

Instructions for users of the STI-HIV model

This note has been prepared for individuals who wish to verify the calculations performed using the STI-HIV model, as presented in the paper “The effect of changes in condom usage and antiretroviral treatment coverage on HIV incidence in South Africa: a model-based analysis” (Johnson, Hallett, Rehle and Dorrington). The note is not intended to provide a comprehensive guide on all aspects of the STI-HIV model – it only concerns those aspects of the model that are relevant to the paper. Queries regarding the model should be directed to Leigh Johnson (Leigh.Johnson at uct.ac.za). The model is very complex and is not user-friendly, and anyone wishing to publish estimates based on the model is therefore encouraged to contact the author to ensure that their interpretation of the model parameters and outputs is correct.

Platforms supported

The code has been tested using the following compilers:

- Microsoft Visual C++ 2010 Express on Microsoft Windows
- MinGW's port of g++ on Microsoft Windows
- g++ on Linux

The program is written in standard ANSI/ISO C++ (C++03) and any C++ compiler should compile it. It is a console application without any operating system specific code. The compiler will generate several warnings. These can be ignored.

Installation

Go to: <https://github.com/leighjohnson/STI-HIV-Model>

If you don't have git, click the ZIP icon and download a zipped file of the contents. Unzip it in a folder on your computer.

If you do have git, you can simply fork the github directory.

Microsoft Visual C++ 2010 Express

Select File | New | Project From Existing Code ...

This opens a Wizard.

On the Welcome window select Visual C++. Click Next.

On the next window, Specify Project Location and Source Files, select the folder where the source has been unzipped and give the project a name. Click Next.

On the final window, Specify Project Settings, make sure to select *Console application project*. If you don't do this, it won't compile properly. Leave everything else unchecked. Click Finish.

To compile the program, press F7 or select Build | Build Solution.

g++ under Linux and g++ under Windows using MinGW

At a command prompt in the folder containing the project, type:

```
make
```

Running the program

In Visual C++, press CTRL-F5 or F5.

If you used MinGW, the executable program is called *stihiv.exe*.

In Linux the executable is called *stihiv*.

Technical note for users of the HIV-STI model

Note: This technical note does not display equations in github. You have to read the generated file **README.pdf** to see the math display properly. Click on the **README.pdf** file in github and then click ***view raw***. It should then be displayed properly by your pdf reading software.

Running the HIV-STI model generates several output files. The output files should be saved in the same folder as the folder in which the other project files are saved. For the purpose of the published paper, the key output files are:

- HIVinc15to49.txt: HIV incidence rates in the population aged 15 to 49
- HIVincByAgeF.txt: HIV incidence rates in women at the start of 2008, by 5-year age group
- HIVincByAgeM.txt: HIV incidence rates in men at the start of 2008, by 5-year age group
- CoVHIVinc.txt: Coefficient of variation in HIV incidence rates in the population aged 15 to 49
- ANCprevTot.txt: HIV prevalence estimates in pregnant women, before adjusting for antenatal bias
- ANCprev15.txt, ANCprev20.txt, ANCprev25.txt, ANCprev30.txt, and ANCprev35.txt: HIV prevalence estimates in pregnant women aged 15-19, 20-24, 25-29, 30-34 and 35-39 respectively (before adjusting for antenatal bias)
- ANCbias.txt: the estimated extent of antenatal bias in each scenario
- HSRCprevF.txt and HSRCprevM.txt: HIV prevalence in 2005, by 5-year age group, in females and males respectively (for the purpose of comparison with the 2005 HSRC prevalence survey)
- HSRC2008F.txt and HSRC2008M.txt: HIV prevalence in 2008, by 5-year age group, in females and males respectively (for the purpose of comparison with the 2008 HSRC prevalence survey)
- RandomUniformSex.txt: the sexual behaviour parameters in the 1000 simulations (expressed as percentiles of the corresponding prior distributions)

To verify that you've compiled and run the program successfully, you can compare (e.g. using a tool like diff) the contents of the generated file HIVinc15to49.txt to HIVinc15to49.cmp. These should be identical. (Note that Linux and Windows carriage returns differently and that if you use diff under Linux, you need to account for this.)

In each of these output files, each row corresponds to a different simulation (1000 randomly generated simulations in total). For the HIVinc15to49.txt, CoVHIVinc.txt and all ANCprev text files, each column corresponds to a different projection year (the third column corresponds to 1985, the first projection year). For the HIVincByAge, HSRCprev and HSRC2008 text files, each column corresponds to a different 5-year age group (the third column corresponds to 15-19 year olds and the last column corresponds to 55-59 year olds). For the RandomUniformSex.txt file, each column corresponds to a different sexual behaviour parameter that is allowed to vary in the uncertainty analysis (the last column is 1 – the condom bias parameter).

These output files contain only the raw outputs. To generate the summary outputs we recommend that users copy the output files into spreadsheets and use spreadsheet formulas to calculate the summary outputs. For example, in Microsoft Excel, the formula '=AVERAGE(cell range)' can be used to calculate the mean of the posterior sample, and the formulas '=PERCENTILE(cell range,0.025)' and '=PERCENTILE(cell range,0.975)' can be used to calculate the lower and upper limits of the 95% confidence intervals respectively. It may also be necessary to use Excel to calculate some of the outputs that are functions of the existing variables. For example, the probability that an HIV-negative 15-year old becomes infected before age 60 is approximated using the formula

$$1 - \prod_{j=3}^{11} [1 - I_g(j \times 5)]^5$$

where $I_a(x)$ is the average HIV incidence in individuals of sex g , aged x to $x + 4$ (as recorded in the HIVincByAgeF.txt and HIVincByAgeM.txt files). To generate this output, it would be necessary to apply the above formula to the HIV incidence rates in each row of the output file, before applying the AVERAGE and PERCENTILE functions to this calculated lifetime risk of infection.

To get the adjusted model estimate of antenatal HIV prevalence, after controlling for antenatal bias, it is necessary to combine the ANCprev output and the ANCbias output, taking into account the model assumption that the adjusted antenatal prevalence, p_a , and the unadjusted antenatal prevalence, p_u , differ by a constant b on the *logit* scale, i.e. $\text{logit}(p_a) = \text{logit}(p_u) + b$. Thus the formula for calculating p_a is

$$\left[1 + \left(\frac{1-p_u}{p_u}\right)e^{-b}\right]^{-1}$$

Note that the antenatal bias output in the ANCbias.txt file is already in exponentiated form, so that the formula one would enter into Excel to calculate the adjusted antenatal prevalence would be of the form $'=1/(1+(1/ANCprev-1)/ANCbias)'$.

Some points to note on the timing of output calculations:

- All HIV prevalence calculations are performed at the middle of the corresponding calendar year.
- The HIV incidence rates are calculated from mid-year to mid-year. Thus the incidence rates for 1985 (in the third column of the HIVinc15to49.txt file) correspond to the period from mid-1985 to mid-1986. Where we refer to 'HIV incidence rates at the start of 2008', we are actually referring to the HIV incidence rate over the period from mid-2007 to mid-2008, on the assumption that any change in HIV incidence over this interval would be approximately linear.
- The HIV incidence rates by 5-year age group are calculated for the period from mid-2007 to mid-2008 (i.e. approximately at the start of 2008).
- The coefficient of variation in HIV incidence rates is calculated as at the middle of 2008.

To generate model results for the counterfactual scenario in which there is no increase in condom use over time, edit out line 6753 in the 'TSHISAv1.cpp' file and remove the two forward slashes in the line above, so that the UpdateCondomUse function is called only at the start of each simulation, and not at the start of each year.

To generate the model results for the counterfactual scenario in which there is no ART, edit in line 6747 in the 'TSHISAv1.cpp' file ("HAARTaccess[CurrYear-StartYear] = 0.0;").

Credits

Leigh Johnson is the author of the STI HIV model coded in TSHISAv1.cpp and TSHISAv1.h.

This 3rd-party code is used by this project:

Agner Fog is the author of the mersenne twister code.

Barry Brown, James Lovato, and Kathy Russell are the authors of the statistics functions in StatFunctions.cpp and StatFunctions.h