to track the overlap (e.g. TB patients are a group, which cut across different other populations, but the proportion of TB patients coming from different groups may be hard to establish, i.e. in this case TB patients would be hard to include as a separate population).

The user does not have to include all the default populations in the analysis and should consider the following:

- The user can decide on whether to disaggregate certain populations by gender or not depending on data availability (for example if over 95 percent of migrants or PWID are male, there may not be a need for separate population groups for female migrants or females who inject drugs).
- In most epidemics, certain key populations like female sex workers, their clients and men who have sex with men play a critical role and should be included.
- Users are encouraged to include the categories of Other males and Other females in order to ensure that the country's total population is reflected in the analysis. This will allow for checking the total population of the country against other demographic data

and projections. It will also make the Optima HIV analysis more comparable to other national HIV estimates, which are based on the country's total population in specific age groups.

Based on the country's epidemic type the Optima HIV user can add more populations, while considering that including more populations will also make the results of the analysis more complex and will require specific data on these populations.

The total sub-groups of populations should reflect the total population numbers as reflected in the general census reports or population and demographic studies.

1.1.4.2. Deciding which populations to include: concentrated epidemics

In concentrated epidemic settings specific other key populations may play important roles. Such populations could include:

- Female partners of MSM,
- Female partners of sex work clients,
- Sexual partners of people who inject drugs,
- Migrants,
- Uniformed services,
- Specific other professional groups,
- Prisoners and others.

What will determine their inclusion as a separate population is their contribution to the epidemic, and whether in the country context these populations are identifiable for delivery of HIV programs. For example, female partners of sex work clients account for a substantial portion of new data availability infections in many settings, but are difficult to distinguish programmatically from other women aged 15-49 years and there are few data on this group in most settings, their inclusion therefore may not be essential.

1.1.4.3. Deciding which populations to include: generalized epidemics

In generalized and mixed epidemics, it could be relevant to further divide the general population of males and females. This could be done by age, preferably using age definitions, which are relevant to the epidemic and for which other data will be available, for example:

- Males 15-24 years
- Males 25-49 years
- Males 50+ years
- Females 15-24 years
- Females 25-49 years
- Females 50+ years

The default options for including children in Optima HIV is to divide them into infants (0-1 years) and children (2-14 years). However, it is also possible to include one population of children 0-14 years. For age definitions for 50+ age groups, enter an "Age to (years)" of "99" to represent a 'maximum' age. If the country would like to specifically analyze new infections, number of HIV-positive infants and deaths among infants, it will be useful to separate the two populations. Otherwise, one population of children may also be sufficient, as it would be assumed that the majority of HIV infections among children would relate to earlier MTCT.

Optima HIV users could also apply definitions, which they know from other epidemic models, for example by HIV risk: people who are not sexually active ("no risk"), people in a "stable relationship", people with casual partners. However, there may be overlaps between such categories and others and data may not be consistently available for such definitions. Therefore, it will be preferable to use definitions, which are supported by most standard DHS reports like disaggregation by age or marital status.

In principle, Optima HIV is sufficiently flexible to include any number of different populations. For example, in some generalized epidemics specific geographical areas, settings like mines, migrants or even groups like long-term sero-discordant couples could be included. In most countries, this would, however, require secondary analysis of DHS or other data in order to come up with HIV prevalence, behavior and service use data for such user-defined populations.

1.1.4.4. Deciding which populations to include: subnational analysis

Optima HIV can also be used to conduct sub-national HIV epidemic analysis, which can be particularly relevant in countries with large epidemics and countries with relatively heterogeneous epidemics such as geographically mixed epidemics (like Nigeria, Indonesia, Kenya, Tanzania, Mozambique and many others). In such cases, the user does not have to define all populations for each region of the country, but could focus on key populations in specific geographical regions.

Below is an example how this could be addressed in a large country with three epidemic sub-regions including one with drug injecting practices and one with an epidemic among mine workers:

Female sex workers, Region A	Female sex workers, Region B	Female sex workers, Region C
Clients of sex workers, Region A	Clients of sex workers, Region B	Clients of sex workers, Region C
Men who have sex w. men, Region A	Men who have sex with men, Region B	Men who have sex w men, Region C
People who inject drugs, Region A		
Other males (15-49), Region A	Other males (15-49), Region B	Other males (15-49), Region C
Other females (15-49), Region A	Other females (15-49), Region B	Other females (15-49), Region C

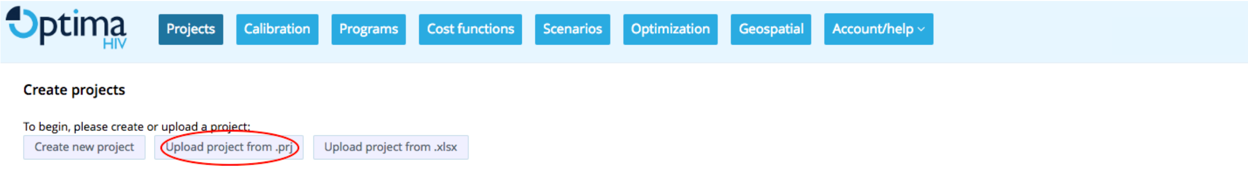
	Mine workers, male, Region B	
Infants (below 1), Region A	Infants (below 1), Region B	Infants (below 1), Region C
Children (1-14), Region A	Children (1-14), Region B	Children (1-14), Region C
Other males 50+, Region A	Other males 50+, Region B	Other males 50+, Region C
Other females 50+, Region A	Other females 50+, Region B	Other females 50+, Region C

1.1.5. Create project and download data entry spreadsheet

Once you have entered a project name (Sec 1.1.1.), a first and final year for data entry (Sec 1.1.2., 1.1.3.) and have selected at least one population for the project, you will be able to click on the "Create project and download data entry spreadsheet" button. This will:

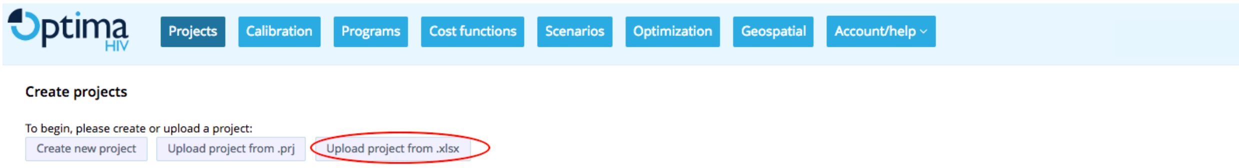
1. Add the current project to the list of available projects associated with the account.
2. Bring up a file save dialog box with the freshly generated blank input spreadsheet. This spreadsheet is created in accordance with the number and names of the populations and programs and the years for data entry. This spreadsheet will be used to input all data used for the Optima HIV analyses.

1.2. Upload project from .prj



This option is useful when sharing work with collaborators. Clicking the button "Upload project from .prj" will bring up a file load dialog box. Optima HIV projects have a file extension ".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.

1.3. Upload project from .xlsx



This option is useful when sharing work with collaborators. Clicking the button "Upload project from .xlsx" will bring up a file load dialog box, where you will have the option to select an Optima HIV data entry spreadsheet. This will create a project with the populations in the spreadsheet.

1.4. Project management options

Once you have projects 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.

1.4.1. Delete 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.

1.4.2. Copy a project

Click the "Copy" button. This will create a copy of your project.

1.4.3. Rename a project

Click the "Rename" button. This will bring up a dialogue box where you can specify a new name for your project.

1.4.4. Upload spreadsheet to a project

Before any Optima analyses can be conducted, the data entry spreadsheet must be filled out correctly and uploaded. Detailed instructions on how to fill out the data sheet are contained in Volume 4 of this user manual. To upload a completed data spreadsheet, click on "Upload" in the column "Data spreadsheet (.xlsx)" 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 the data sheet.

1.4.5. Download spreadsheet from a project

If you wish to view the data entry spreadsheet for a given project, click on "Download" button in the column "Data spreadsheet (.xlsx)" of the project table. This will bring up a file open dialog box, where you can save the file you wish to download.

1.4.6. Download project

If you wish to share a project with a collaborator, the easiest way to do so is to download the project. There are <u>three</u> options for doing this, two of which can be found in the "Project file (.prj)" column of the project table:

1. The "Download" button will download a version of your project <u>without</u> the results. A project file without results is much smaller in size. It 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), but it would not contain the results. Therefore, if you take this option, you would need to rerun the analyses in order to see the results.
2. The "Download with results" button will download a version of your project <u>with</u> the results. A project file without results is much bigger in size, but it means that you would not need to rerun the analyses in order to see the results.

The final option for downloading projects is to select the checkbox in the leftmost 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.

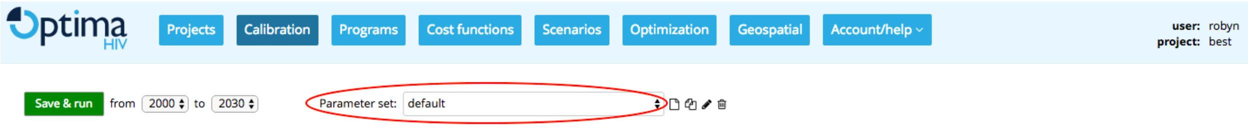
2. Calibration

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 is 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. The relationship between these indicators and their dynamics are determined by the various demographic, behavioral and epidemiological data, and population interactions defined in the input data. Projecting the epidemic into the future provides a baseline for scenarios and optimizations.

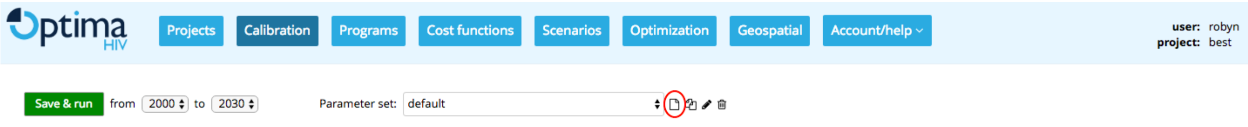
In this section, the elements of the calibration page are introduced.

2.1. Parameter sets



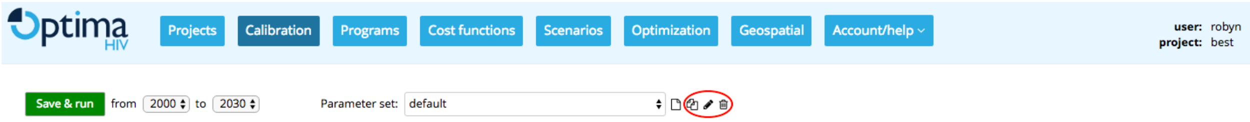
> Multiple versions of calibrations can be saved as 'parameter sets'. This functionality is particularly useful when investigating the impact of different assumptions. Uncertainty around HIV prevalence, background mortality and population sizes can have potentially significant effects on results that are not immediately apparent during calibration. Test the sensitivity of results to these alternatives by running scenarios and optimization analyses across alternative calibrations. For example, where it is known that a particular source is likely to overestimate HIV prevalence in the female general population, calibrations with alternative lower prevalence curves can be saved and used later to evaluate the effect on scenario and optimization results.

2.1.1. Create new parameter set



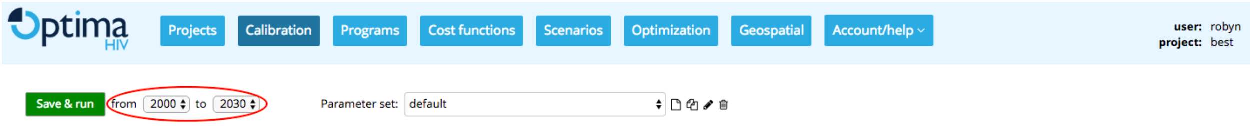
To create a new parameter set, click the "New" button to the right of the parameter set selection button. This will launch a dialogue box where you can decide on the name of your new parameter set.

2.1.2. Copy/rename/delete new parameter set



To rename, copy or delete a parameter set, use the buttons to the right of the parameter set selection button.

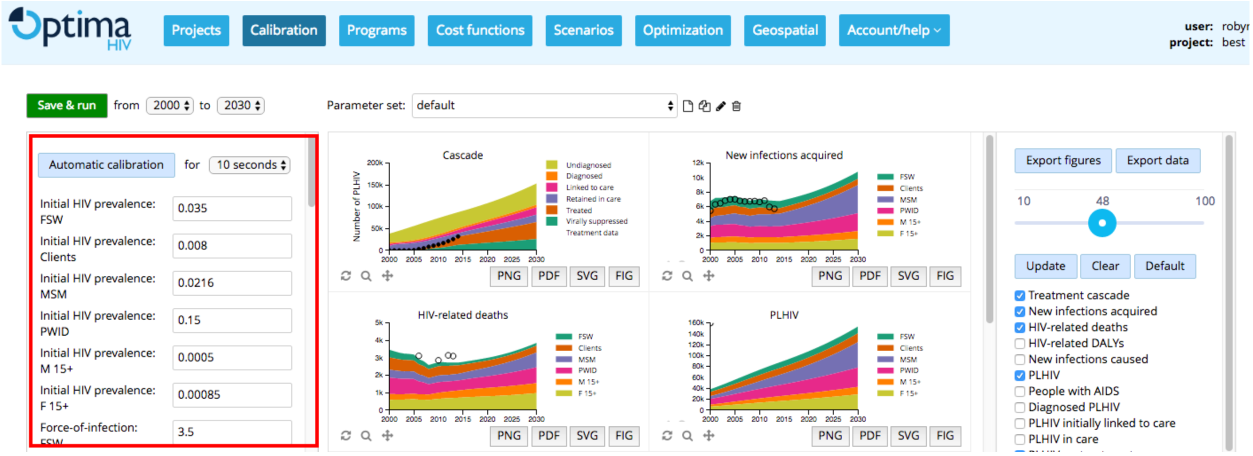
2.2. Years for running model



Using the drop-down year menus, you can select the years that you want to run the model for. The first available year will be the first year for which data was entered for this project (see Sec 1.1.2.).

2.3. Calibration panel

This panel provides the main functionality associated with this page: calibrating the model.



2.3.1. ART projection assumptions

Optima provides two different possible assumptions regarding projecting numbers of people on ART in future. By default, Optima assumes that the proportion of diagnosed PLHIV on ART will remain constant from the last-entered value of the treatment data. Under this assumption, the number of PLHIV on ART will increase if the number of diagnosed PLHIV increases. Alternatively, one can assume that the *number* (not proportion) of PLHIV on ART remains constant. For epidemic projections, the former assumption tends to be more realistic. However, it cannot be used in the context of budget scenarios or optimizations where ART is included as a program, since the assumption of a fixed proportion of PLHIV on ART overrides the proportion that would be calculated from the scenario or optimization. Thus, in many cases it may make sense to create two parameter sets: one assuming a fixed proportion, for use in epidemic projections; and a second assuming a fixed number, for use in budget scenarios and optimizations.

2.3.2. Automatic calibration

The first (and in many cases only) 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 “Automatic calibration”. The time required will depend on the complexity of the model being calibrated and the quality of the data entered into the model. More complex models, or models with poor quality data, will take longer to calibrate. In general, automatic calibration should be run for as long as possible (i.e. 1 hour). Shorter options are available for quick explorations, or for use in conjunction with manual calibration (see below). If automatic calibration is rerun, it will start from the last point found.

2.3.3. Manual calibration

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

The parameters that are available for calibration are not parameters that are used directly in the model, but are instead “meta-parameters” (i.e., parameters that affect other parameters).

The majority of Optima models can be calibrated by adjusting only

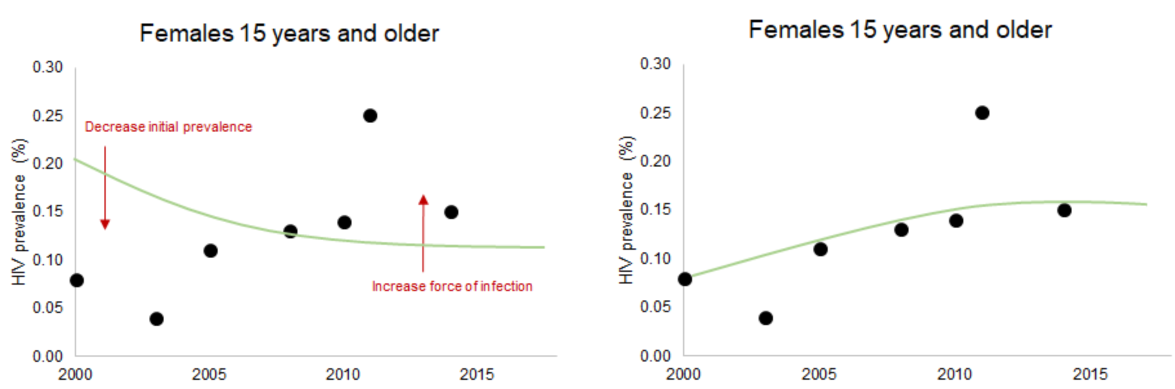
1. Force of infection (FOI) metaparameter (scaling the FOI derived from data inputs) and
2. Initial prevalence for each population.

Initial population sizes are read in from the data but can be adjusted directly during calibration. Behavioral, biological and other constants are informed by the latest evidence and are generally left unadjusted. Users are free to override as demanded by context-specific evidence. Calibration is a context specific process with no strict rules for manual calibration, the following steps are broadly applicable.

To manually adjust a calibration:

1. Start by trying to match calibration curves to the most reliable data or estimates. Depending on the nature of the epidemic (e.g. concentrated, generalized, or mixed), adjusting data that shows the worst fit to the curve or adjusting the population with the largest number of PLHIV generally provides a good starting place as these have the greatest impact on other populations
2. Modify parameter values from the calibration panel. The most commonly adjusted parameters for a given population are:
   a. Initial prevalence
   b. Force of infection (FOI)
      i. Usually adjusted between 0.1 and 10. If the FOI has to be adjusted outside these values to fit the curve, databook values should be reviewed.
   c. Extreme (in)sensitivity to changes in force of infection may justify an interrogation of data and assumptions (e.g., number of unprotected casual acts unreasonably high).
   d. Inhomogeneity: adjust to alter the trajectory of the curve, as it increases the FOI below a given prevalence value and decreases the FOI above a given prevalence value.
   e. If changing initial prevalence, FOI, or inhomogeneity does not result in a good fit, examine interaction and transitions between populations.
3. **Prevalence** data: Adjusting the prevalence for a given population will also adjust prevalence for the other interacting populations. The magnitude of this impact will depend on the intensity of interactions and behavioral indicators.
4. **Diagnosis** data: consider adjusting testing rates as the model assumes all in a given population are equally likely to be tested, however, in practice some get tested regularly and some are never tested. Reducing test rates might result in a good calibration fit, as these are frequently over-reported in respondent surveys.

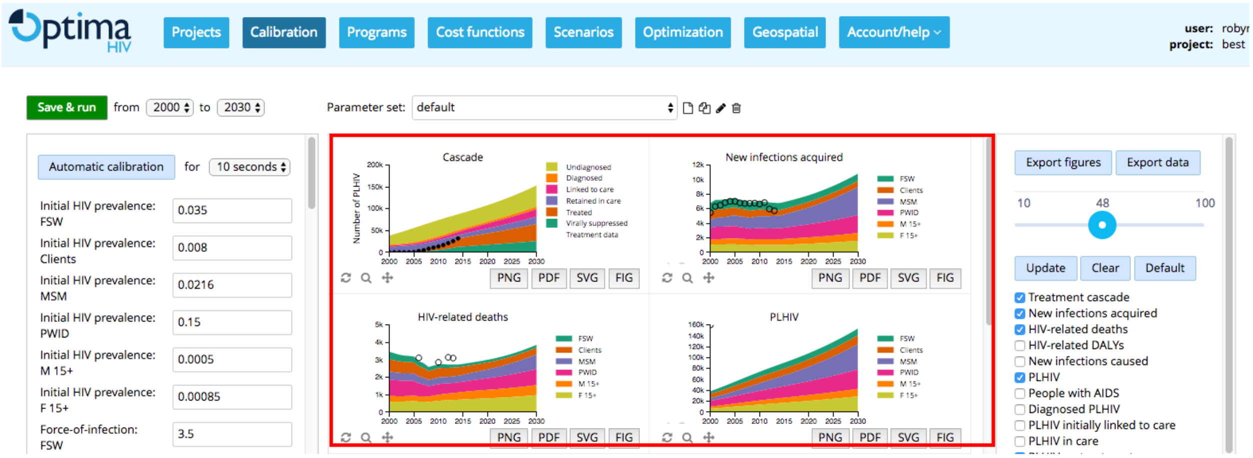
5. **Treatment** data: consider adjusting death rates for those on treatment or the treatment efficacy rates if treatment may not be as effective as expected in the given setting.
6. Review databook values:
    a. In particular the number of regular, casual, and commercial acts, condom use by type of sexual partnership (e.g. are values lower than expected due to poor usage or was their over-reporting of condom use?), condom efficacy constant value, HIV testing rates, number of injections, and percentage of people who receptively shared a needle/syringe at last injection
    b. Are there dramatic changes in annual trends from year to year (e.g. for people who inject drugs the number of injections in 2015 was 7, but was increased to 700 in 2016) may cause difficulties in achieving a well-matched calibration for certain populations.
7. Review the databook for inconsistencies:
    a. Model assumptions: e.g. the model will balance the number of sexual acts for females and males, are input values balanced?
    b. Incorrect entries in the databook: e.g. are the females aged 50 years and old incorrectly specified to age into the females 0-14 years group? Correct any incorrect values in the databook.
8. Click the green "Save and run" button at the top left.
9. For additional parameters to adjust, click the 'Advanced options' at the end of the parameter list.



Before (left) and after (right) example of a manual calibration, adjusting FOI and initial prevalence to calibrate the F15+ prevalence curve. The 2011 estimate was identified as an outlier and considered with less confidence than the other estimates. However, it was not completely irrelevant, as the curve still passes above the 2014 estimate. Had the 2011 estimate been considered irrelevant, the curve would have been made to pass through the 2014 estimate.

2.4. Graph display panel

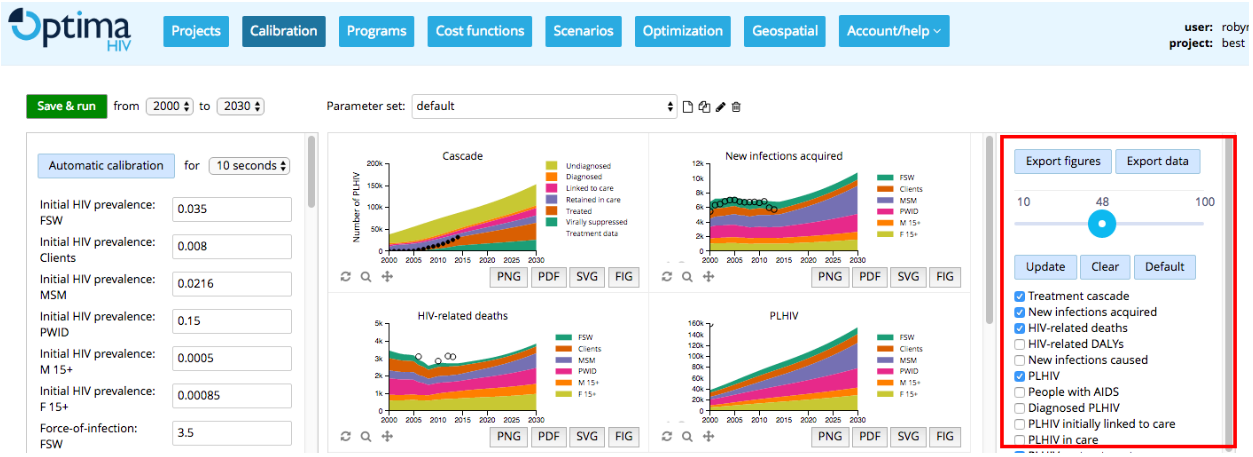
This panel allows you to view the model's projections of epidemic indicators.



2.4.1. Export a single figure

Underneath each figure, there is the option to save it as a PNG, PDF, SVG or FIG file. In most cases, you will want to export as a PNG (Portable Network Graphics), as this is the most flexible format for use in publications. For sharing with collaborators, either PNG or PDF (Portable Document Format) is preferred. If you wish to edit the figure in an application such as Adobe Illustrator, use either PDF or SVG (Scalable Vector Graphics) format. Finally, FIG (Figure) format exports the raw Python data, and is intended for advanced users.

2.5. Graph selection panel



2.5.1. Export figures

Click the "Export figures" button at the top left of the graph display panel to save a .pdf of all figures.

2.5.2. Export data

Click the "Export data" button to export the model's projections of the epidemic indicators to an xlsx file. Data is downloaded for all figures that are selected for display.

2.5.3. Graph size adjuster

The slider at the top of the graph selection panel allows you to modify the size of the graphs being displayed.

2.5.4. Graph selection

In the graph selection panel, you will see a list of all the possible results that you may wish to see. You can select/deselect different options using the checkboxes, and then click "Update" to update the display. Clicking "Clear" will deselect all graphs. Clicking "Default" will select the default selection of results to display.

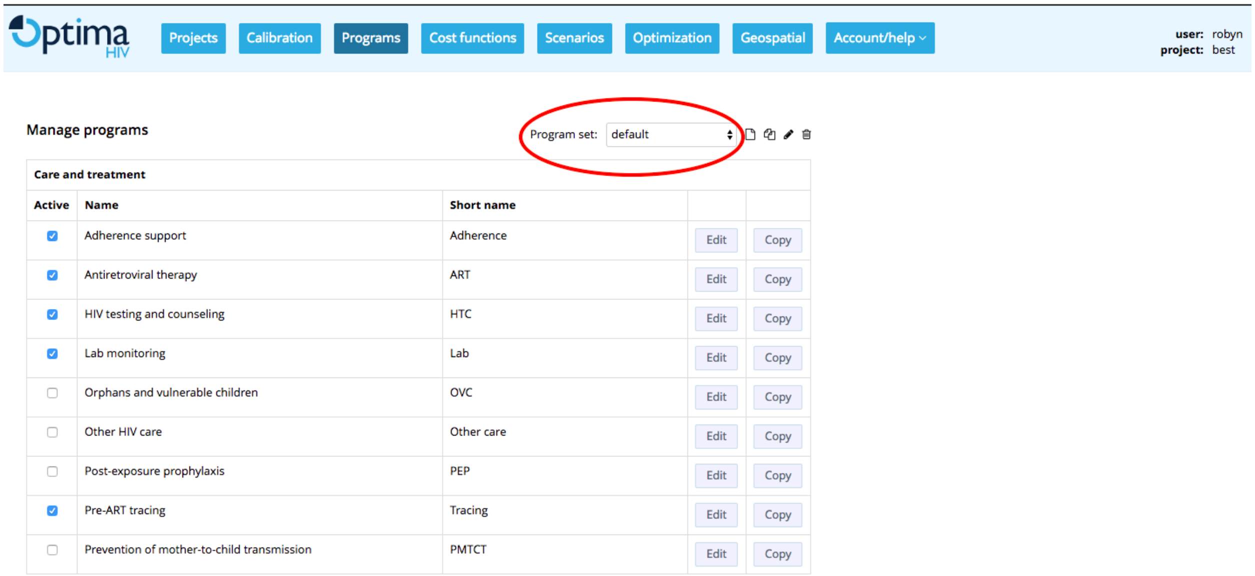
At the bottom of the graph selection panel, there is a button "Advanced". Clicking this will update the list of available results, and generate a longer list with more possibilities. Clicking "Normal" will revert to the previous list.

3. Programs, cost functions and outcomes

Once the epidemiological model is calibrated, a programmatic response is defined. This involves collating data on the current allocation of resources, and relationship between program spending, coverage and population-specific outcomes.

A set of programs is gathered together in a **program set**.

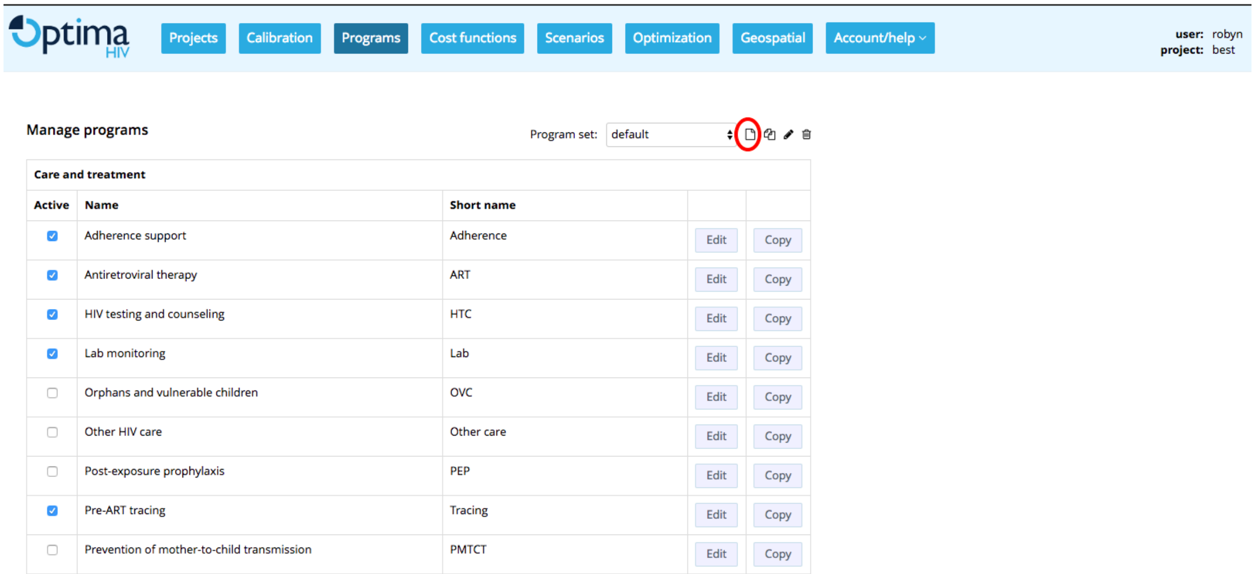
3.1. Program sets



Multiple versions of different programmatic responses can be saved as 'program sets'. Each program set contains details about the programs and how they operate (e.g., their costs and impact). Having multiple program sets is useful for a number of reasons:

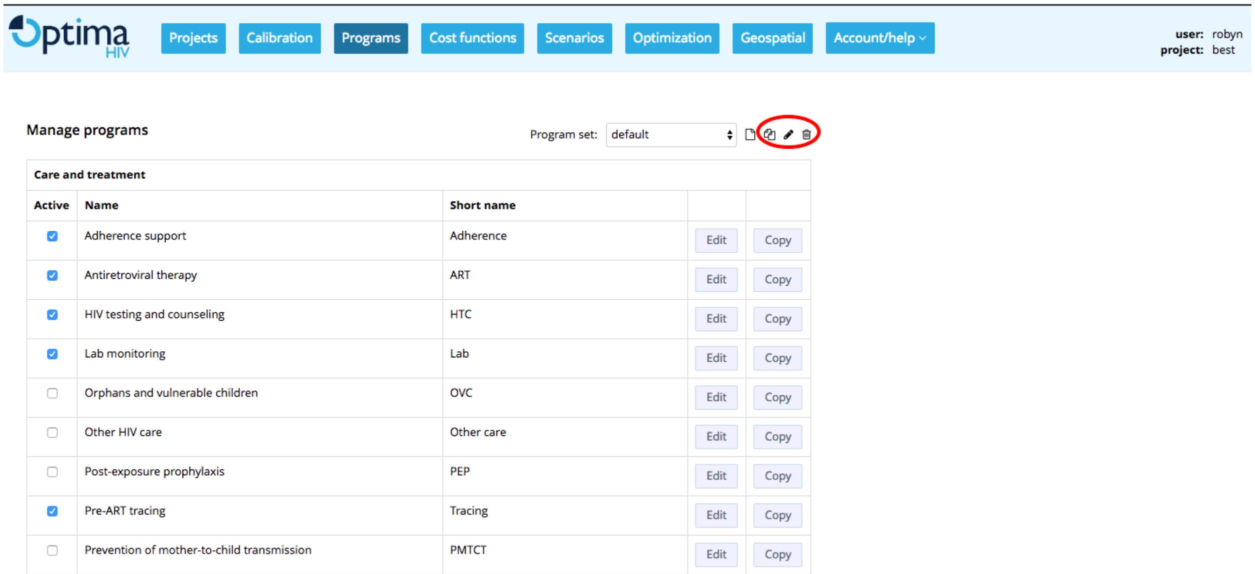
1. **Conducting technical efficiency analyses**. To do this, you could create one program set where the unit costs of programs were set to their current values, and an alternative version where the unit costs were 10% lower. You could then test the sensitivity of results to these alternatives by running scenarios and optimization analyses across alternative program sets.
2. **Conducting analyses for different funding sources**. Suppose you want to run two different analyses: one in which you investigate the optimal distribution of all HIV funds, and one in which you want to investigate the optimal allocation of national government expenditure only, assuming that international funders do not shift their allocation of funds. To do this, you would create two program sets. In the first, you would select the full set of programs operating in the country. In the second, you would only select the program operated by the government.

3.1.1. Create new program set



To create a new program set, click the "New" button to the right of the program set selection button. This will launch a dialogue box where you can decide on the name of your new program set.

3.1.2. Copy/rename/delete new program set



To rename, copy or delete a program set, use the buttons to the right of the program set selection button.

3.2. Programs

Optima HIV comes with 27 predefined programs. These reflect the basic National AIDS Spending Assessment (NASA) categories that form the basis of understanding HIV resource allocations. The rules for editing, copying, and adding new populations are identical with programs; the only difference is the options available. As with populations, the "Short name" is what will be used in Optima outputs (e.g., plot labels and legends).

> Optima HIV classifies HIV programs into two types. Direct programs are those with a clear potential impact on reducing HIV transmission, morbidity or mortality. Indirect programs are those that do not have a direct impact on health or where data limitations or classifications prevent a clear attribution to population specific outcomes. These include human resources, enabling environment, social protection, monitoring and evaluation, programmes for HIV-related orphans and vulnerable children (OVC), other HIV care, and other non-targeted programmes.
>
> Management costs may be classified separately to the specific program they relate to. For instance, ART costs may only include procurement of the drugs themselves and exclude staff wages and facility overheads relating to the provision of this treatment. Where possible, such management costs that can be reliably attributed to a specific program should be incorporated into the unit cost for that program.
>
> Where overheads contribute towards goods and services across programs or even outside the HIV budget (e.g., blood monitoring), values should be apportioned appropriately. See the Indicator Guide for a full treatment.

> You can include programs or modalities of service delivery which do not currently exist but are being considered for future implementation.

The following criteria should be used for dividing the national HIV response into specific programs:

- the program plays an important role in the national HIV response;

- there is historical program coverage information for the program if it is a ‘direct’ program targeting a population (or in case of new programs at least an understanding of potential coverage targets);
- past expenditure data (or budget allocation data) are easily accessible for the program (for example: the program corresponds to a standard category in the National AIDS Spending Assessment);
- there is research evidence on the efficacy or effectiveness of a HIV program for a specific population.

Just as for populations, an important consideration is to keep the number of programs manageable, for the resulting analysis to be robust.

Other programs should be included in categories defined in Optima HIV such that the entire budget can be assessed.

In principle all programs can be linked to a specific parameter in the model. For example, condom promotion and distribution would be linked to the percentage condom use in casual sexual acts. However, not all programs have to be linked to a specific outcome. Some cost categories, in particular enablers or synergies will not have a specific measurable effect on an impact metric in the model – such as HIV incidence or AIDS deaths - and will need to be treated as fixed costs. Fixed costs could be reduced through technical efficiency assumptions.

The choice of programs should correspond to the choice of populations groups included in the Optima HIV analysis. Therefore if there is an FSW and an MSM population it will be useful to model FSW and MSM focused programs. This could also be useful, if such programs are not yet in place, as the Optima HIV analysis could then be used to assess the potential impact of such a program. Optima HIV can also model the effect of new programs like oral Pre-Exposure Prophylaxis based on cost estimates and efficacy data from Randomized Control Trials (RCTs), along with assumed adherence levels.

With Optima HIV , it is also possible to add user-defined programs as required, for example programs for user-defined populations such as prisoners, migrants or others. In concentrated epidemics, these may include particularly intensive outreach programs for a key population or a program targeted at a specific population like prisoners. In generalized epidemics, these may include a particularly intensive multi-pronged program for young women and their male partners, which would be expected to have higher efficacy than general SBC spending. For adding any such additional programs, it will be important to consider whether information on program effects is available from studies or from evaluations.

The total group of programs should reflect the total national HIV spending as reflected under a NASA (past spending) or a National Strategic Plan (future budgets). If there is a group of small expenditure items, which do not fit under any of the categories, there could be room to include them under a separate category like "Other HIV programs".

3.2.1. Add a program to a program set

There are three different ways of adding a program to a program set:

1. Any predefined program can be added to a program set by selected the checkbox to mark it as active.
2. The "Add program" button at the bottom of the program list can be used to add a new program that isn't on the list
3. Programs can be copied from existing programs by using the "Copy" button

3.2.2. Editing a program

Click the 'Edit' button to define exactly which populations and outcome parameters are affected by the program.

3.2.2.1. Name

Enter the full name of the program.

3.2.2.2. Short name

As with populations, the "Short name" is what will be used in Optima HIV outputs (e.g., plot labels and legends).

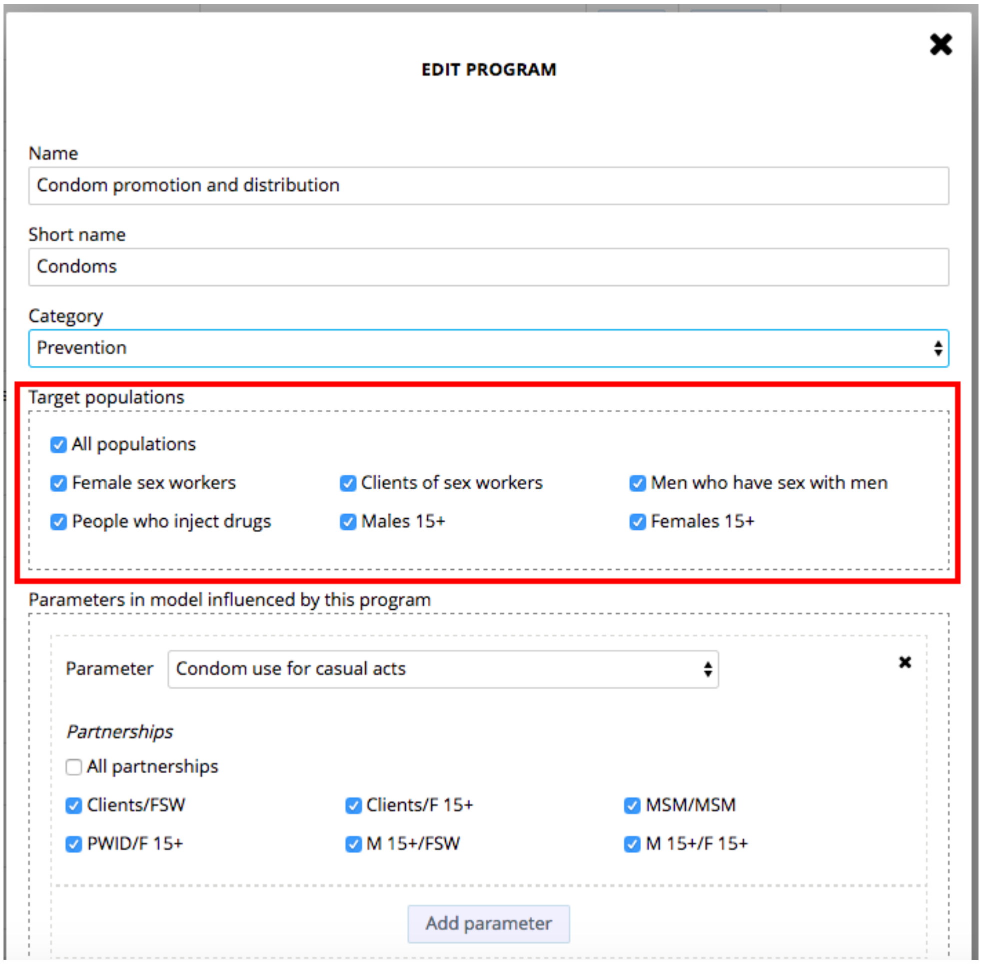
> You won't be able to save changes to the program until you have given it a short name.

3.2.2.3. Category

Programs can be classified as "prevention", "care and treatment", "management and administration" or "other". These categories are not used within the model; they are for user guidance only.

3.2.2.4. Target populations

Here, you select the populations that are targeted by the program.



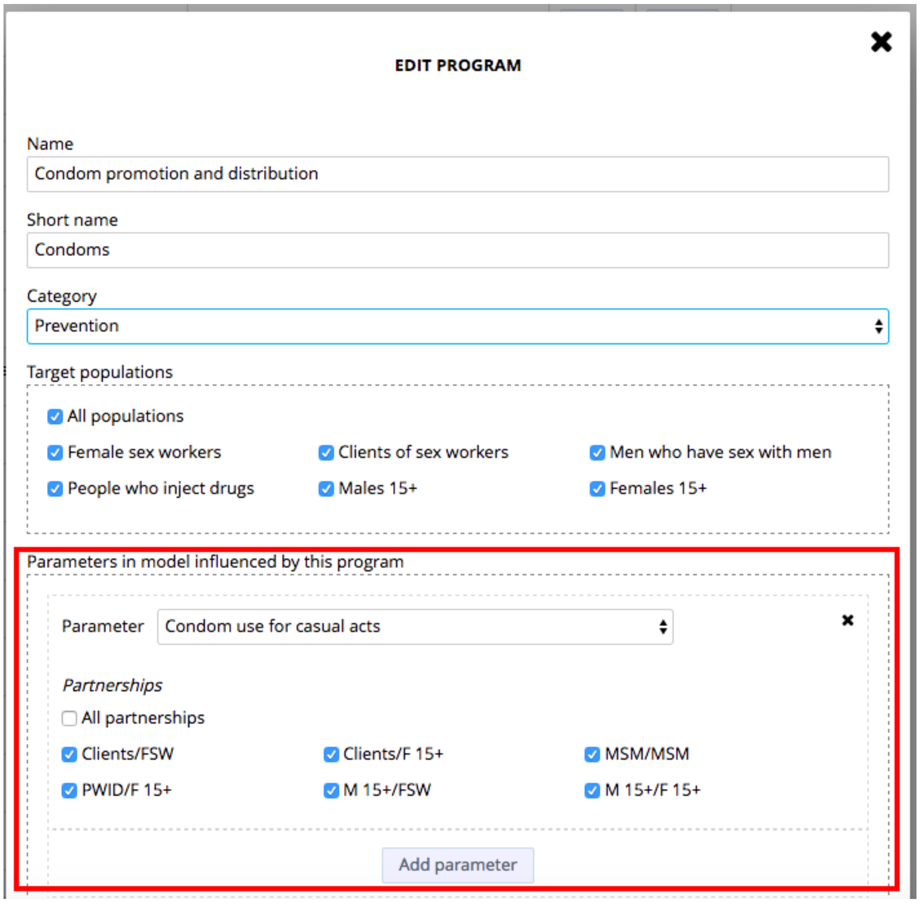
NOTE: For some programs, the default selection of the target population is "All populations", and there is no possibility of modifying this. This is the case for ART programs and lab monitoring programs, for example. This is because programs like this are not typically targeted according to particular sub-populations; ART is offered to all people regardless of their age, sex or behavioral tendencies. There are, however, distinct programs that aim to get sub-populations into treatment, and these are considered separately to the ART program itself. For example, you may wish to define a program for "Linking FSW to treatment". To do this, you would:

1. Create a program using the "Add program" button
2. Enter the program name, short name and category
3. Click the "Add parameter" button
4. Select "Female sex workers" as the target population
5. Select "Average time taken to be linked to care (years)" from the drop-down list of parameters
6. Select "Female sex workers" as the population associated with this parameter

7. Save the program.

3.2.2.5. Parameters in the model influenced by this program

Here, you select the model parameters that are targeted by the program.



- ➔ Pre-defined programs come with recommended default setup but these can be edited. Select target populations covered by the program, parameters influenced by the program and, for each, specify whether the program influences the parameter for all or only some of the covered populations.
- ➔ To add a parameter from scratch, simply "Add parameter" and repeat the preceding three steps.

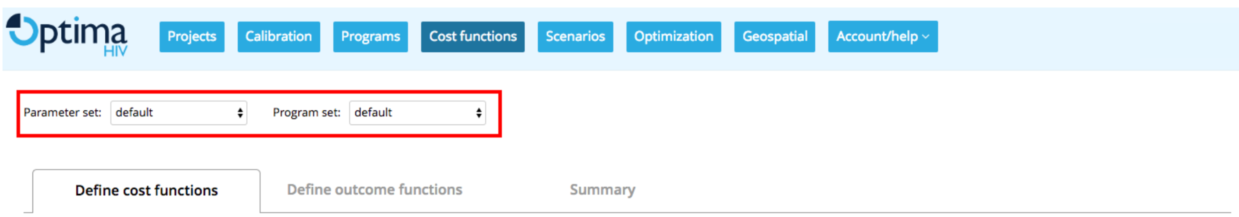
Users are encouraged to include only the most important 1 or 2 channels through which the program will directly affect the epidemic. For instance, a program for MSM would typically target casual condom use by MSM as well as their testing rates.

3.3. Defining cost functions

At this step, you provide more details about how the programs that you defined in section 3.2. operate. This means that you will (a) add data on how much has been spent on the programs and the coverage levels that they have achieved, (b) add information about the unit costs of operating the programs and (c) add information about the impact of the program on those who are covered by it.

The "Cost functions" page has three tabs: "Define cost functions", "Define outcome functions" and "Summary".

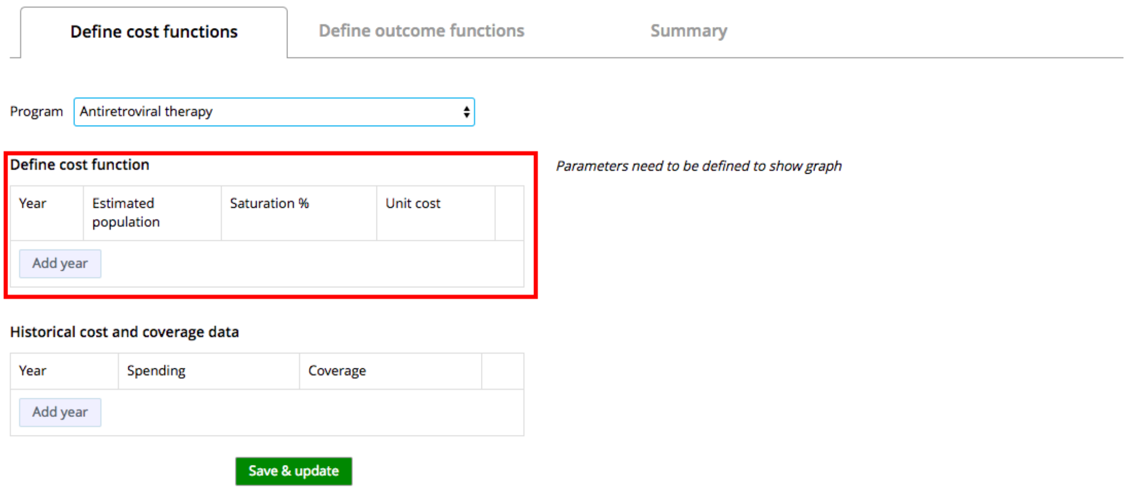
As the first step, you need to select the parameter set and program set that you want to define the cost functions for.



3.3.1. Define cost functions

On the page "Define cost functions", a drop-down menu will list all the programs associated with your selected program set.

3.3.1.1. Define cost function



You can add different cost functions for different years. To get started, click "Add year". This will add a row to the table, automatically populated with the year 2016 in the first column. You can change the year, and this will automatically update the number in the "Estimated population" column, to give an estimate of the target population size in the chosen year.

You will need to enter values for the **saturation** and **unit cost** of the program. There are two boxes for each of these, enabling you to enter a range.

> Optima uses non-linear cost curves to incorporate increasing marginal costs at higher levels of coverage, reflecting the increased expense of accessing the 'last mile' for a given population. The **saturation** point specifies the proportion of the population that is accessible by a given program, as a percentage of the population remains inaccessible regardless of the level of coverage of a given program, typically due to geographical or social (stigmatization) factors. The functionality could also be used to model a program where accessibility is limited to a proportion of a given population, such as public-sector workplace programs (where saturation/maximum coverage would reflect public-sector employment share).
>
> The relative gradient of the cost curve is a key driver of the optimization results. A steeper curve implies a lower **unit cost** (greater coverage for given expenditure) and a generally greater likelihood of being prioritized in the optimal budget mix.

Define cost function

Year	Estimated population	Saturation %	Unit cost	
2016	25773258	-	-	
Add year				

Defining a cost function for only a single year implies a constant unit cost over time, 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 curves 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, the user could split the program into two programs with different names (eg. provider-initiated HTC and HTC outreach). These programs would have two separate cost-outcome curves and the model would assess which one will be more effective.

3.3.1.2. Past spending and coverage data

Click "Add year" to enter historical cost and coverage data. These are not directly used to generate optimization results but rather are used to fit the cost curve. That is, ultimately, it is the cost curve that drives results, not historical data.

Historical cost and coverage data

Year	Spending	Coverage	
2016			
Add year			

> All cost and spending data must be in nominal terms and in a consistent currency.

> Coverage values must be in terms of people covered *not* goods or services provided (e.g., numbers of people covered by a condom or testing program, not the number of condoms distributed or tests conducted). The Indicator Guide provides detailed guidance.
>
> To include a program that doesn't presently exist, simply enter zero spend and zero coverage.

The standard approach in Optima HIV is to use actual HIV spending data from NASA or comparable reports produced for Global AIDS Progress Reporting for determining program costs.

The quickest way to get from NASA data to the actual spending data required in Optima HIV will be to do the following:

- Add a column to a NASA/GARPR expenditure table,
- Enter in this column a code (example short name) for each program the user selected in Optima HIV,
- Calculate the totals for each Optima HIV program category,
- Check that all spending in the NASA/GARPR is assigned to a program in Optima HIV.

In some cases, the country team will have additional knowledge on how to classify specific expenditure items. The user will also be required to enter specific coverage information for each program, Optima HIV will establish relationships between cost and 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 costs in the table described in section 3.3.1.1. 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, and use this to infer future relationships.

The historical program spending data should include all investments, which would not exist, if the program was not being implemented. Therefore program specific management cost (eg. administrative cost of an NGO working with female sex workers) should be included in the program.

Example: The national prevention program for female sex workers reached 300 FSW for a total cost of 150,000 USD. The cost includes a national program manager, M&E officer and administrator who all spend a good part of their time planning for program scale up. The cost per person reached in this program was 500 USD, but it is very likely that this will go down if more FSW are reached. In this case you can refer to default cost-coverage data from international literature. For example you could estimate that fixed cost is 100,000 USD and that actual service delivery cost per FSW reached is 80 USD. In this case, total cost in the next year with a target of 2,000 FSW would be 260,000 USD (100,000 USD + 2,000*80 USD) and in the year after with 5,000 FSW reached 500,000 USD (100,000 USD +5,000 *80 USD). This information would shape your cost-coverage curve in a way that will reflect fixed and variable costs.

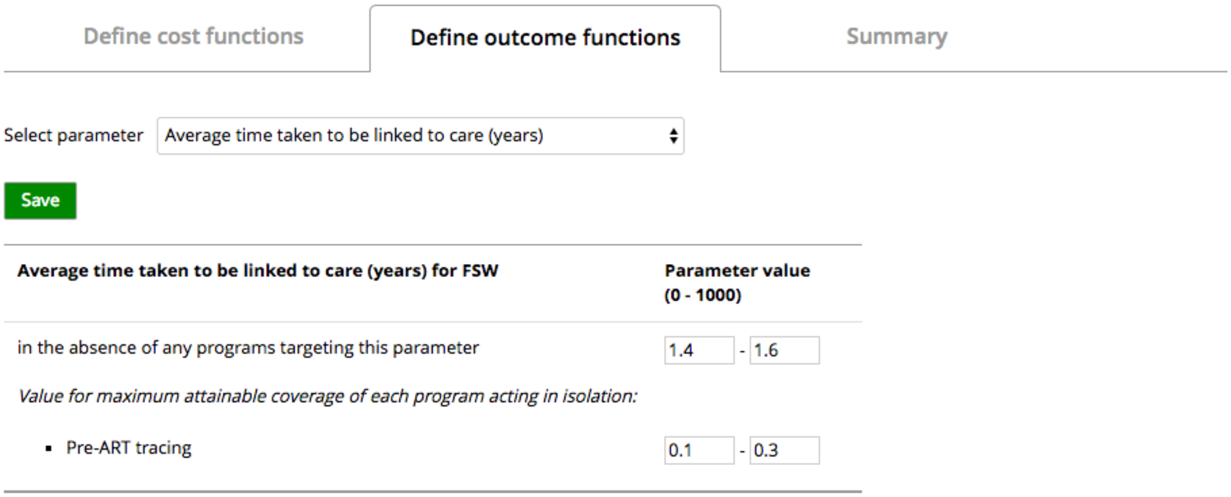
3.3.1.3. Cost function graph panel

Once you have entered data in the "Define cost function" table, a graph will appear on the right reflecting the cost function generated with these values. There are options to adjust 'maximum spending shown' to zoom into/out of the curve, and notes on any underlying assumptions and limitations can be left in the 'remarks' box.

3.3.2. Define outcome functions

Having defined the relationship between program spending and coverage, we now need to specify what happens when someone is covered by the program. When condoms are distributed to an individual, how does their condom use change? What proportion of individuals invited to a testing facility actually turn up and get tested?

On the page "Define outcome functions", a drop-down menu will list all the parameters associated with your selected program set.



For each parameter, there will be a table listing all the populations to whom this parameter applies. This is specific to a particular program set; for example, you may have defined one program set in which there is only one testing program, and another program set in which there are multiple testing programs.

For each parameter you must enter some data specifying how it is affected by particular programs.

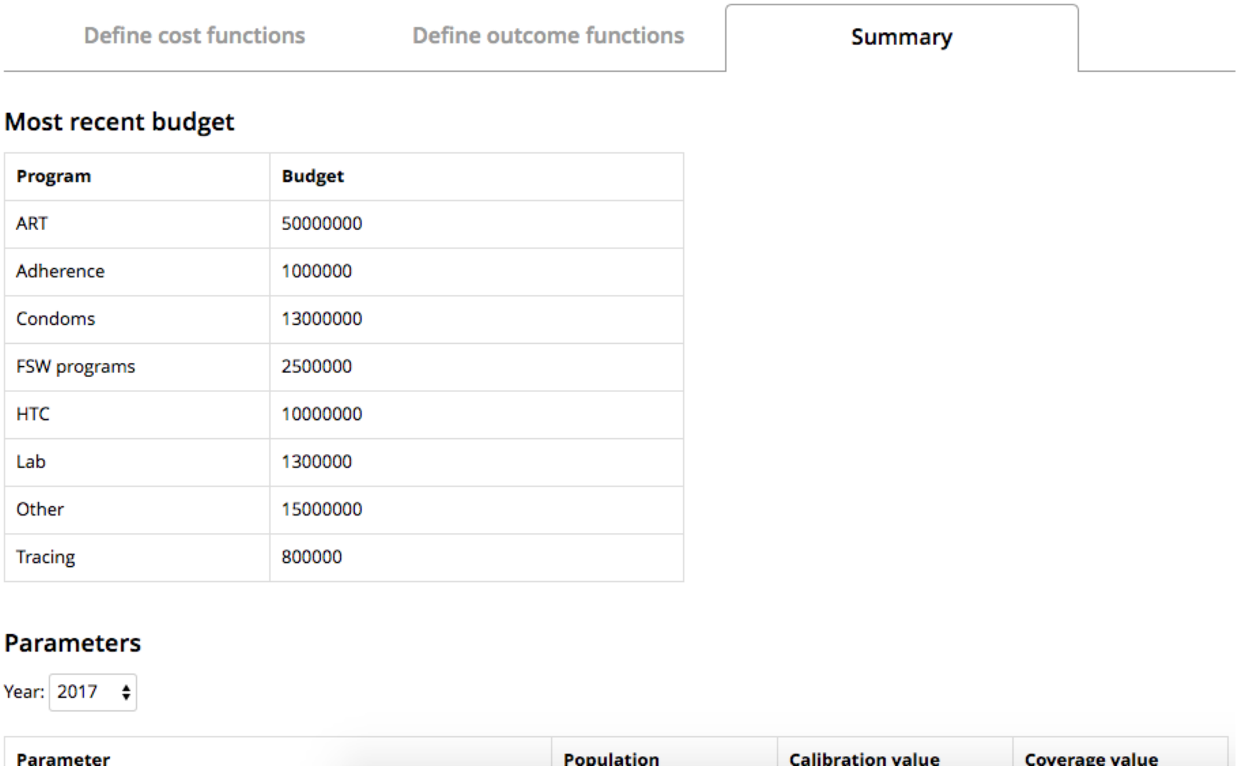
- ➔ Begin by defining outcomes: "...in the absence of any programs targeting this parameter". These values reflect background behaviour (say condom use or testing rates) in the absence of the program being considered and should incorporate goods and services accessed through alternative programs or private means (e.g., privately funded treatment).

- ➔ Define value "under maximal attainable coverage" of the program in isolation to reflect differences in real-world effectiveness between programs and modalities, and behavioral response of populations. See the Indicator Guide for further information.

  Certain parameters, notably *Numbers of people on treatment* and *numbers of pregnant women on PMTCT* require only background outcomes to be defined. For these parameters, there is no scope for behavioral response as individuals are either on or off treatment. Adherence and loss to follow up are modeled as separate parameters.

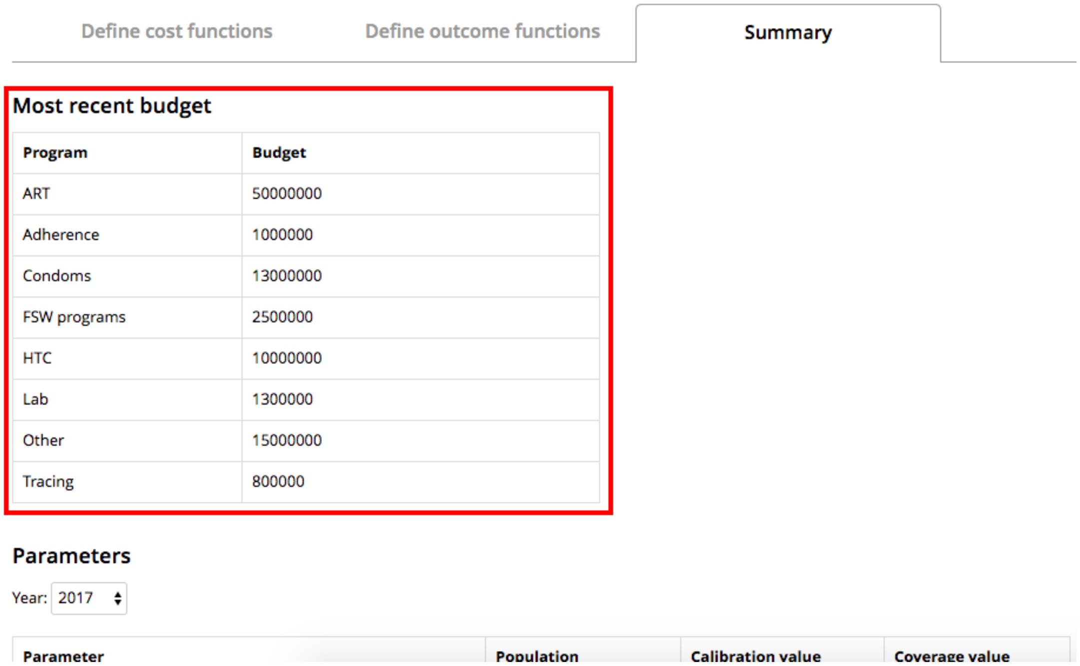
3.3.3. Summary

On this page, you will see a summary of all the programmatic data that you have entered.



3.3.3.1. Most recent budget

Here, you will see an overview of the most recent spending on each of the programs in your selected program set.



Program	Budget
ART	50000000
Adherence	1000000
Condoms	13000000
FSW programs	2500000
HTC	10000000
Lab	1300000
Other	15000000
Tracing	800000

The values in the "Most recent budget" will serve as the *baseline allocation of funds*. This is important when it comes to generating scenarios and optimizations. The status quo scenario assumes that the **total budget available is equal to the sum of the most recent spending on each program.** Moreover, status quo assumes that this same amount will be annually available and allocated as it was in the most recent year. It is therefore very important to check that the values in this table are an accurate reflection of the budget. To modify the amounts that appear here, you will need to go back to the "Historical cost and coverage data" table described in section 3.3.1.2., under the "Define cost functions" tab.

3.3.3.2. Parameters

Here, you will see an overview table comparing (1) the model parameter values defined by the calibration and (2) the model parameter values implied by the most recent budget (section 3.3.3.1). **It is very important that these align**. At the bottom of the table, there is a button "Reconcile" which will automatically reconcile any differences. You can also adjust these parameters by (a) modifying the parameters on the calibration page, or (b) adjusting the cost-coverage or coverage-outcome functions.

4. 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 Optima HIV analyses, scenarios and optimizations. In this section, we discuss scenarios.

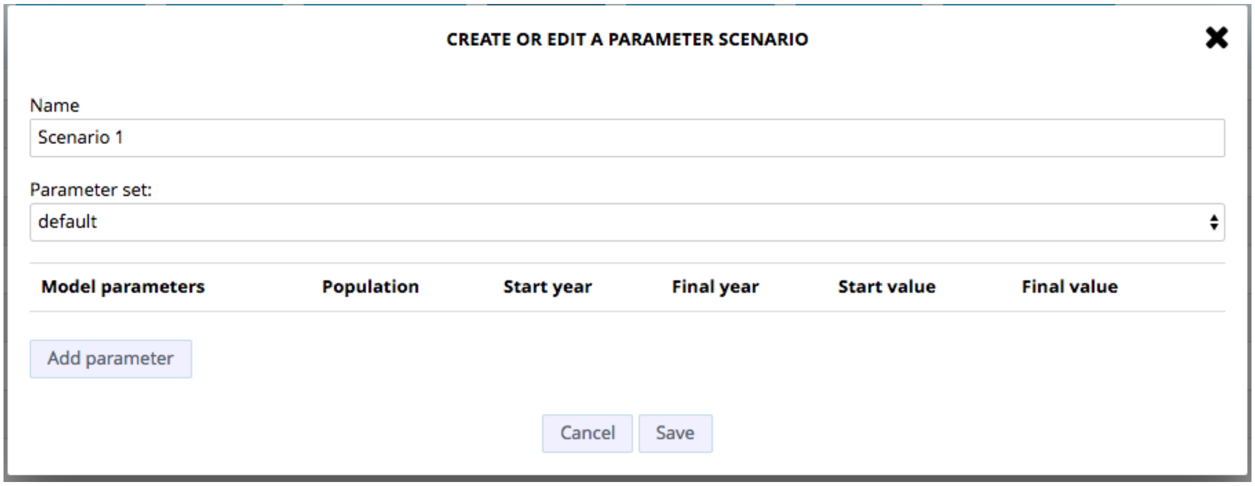
Scenarios are run when model parameters such as testing rates and treatment are adjusted directly or indirectly by varying program budgets or coverage. Scenario results reflect the impact of specified adjustments on the epidemic with all other variables unchanged from baseline.

There are three basic types of scenario: budget scenarios, coverage scenarios and parameter scenarios.

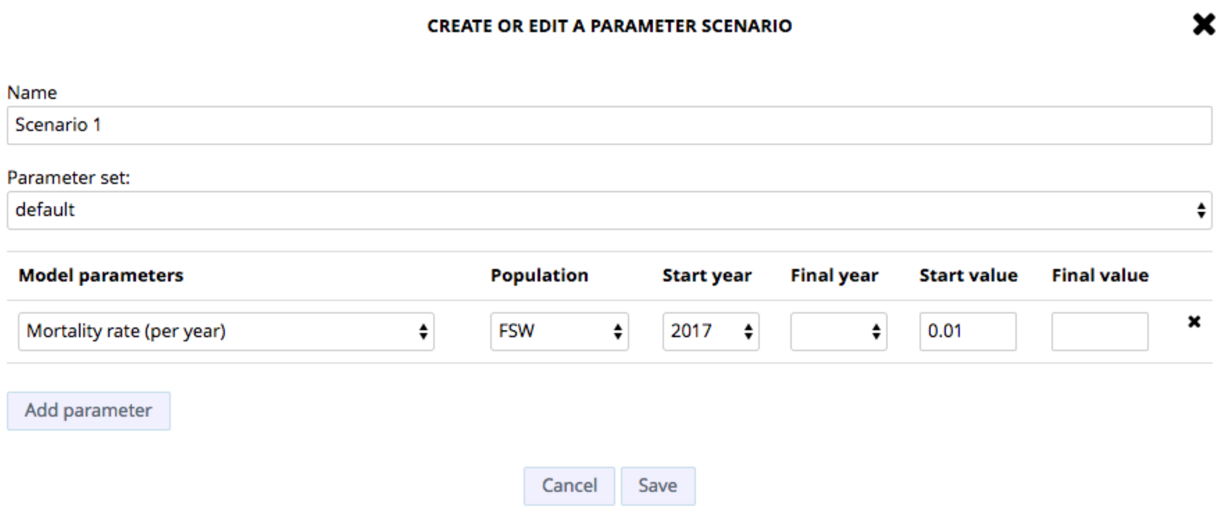
4.1. Parameter scenarios

These are the simplest kind of scenarios. You can run a parameter scenario even if programs have not been defined. The purpose of a parameter scenario is to investigate direct changes in epidemiological, clinical and behavioral model parameters. E.g., how many deaths could be averted by improving linkage to care?

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



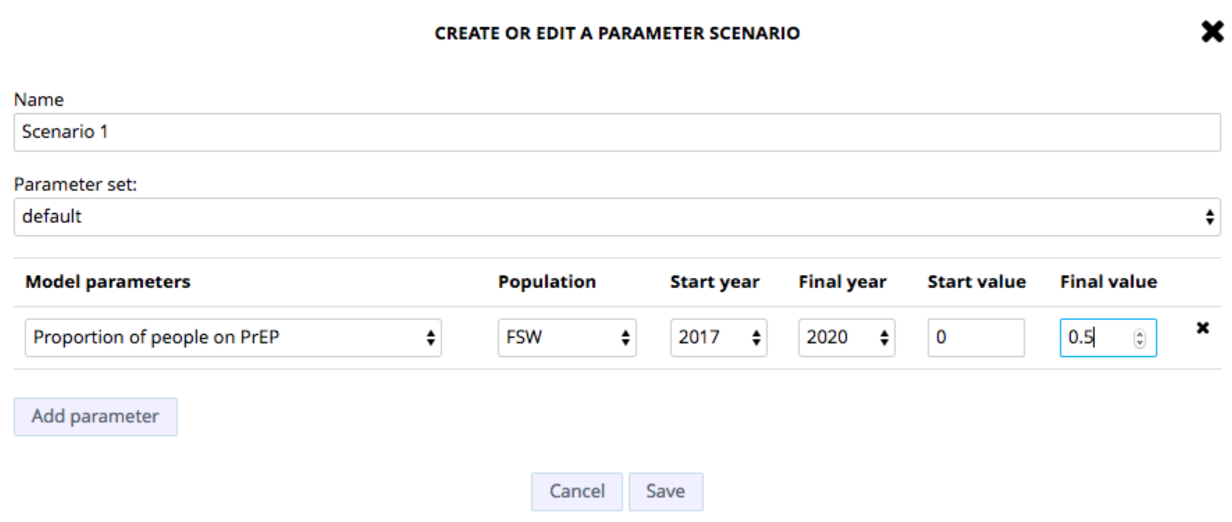
To start, give the scenario a name, and select the parameter set that you want to use as a baseline, then click "Add parameter".



1. Select the parameter that you want to modify from the drop-down list.
2. Select the population that you want to modify this parameter for.
3. Select the year that you want to modifications to begin. By default, this is 2017, but you could modify this if, for instance, you want to investigate a scenario starting in the future or evaluate what would have happened in the past.
4. [Optional] Select the final year for the parameter change. This is optional: if you do not select a final year, then the value that you enter for the "Start year" will apply to *all future years.*
5. The "Start value" will automatically pre-fill with the value of the parameter in the year selected as the "Start year". You can modify this if you want to; for example, you might want to investigate the effects of an immediate increase in condom use. More commonly, you would leave this as it is.
6. The "Final value" sets the parameter value in the "Final year".

Take care any rates or proportions are entered as decimals.

Example: The example set-up below shows how you would fill in the scenario values if you wanted to investigate the impact of increasing PrEP coverage among FSW from 0% in 2017 to 50% in 2020.



You can also define complex multi-parameter scenarios, by adding more parameters using the "Add parameter" button.

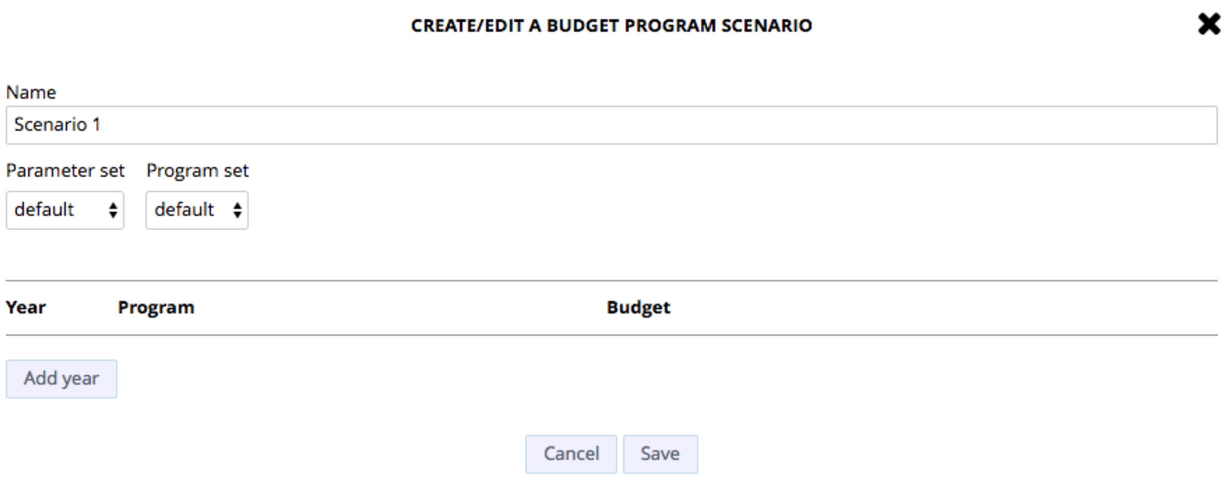
Once you have defined your scenario, click "Save".

4.2. Budget scenarios

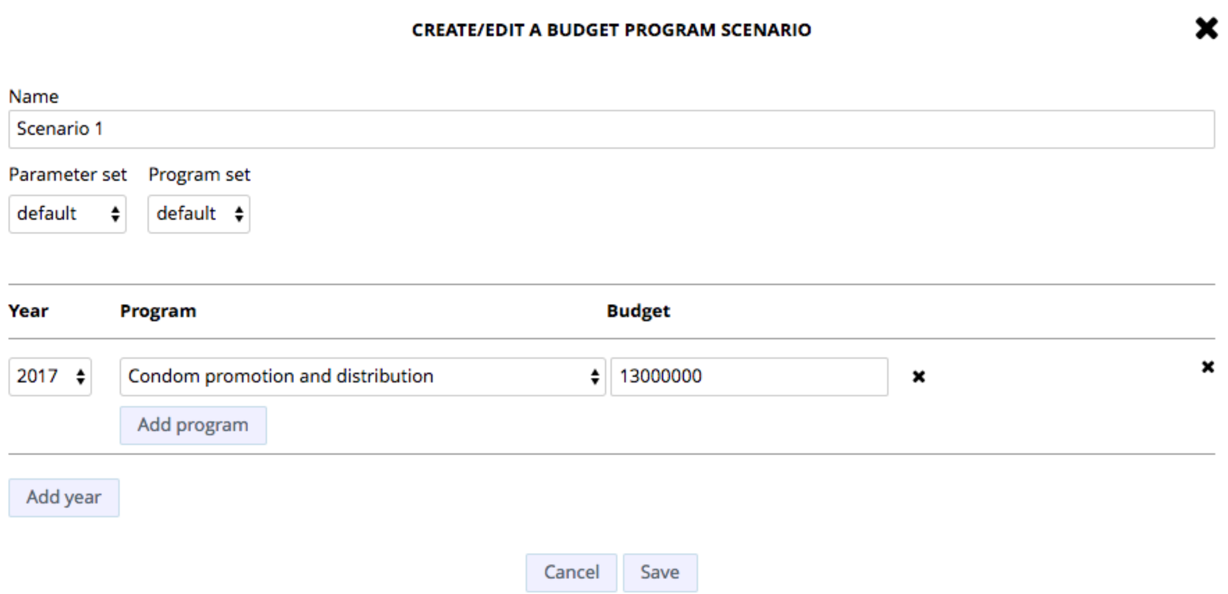
Budget scenarios estimate the impact of a change in spending for one or more programs. Only parameters targeted by the program are affected.

> A budget scenario example: How many infections and deaths have been averted by the introduction of the FSW program a decade ago? Reducing the FSW program budget to $0 effectively sets the targeted FSW testing and condom use parameters to their background values. All other programs in the model remain at baseline values. By isolating the parameters targeted by this program, we can evaluate the program's impact.

To add a budget scenario, click "Add budget 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 and program set that you want to use as a baseline. Then click “Add year” and then “Add program”.



1. Select the program that you want to modify from the drop-down list.
2. The “Budget” box will automatically pre-fill with the most recent budget for this program. The **baseline** scenario will assume that this same amount is annually available for the program. Enter a different value here to explore the impact of different budgets.
3. If you want to specify different budgets for multiple years, click “Add year”. This is how you would set up a scenario in which a program was gradually defunded over time.
4. If you want to specify different budgets for multiple programs, click “Add program”. This allows you to define budget scenarios for multiple programs at once.

4.3. Coverage scenarios

In coverage scenarios, policy questions are framed in terms of coverage numbers. A coverage scenario can be used to estimate the impact of changes in coverage for one or more programs. Again, only parameters targeted by the program are affected but this time by adjusting program coverage instead of spending.

> Take care to enter numbers of people when setting up coverage scenarios.

To add a coverage scenario, click "Add coverage scenario". This will launch a dialog box where you can define features of the scenario. This dialogue box looks identical to the "budget scenario" one.

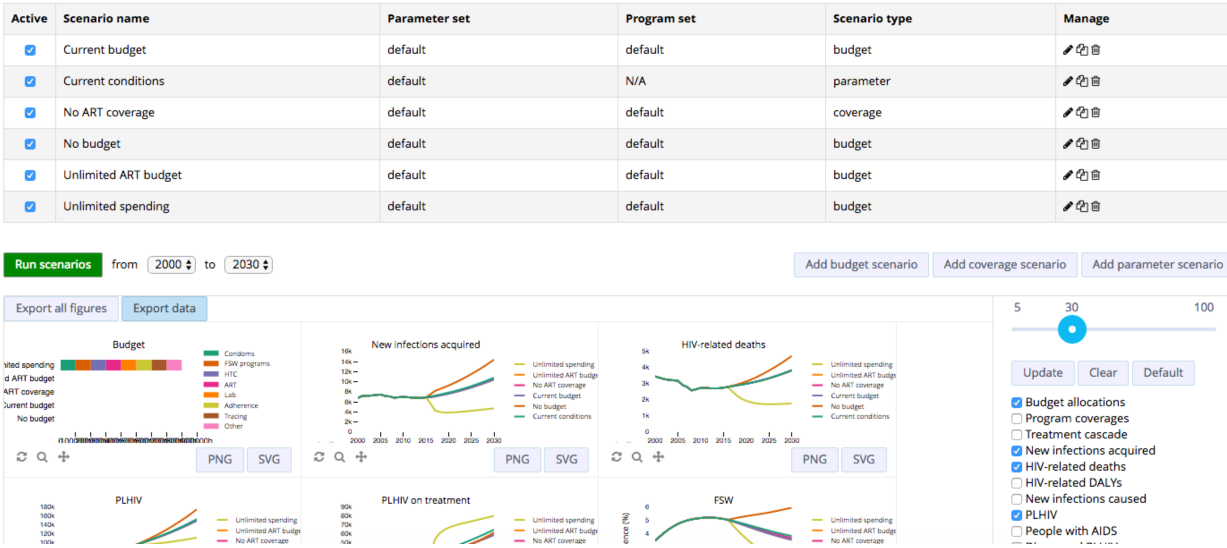
To start, give the scenario a name, select the parameter set and program set that you want to use as a baseline. Then click "Add year" and then "Add program".

1. Select the program that you want to modify from the drop-down list.
2. The "Coverage" box will automatically pre-fill with the most recent coverage level for this program. The **baseline** scenario will assume that this same number of people is annually covered by the program. Enter a different value here to explore the impact of different coverage levels.
3. If you want to specify different coverage levels for multiple years, click "Add year". This is how you would set up a scenario in which coverage of a program gradually decreased over time.
4. If you want to specify different coverage levels for multiple programs, click "Add program". This allows you to define coverage scenarios for multiple programs at once.

4.4. Viewing results of scenarios

After you have defined your scenarios, the next step is to run them.

Click the green "Run scenarios" button to generate results. There is an option to run the scenarios over different date ranges.



The results are displayed in a **graph display panel,** with additional options available in the **graph selection panel.** Please refer to sections 2.4. and 2.5. for instructions on how to use these elements.

5. Optimizations

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.

5.1. Add outcomes optimization

To add an outcomes optimization, click "Add outcome 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 and program set that you want to use as a baseline.

5.1.1. Set objectives

5.1.1.1. 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 (2017). 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. Typically, this period will correspond to the period of a national strategic plan. Analyses are typically performed for 3-10 years; it makes relatively little sense to optimize budgets for fewer than 3 years given the relatively limited impact it will have. Conversely, there is usually too much uncertainty in the future trends in the epidemic to optimize budgets for periods longer than 10 years. The most common period over which to optimize budgets is thus 5 years.

> A timeline example: How can existing funds best be allocated in order to get as close to possible to the national strategic targets, which call for a 50% reduction in new infections by 2020? For this policy question, you would set the end data to 2020.

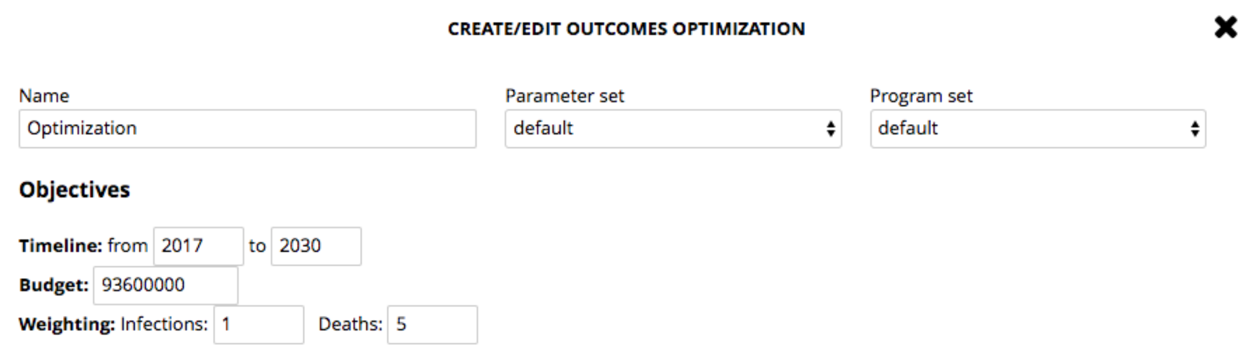
5.1.1.2. Budget

The budget is pre-filled with the sum of the most recent spending on each program in the program set (see section 3.3.3.1.). 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 5.1.2.1.), and will be allocated according to the allocation between programs given in section 3.3.3.1.

5.1.1.3. Weighting

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 or (c) a weighted combination of the two. Here, you specify the weighting.

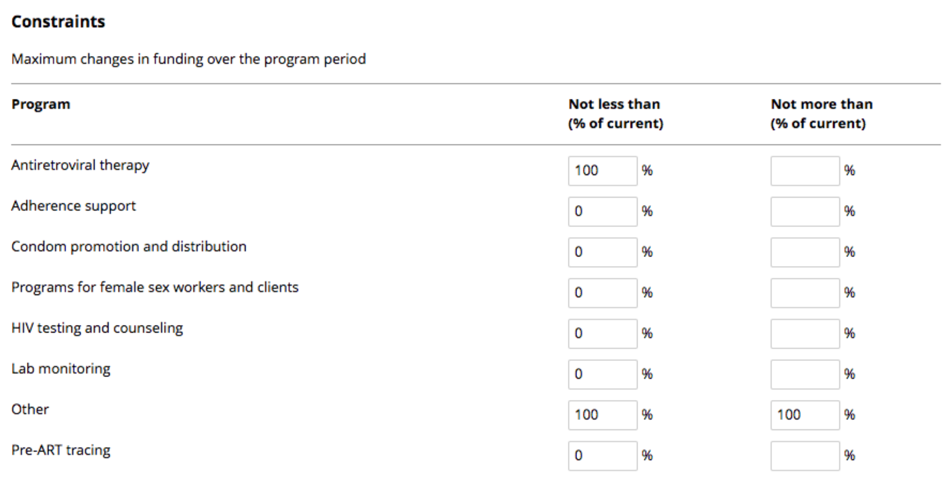
- Setting "0" for infections and "1" for deaths will minimize the total number of deaths over the optimization timeline (see section 5.1.2.1.). This will typically prioritize treatment programs, especially over the short-term, since it takes a long time for prevented infections to lead to prevented deaths.
- Setting "1" for infections and "0" for deaths will minimize the total number of infections over the optimization timeline (see section 5.1.2.1.). This will typically prioritize prevention programs, as these have the greatest impact on the number of new infections.
- The default weighting is "1" for infections and "5" for deaths, which is loosely equivalent to minimizing the total number of HIV-related DALYs. This option (which is recommended for most analyses) will produce an allocation balanced between those optimal for minimizing infections and those optimal for minimizing deaths, but will be much closer to the allocation that minimizes deaths, since deaths are by far the biggest driver of HIV-related DALYs.



5.1.2. Set 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 HIV 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 "Not less than" and "Not more than" are set to 100%, then that program will be treated as a fixed cost and will not be optimized.

They can be defined for each program (see example below).



Typical constraints:

1. Analyses would specify that funding to the ART program cannot fall below 100% of what it is currently. This is due to ethical constraints in taking people off treatment. Similar logic could apply to PMTCT programs and opiate substitution programs.
2. Analyses often specify that funding to non-direct programs (i.e., management or other administrative programs) should be constrained to remain at its current levels. This is because these programs are often essential for the continued running of the HIV response, but their effects cannot be quantified. Since the effects can't be quantified, it would not be reasonable to include them in an optimization analysis.

5.2. Add money optimization

To add a money optimization, click "Add money 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 and program set that you want to use as a baseline.

5.2.1. Set objectives

5.2.1.1. Timeline

When defining a money minimization, the analysis will find the smallest possible amount of money necessary to achieve some predefined target for reduction in new infections or deaths. Typically, this target is measured relative to some baseline year; for example, the target may specify a 50% reduction in new infections by 2020 relative to 2012. In the timeline section, you can define the years to establish these targets.

For example, if 2012 is set in the "compare to base year" box, and the analyses are run from 2017-2030, the optimization will find the optimal allocation of funds over the period 2017-2030 in order to achieve the objectives by 2030, where the objectives are measured relative to 2012.

5.2.1.2. Budget

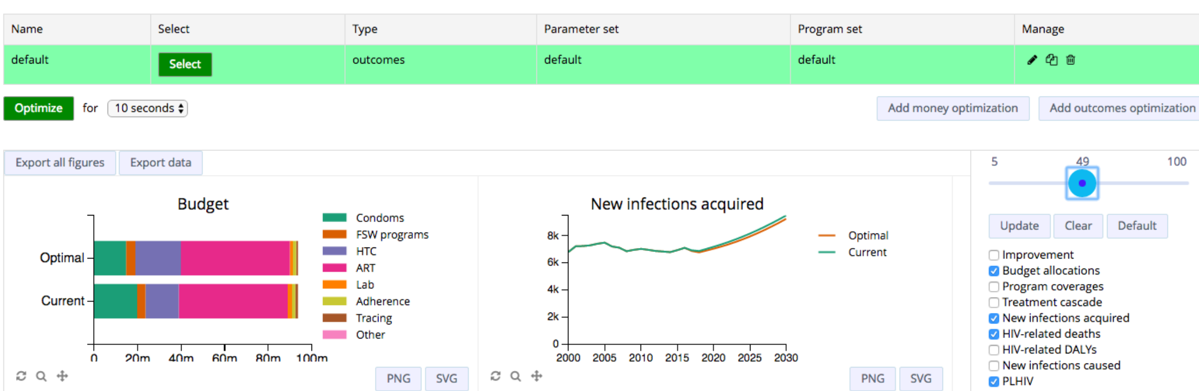
The budget is pre-filled with the sum of the most recent spending on each program in the program set (see section 3.3.3.1.). However, if you want to investigate a different amount, you can modify the budget total here. The baseline assumption is that the amount that you enter here will be annually available in each year over the period defined in your optimization timeline (see section 5.1.2.1.), and will be allocated according to the allocation between programs given in section 3.3.3.1.

5.2.1.3. Weighting

This specifies what the objectives are, and should reflect the targets and be linked to the timelines (see section 5.2.1.2.). For example, if the "Fraction of deaths to be averted" is set to 0.5, then the analysis will find the smallest amount of money required in order to achieve a 50% reduction in deaths. That is, it would find the smallest amount of money such that the number of deaths in the final year of the analysis is 50% of the number of deaths in the baseline year of the analysis. Multiple objectives can be defined by setting non-zero numbers for both new infections and deaths.

5.3. Viewing results of optimizations

After you have defined your optimizations, the next step is to run them. Click the green "Optimize" button to generate results. There is an option to set the time limit for running the optimization. The length of time required to perform an optimization depends on the complexity of the model and the number of programs being optimized. In most applications, 10 minutes typically suffices.



The results are displayed in a **graph display panel,** with additional options available in the **graph selection panel.** Please refer to sections 2.4. and 2.5. for instructions on how to use these elements.

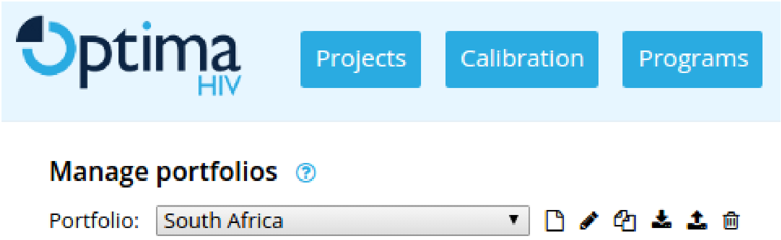
6. Geospatial optimization analysis

Geospatial dynamics of HIV epidemics are a primary consideration for decision makers and planners in developing an effective response. Where regions display different transmission dynamics and epidemiological trends, geospatial optimization analysis becomes increasingly important.

The geospatial analysis module of Optima allows users to optimize resources across budgets and, simultaneously, across districts. At minimum, a geospatial analysis needs region-level prevalence and population data. All other behavioral, demographic, and clinical parameters can be drawn from the national-level project.

6.1 Portfolios

Geospatial analyses are organized into *portfolios*. A portfolio is simply a container of two or more projects. Portfolios can be created, uploaded, downloaded, and deleted in much the same way projects can:



Because portfolios are containers for projects, any projects in your account that are ready to be optimized are available to add into the portfolio. When performing a geospatial optimization across disparate settings – for example, different countries – the typical approach would be to create each project individually, including the data entry, calibration, and definitions of programs and cost functions.

When projects are added to a portfolio, they are *copied* into that portfolio. Further changes you make to a project after adding it to the portfolio will not affect the portfolio. However, projects can be added to or removed from a portfolio at any time.

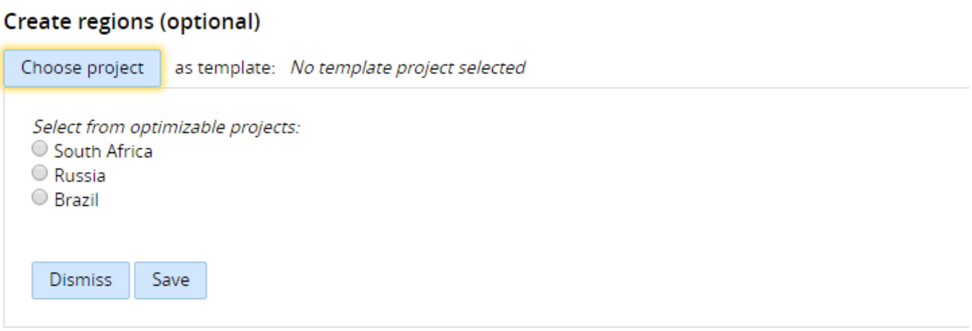
6.2 Creating regions

While geospatial analyses can cover regions with very different epidemics, typically, they cover regions within a similar setting – such as different districts within a country. In this case, region-level projects can be automatically generated from a single master project.

The master project is typically at the country level, which is then subdivided into two or more regions. The master project should be a completed Optima project: in particular, it must have all data uploaded, be calibrated, and have defined programs and cost functions. Any project that meets these criteria is available to be chosen as a master project.

Once selected, a master project is used to create a spreadsheet for filling in data on the region-level epidemics. The only data that are required at the region level are population size and prevalence. Furthermore, they are only required for a single year. If not all data are available in a single year, then choose the best recent year. It is expected that there are data on both prevalence and population size for each district (and preferably all estimates within several years of each other). If these data are not available, there may not be sufficient information to perform a robust geospatial analysis. All other data, including program data and behavioral data, are automatically transferred or scaled as appropriate to the region-level projects.

To select a master project, use the "Choose project" button:



Once clicked, it will bring up a list of available projects to select. Once one of these has been chosen, the option becomes available to create a spreadsheet based on this project:

The first option controls the number of subregions to create. This will be the number of regions in the geospatial analysis. The second option controls which year of the master project calibration to base the region-level projects on. This is used to calculate the population size and prevalence in the regional subdivision spreadsheet. While not used in the analysis directly, these numbers are provided since the total population size across all regions and the average HIV prevalence should match those of the master project.

> **NOTE:** The parameter and program sets that will be used to produce these estimates will be the most recent one created. If you wish to use a different parameter set, go to the project management screen, copy it to a new project, go to the calibration screen, delete any extraneous parameter sets, then go to the programs screen, and delete the extraneous program sets.

The created spreadsheet contains columns to enter the district-level population size and prevalence. Data entry should occur in the shaded column:

	A	B	C	D	E	F	G
1		M 15-49	F 15-49		Total (intended)	Total (actual)	
2	South Africa - region 1	0.00	0.00	OR		0.00	
3	South Africa - region 2	0.00	0.00	OR		0.00	
4	South Africa - region 3	0.00	0.00	OR		0.00	
5	South Africa - region 4	0.00	0.00	OR		0.00	
6	South Africa - region 5	0.00	0.00	OR		0.00	
7	South Africa - region 6	0.00	0.00	OR		0.00	
8	South Africa - region 7	0.00	0.00	OR		0.00	
9	South Africa - region 8	0.00	0.00	OR		0.00	
10	South Africa - region 9	0.00	0.00	OR		0.00	
11	---						
12	Calibration 2017	26,342,221.00	28,436,845.00		54,779,066.00	54,779,066.00	
13	District aggregate	0.00	0.00		0.00	0.00	
14							

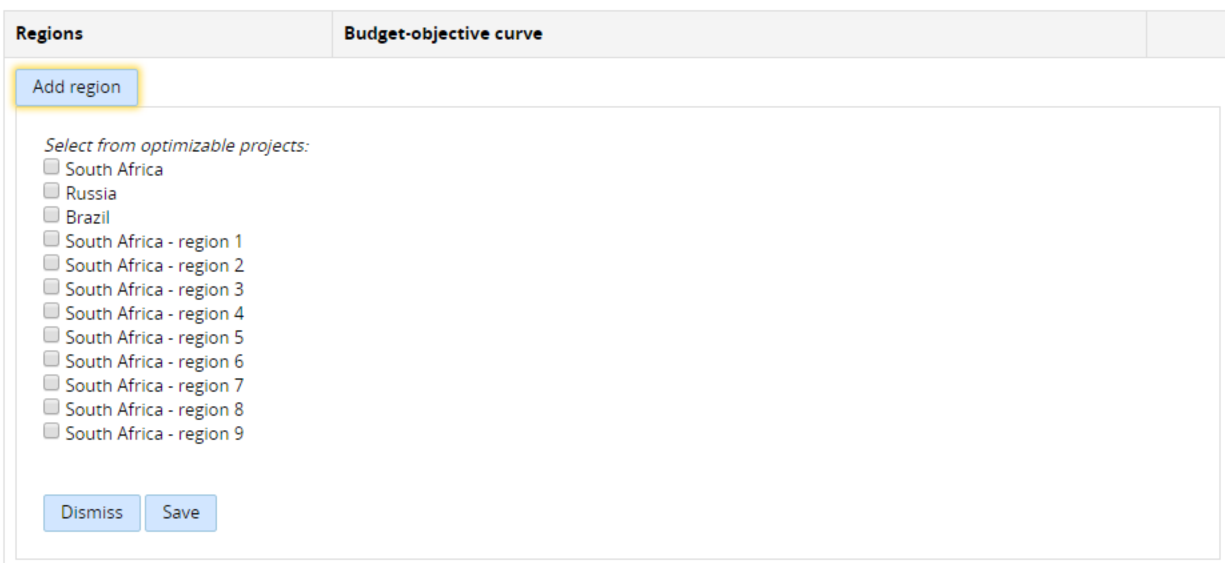
The remaining columns will be automatically populated. Please check that the district aggregate matches the calibrated value for each column. Once the population size spreadsheet is complete, follow a similar procedure for the prevalence worksheet. Unlike the main Optima data entry spreadsheet, no cells can be left blank.

Once filled out, click on "Upload spreadsheet", and your new projects will be created:

These projects will have the population size and prevalence specified in the subdivision spreadsheet, with other parameter values taken from the master project. Note: it is important that at this stage you open each project and check that it is sufficiently well calibrated. If any adjustments are required, perform calibration as per normal.

6.2.1. Adding regions

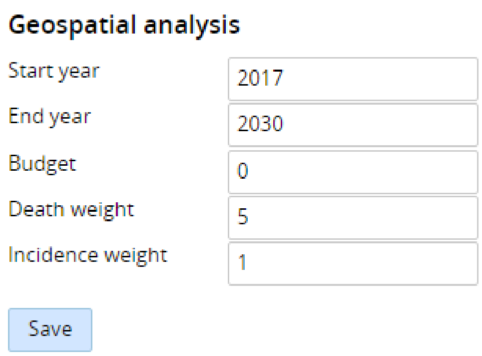
Newly created projects, as well as all valid projects previously present in your account, are available to be selected as regions in the portfolio:



Note that regions can be added or removed at any time (if, for example, a project needs to be updated). Remember that regions are projects that are *copied* into a portfolio, and thus future changes in the source projects will not affect the region projects in the portfolio. If you need to update a region project, delete it from the portfolio, load it in the project management page, modify it as desired, and re-add it to the portfolio.

6.3 Running analyses

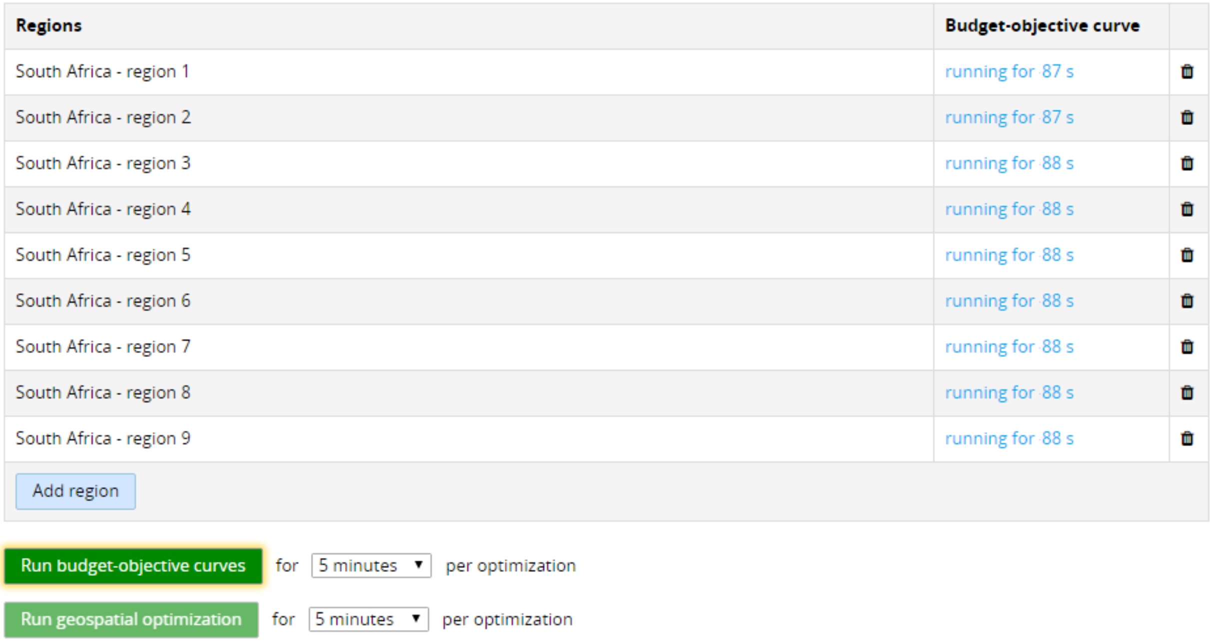
Because geospatial analysis is a type of optimization, the options for running are similar (see section 5.1.1 for an explanation of each field):



"Budget" refers to the total budget across all regions. If it is left as zero, then it will be automatically calculated as the sum across all regions. Note that constraints are not defined for portfolios. This is because each of the projects in the portfolio has its own set of constraints, and these are used instead.

6.3.1. Running budget-objective curves

Once the region-level projects have been added to the portfolio and the geospatial objectives have been set, the portfolio is ready to optimize. Click on the "Run budget-objective curves" button to begin this process:



An optimization is run independently for each district at different budget levels generating an optimal outcome (deaths and infections averted) for each budget level. This is known as a budget-objective curve (BOC), and it summarizes the relationship between a given budget and the maximum possible health outcome. Each district has a unique BOC reflecting differences in the epidemic. Districts that demonstrate a greater return on budget increases will be prioritized. Because the budget-objective curve has to be calculated at many points, and depending on the number of regions, the total run time may be up to approximately 10-50 times longer than the time per optimization shown. The right-hand column shows the status of the optimization for each region. Note that there is no need to keep your browser open during this time -- you may close it and come back to this screen at any time, and the optimizations will continue running in the background.

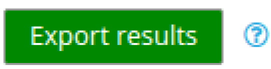
6.3.2. Running geospatial optimization

Once all budget-objective curves have been calculated, click on "Run geospatial optimization" to calculate the optimal allocation of resources between regions:

This algorithm examines every possible funding level to every region for a given budget amount, and assigns funds to the region with greatest marginal cost-effectiveness. This process is repeated until the entire budget has been allocated. Next, each individual district is re-optimized with its final total budget amount for the amount of time show. Because of this, you should expect a total running time somewhat longer than the number of regions multiplied by the time limit shown (for example, a 10-region project run for 5 minutes per optimization should complete in roughly 50 minutes). Note that this is an upper limit on run time, and in practice run times may be significantly less, if the algorithm is able to quickly converge on a solution.

6.3.3. Exporting geospatial results

Once geospatial analysis has finished running, results can be exported to Excel format:



(Note that due to the complexities of mapping geospatial results, graphical output is not available in this version of Optima.) The exported spreadsheet allows you to view optimal budget allocations i) across programs within each district and ii) across district budgets. Corresponding district-level program coverage and aggregate health outcomes (new infections and deaths) are also provided:

	A	B	C	D
1	**Geospatial analysis results: minimize outcomes from 2017 to 2030**			
2				
3				
4			**Initial**	**Optimal**
5	**Overall summary**			
6		**Portfolio budget:**	22,077,727.00	22,077,727.00
7		**Outcome:**	218,094.00	121,797.00
8		Deaths:	32,513.00	16,283.00
9		New infections:	55,529.00	35,583.00
10				
11				
12			**Initial**	**Optimal**
13	**Project: "Australia - WA/NT/TAS"**			
14				
15		**Budget:**	6,493,449.00	679,215.00
16		**Outcome:**	2,074.00	2,536.00
17		Deaths:	284.00	345.00
18		New infections:	654.00	744.00
19				
20		**Allocation:**		
21		HTC	5,903,135.00	425,777.00
22		ART	590,314.00	253,438.00
23				
24		**Coverage (2017):**		
25		HTC	440,085.00	56,499.00
26		ART	984.00	422.00
27				
28				
29			**Initial**	**Optimal**
30	**Project: "Australia - NSW/VIC/QLD/SA"**			
31				
32		**Budget:**	15,584,278.00	21,398,511.00
33		**Outcome:**	216,020.00	119,261.00
34		Deaths:	32,229.00	15,938.00
35		New infections:	54,875.00	34,839.00
36				
37		**Allocation:**		
38		HTC	14,167,525.00	6,113,860.00
39		ART	1,416,753.00	15,284,651.00
40				
41		**Coverage (2017):**		

The left column shows the initial program spending and coverage in each region, as well as the outcome (in terms of deaths and new infections) The right column shows the geospatially optimized equivalent. Note that although geospatial optimization will never increase the overall total outcome, it may increase the outcome for certain subregions.

7. Additional support

Additional support for using Optima is provided via training workshops. Online support is also available. To inquire about an Optima training workshop or for any questions on using Optima not covered here, please email us at **info@ocds.co**.