

Guidance on model outputs: how to generate and upload estimates

Guidance for touchstone 202310

Overview

For the 2023 full model runs, you will need to provide the following:

Item	How to create	Where to upload complete file(s)	Due by
Central estimates (multiple files)	Download central burden estimate template from Montagu, use this to create one file per scenario	Montagu	4 December 2023
Stochastic estimates (multiple files)	Download stochastic burden estimate template from Montagu, use this to create as many files as you need.	Dropbox	29 January 2024
Parameter set (1 file per disease)	Download stochastic parameters template from Montagu, use this to create your parameter set.	Montagu	29 January 2024
Parameter certificate (1 file per disease)	You will be able to download this from Montagu once you have uploaded your parameter set.	Dropbox	29 January 2024

Countries

Modellers of hepatitis B, HPV, measles, COVID-19 and rubella will provide central and stochastic estimates for 117 countries. Modellers of cholera, meningitis A, typhoid,

To contact us, please email montagu-help@imperial.ac.uk.

malaria and yellow fever will provide central and stochastic estimates for the agreed subset of these countries.

State of your model

You should complete any key improvements to your model before you start on the model runs, and keep track of this 'frozen' state of your model.

Age groups

The age groups in your burden estimate templates must be 1-year age groups. If your model uses larger age groups, you will need to disaggregate these.

DALYs guidance

[Download report with detailed guidance on DALYs.](#)

Main output columns:

For most groups, the main output columns in the burden estimate template are cases, DALYs, deaths and YLL. In the example row below, in cells G2/H2/I2/J2, you should enter the number of yellow fever cases/DALYs/deaths/YLL that occurred in the year 2000 among 0-year-olds in Angola, for the scenario in question.

	A	B	C	D	E	F	G	H	I	J
1	disease	year	age	country	country_name	cohort_size	cases	dalys	deaths	yll
2	YF	2000	0	AGO	Angola					

Cohort size

The cohort size is the number of people alive in a given birth cohort specified by the calendar year and age during that year – so it will be the same across all scenarios. Taking the above row as an example, in cell F2 you should enter the number of 0-year-olds in Angola in the year 2000. You must include the actual cohort size for each row; Montagu will not accept 'NA' as a value.

We will then be able to calculate the number of FVPs (fully vaccinated persons) by multiplying this with the relevant coverage. The cohort size should match to the interpolated population provided on Montagu. The cohort size should reflect the age range, time range and gender (female, male or both) for which your model is tracking the population.

-----, please enter [country name](#) in column E

Checklist for avoiding errors when uploading to Montagu:

- Your file must not contain any empty cells or 'NA' values
- Your file should not contain negative burden outcomes
- Your file must contain the exact same rows and columns as your burden estimate template
- Values should not contain commas (e.g. 1395 not 1,395)
- The demographic/coverage data may include years that are outside the scope we are asking you to provide estimates for. Therefore, you should go by the years that appear in your template.
- If your model doesn't provide something that the template requires, please [contact us](#).

*Please provide central estimates with **at most** 2 decimal places or precision, for stochastic estimates, please round to the nearest integer.*

Central estimates

First, download your central burden estimate template from the Responsibilities page (listed under 'Scenarios'). These templates are customised for each model.

You will need to use this template to create one file for each scenario, ensuring that you fill in all rows and columns. Montagu will now only accept your central estimate files if these contain the same number of rows and columns as your template. Therefore, if your template differs from what you are expecting, please [contact us](#).

Please refer to 'Guidance on Model Inputs' for more information about the scenarios.

Once you have completed one output file for each scenario, you should upload each file to Montagu, using the '**Upload central burden estimates**' buttons on the Responsibilities page.

There is no specific filename format to use. This is because when you upload through Montagu, the URL of the page you are on will determine the scenario.

Montagu will confirm whether each central burden estimate file uploads successfully and show you some quick diagnostic graphs.

Registering how your central estimates have been calculated

When uploading your central estimates, Montagu will ask you to register how these have been calculated.

To contact us, please email montagu-help@imperial.ac.uk.

We will ultimately derive your point estimates by taking the mean of your stochastic estimates, rather than by using the central estimates uploaded to Montagu. (We realise many groups produce their central estimates in this way in any case.) However, it is still important that the central estimates that you upload to Montagu by the central estimate due date are meaningful – both for quick error-checking, and to allow us to share estimates with stakeholders quickly.

If possible, please generate your central estimates as the average of your stochastic estimates, as this will ensure consistency with the final estimates.

If this is not possible to do by the central estimate due date, please specify in your answer to the registration questions how your central estimates have been calculated.

If your answers to the registration questions change between uploading your central estimates and your stochastic estimates, you should complete this registration step again and re-upload your central estimates to Montagu.

Stochastic estimates

By ‘stochastic estimates’, we mean individual outputs from your model that capture some of the variation of the modelling process. This variation could be due to uncertainty in your input parameters or, if you have a stochastic model specifically, this variation could be due to individual model runs.

The aim is to capture some of the uncertainty and potential reasonable variation that is inherent in modelling transmission of the disease.

We do not necessarily aim to capture all the potential uncertainty that could occur; for example, so far we only include median projections of population size and do not account for the unknowns and variability in vaccination coverage.

In essence, we wish to answer the question “Given these model inputs of coverage and demography, what is the variation in burden we may expect due to uncertainty in disease transmission and occurrence?”.

Requirements

We require 200 model runs for each scenario, each of which represents a random sample from the uncertainty distribution of your model outputs. Optimally, this would be samples of a posterior distribution representing all the parameter uncertainty in your model. For models not using formal Bayesian inference methods, you should sample from a realistic joint uncertainty distribution of your parameters. If the joint uncertainty distribution is unknown, Latin hypercube sampling over

To contact us, please email montagu-help@imperial.ac.uk.

realistically broad ranges of all parameters is preferred. As we want to compare the runs across scenarios to calculate the impact, the same parameter samples must be used across all scenarios, and the runs labelled to ensure we can identify them.

However, for the 2023 full model runs, you must not vary demography and you should use only the demographic inputs we provide. Similarly, coverage of the vaccine being modelled should not be varied within your stochastic runs, as this is already specified for each scenario. For the same run id (mentioned later) for different scenarios, please set the same random seed.

Parameters

The modellers are subject experts and VIMC will defer to individual groups' expertise respecting the specific parameters/ characteristics within their models. However, this updated guidance aims to give some clarification.

Any parameter that is varied as part of your stochastic runs should be uploaded as part of your parameter set, with 200 parameter values (to match the number of model runs). You should give your parameters informative names (e.g. 'vaccine_efficacy', and provide a brief explainer of what each parameter is (or suite of parameters if they vary by country or age) in your cheatsheet.

What to vary

As an example, the IC-Garske model is estimated in a Bayesian framework and there is a posterior predicted distribution for each country's force of infection (FOI) and the vaccine efficacy. Infections are calculated using 200 samples of these FOI and vaccine efficacy. These are then scaled by the probability of being a severe infection (case) and the probability of a case dying (death). Both of these probabilities are sampled from distributions taken from a published analysis.

In general, if a model is estimated within a Bayesian framework or similar, the stochastic parameters should be sampled from the resulting posterior to produce the estimates of burden.

Where a model is stochastic, each run should represent an individual realisation (see note on random seeds).

For models that are not fitted but take input parameters from the literature, plausible parameter distributions should also be taken from the literature and sampled accordingly.

To contact us, please email montagu-help@imperial.ac.uk.

Parameters varied should pertain to the disease transmission and occurrence of interest. Vaccine efficacy is a good example of a specific parameter that relates to disease transmission and occurrence.

As noted in the 'Requirements' section, you should not vary demography or coverage within your parameters.

Generating and uploading stochastic estimates

First, download your stochastic burden estimate template from the Responsibilities page (listed under 'Scenarios').

The format is almost identical to the central burden estimate template, but there is one additional column 'run_id'. This column labels the particular run, and should link the run to the parameter value detailed in the parameter set file. Importantly, the runs across all scenarios with the same run id should be based on the same parameter values.

We require 200 independent realisations in the stochastic estimates. The template file only contains all rows for a single realisation, so you will need to generate 200 times as many rows.

You will need to use the stochastic burden estimate template to create one or more files for each scenario, ensuring that you fill in all required rows and columns. If you choose to break the data up into multiple files it does not matter how you distribute the rows among files (e.g. by country, by run_id, by year or even randomly), as long as the data are complete, and scenarios are kept separate. The scenarios are the same as for your central estimates.

Next, rename your stochastic estimate files. The filename format should be, for example, *stochastic_burden_est_YF-IC-Garske_yf-routine-default_1.csv*. The first part is from the template filename, the second part is the scenario ID (as it appears in Montagu), the final number is an arbitrary way to distinguish between different files for the same scenario if you choose to split the estimates across several files.

Once you have completed all files for each scenario, you should upload each one to Dropbox, to the specific folder that we have emailed you. We will then use scripts to automatically process the uploaded files and import them into Montagu.

If your model changes between uploading your central estimates and your stochastic estimates, please [contact us](#).

Parameter set

To contact us, please email montagu-help@imperial.ac.uk.

First, download your stochastic parameters template from the Responsibilities page. You should use this template to create one file (a parameter set) that will show us the underlying parameter values of your stochastic runs.

It is essential that the runs across all scenarios with the same 'run id' are based on the same parameter values.

Your parameters file should contain 200 rows (i.e. in addition to the row showing the column headings).

The column headings in the template are labelled <param_1> and <param_2> but you should rename these to the actual parameters you are using, and add extra columns if necessary.

If the model uses country-specific parameters, each country-specific parameter should be given in a separate column, using the naming convention <parameter_name>:<ISO>, where ISO is the 3-letter country code.

Once you have completed your parameter set, you should upload this file via Montagu.

Montagu will then give you a 'parameter certificate'. After you have downloaded this, please upload it to Dropbox, to the specific folder that we have emailed you.

You should only upload one parameter certificate to Dropbox. This must correspond to the exact parameters that underlie your stochastic estimates. Therefore, if you discover a mistake in your stochastic files or parameter set after you have uploaded these to Dropbox, please do all of the following:

- a) [Contact us](#) to request a new Dropbox link.
- b) Upload your amended parameter set to Montagu and get a new parameter certificate
- c) Upload your new parameter certificate and all stochastic estimate files to the new Dropbox link

To contact us, please email montagu-help@imperial.ac.uk.