Step-by-step instructions on how to install R (free software) and run the 5-HTTLPR meta-analysis script at your site

Step 1. Obtain and install R

R is a free software package available at the link below. Please follow these steps to install it on your computer.

The 5-HTTLPR meta-analysis script must be run using R version 2.14 or 2.15. (In particular, the script may not run correctly for R version 3.x.)

1.1. Download R. Version 2.14 for your operating system can be found at the following links:

A. For Mac users: http://cran.us.r-project.org/bin/macosx/old/R-2.14.1.pkg

B. For Microsoft Windows users: http://cran.us.r-project.org/bin/windows/base/old/2.14.1/

C. Due to the large number of possible linux OS specifications, the links will not be enumerated, but will be under http://cran.us.r-project.org/bin/linux/.

1.2. Install R. Once the download finishes, install the software using the default settings. (For example, to begin the process on a Mac, double click on the R-2.14.1.pkg file and follow the instructions in the pop-up window.) Be sure to note the location where R is installed because to run the 5-HTTLPR analysis script, you will need to invoke the program from the command line of a terminal window.

If you have any difficulty installing the R software, please contact the Coordinating Group (achorton@wustl.edu)

Step 2. Format your data into a .csv file

For your data to be read by the analysis script, it must be formatted using commas to separate the variables and the file name must end with the extension .csv. (e.g. *your_data.csv*). Instructions for formatting, naming, and defining the variables are contained in the companion document:

Formatting_data_for_5-HTTLPR_meta-analysis.doc

Step 3. Download and unpack the script

3.1. Download the script

You will shortly receive an email from achorton@wustl.edu regarding a package waiting for you at https://dropbox.wustl.edu/ Follow the instructions in the email to download the script directory tarball, SD_5HTTLPR_v8.1.tgz, into your desired directory for storing the script, datafiles, and results files. Make sure you place the .tgz file in a location with at least 300Mb space before proceeding.

3.2. Uncompress the script directory tarball

This can be done in a terminal window on a Mac or Unix/LINUX system by changing directories to the directory containing the .tgz file and then typing:

<div align="center">

tar -xpvzf SD_5HTTLPR_v8.1.tgz

</div>

or by using a utility such as 7-zip (free for Linux, or Windows), Stuffit Expander or Keka (free for Mac), or WinZip (Windows).

Expanding the tarball will create several files and a subdirectory, **subscripts_v8.1/**. Please check that you have all the files listed in the new file, **list_of_included_files.txt**

Step 4. Run the 5-HTTLPR script on the included test data

In order to confirm that R is properly installed and that the script will run accurately on your OS, please run the script using the input file TEST_DATA.csv BEFORE running it on your data. This section lists the required steps for running the script.

4.1. Edit the wrapper script, SD_5HTTLPR_script_v8.1.r.

One of the files in the main directory will be **SD_5HTTLPR_script_v8.1.r.** This file contains 22 questions about the data over 2 sections. For running on the test data, only edit the first answer (1.1 - the location of the unpacked script)**.** Save the changes to the file.

4.2. Open a terminal window and cd to the directory containing this file if you have not already done so.

For Windows 7: go to Windows>All Programs>Accessories>Command Prompt
For VISTA (or earlier): go to Start>Run. In the new window, type cmd <enter>.

For Mac OSX: click on the /Applications/Utilities/Terminal.app

4.3. At the command prompt type the following:

/path/to/R <space> **CMD** <space> **BATCH** <space> **/path/to/script** <space> **/path/to/log** <enter>

where

/path/to/R is the full pathname of the R program,

/path/to/script is the full pathname of SD_5HTTLPR_script_v8.1.r,

/path/to/log is the full pathname of the logfile
    (should be the same at /path/to/script with the '.r' replaced by '.log'

<space> means typing a blank space and <enter>* is the enter or return key

Examples:

/path/to/R might be

    (for Windows) C:\Program Files\R\R-2.15.0\bin\R.exe
    or
    (for Mac or Linux) /usr/bin/R

/path/to/script might be

    C:\Documents and Settings\userid\SD_5HTTLPR_v8.1\SD_5HTTLPR_script_v8.1.r
    or
    /Users/userid/Projects/SD5_HTTLPR_v8.1/SD_5HTTLPR_script_v8.1.r

/path/to/log should be the same as /path/to/script with the .r replaced by .log.

Running the script on the test data should require between 5 and 60 minutes. You can check progress via messages printed to the logfile.

NOTE: The script depends on an additional R package, *rms*, which it will attempt to download automatically at the beginning of its run. Please **check the logfile within 5 minutes** of starting the script. If there is an error, you may need to manually install this package. For help with manually installing *rms*, please contact achorton@wustl.edu.

Step 5. Check the execution of the script:

5.1. Examine the end of the logfile: The command in step 4.3 ran the analysis program. The prompt will return in the command window when the script has finished running. If it has completed successfully, within the final 10 lines of the log file (SD_5HTTLPR_script_v8.1.log) you should find the statement:

        === Congratulations. The script has run to completion.

If you do not see this message, please contact the coordinating team with your logfile.

5.2. Compare the list of files created in the Results_TEST_EU/ folder with the file **list_of_files_generated_by TEST.txt**. If some of these files are missing, please contact the coordinating team with your logfile and a list of files present in the Results_TEST_EU/ folder.

5.3. Once you have a successful test run, delete the logfile

Step 6. Run the 5-HTTLPR script on your data

If you have multiple datasets for analysis, **repeat Steps 6 through 9 for each dataset.**

6.1. Edit SD_5HTTLPR_script_v8.1.r for YOUR DATA

Answer the 22 questions at the top of the script (Note: your answer to 1.1 is unlikely to change between runs.)

6.2. Run the script on your dataset. At the command prompt type the following:

/path/to/R <space> **CMD** <space> **BATCH** <space> **/path/to/script** <space> **/path/to/log** <enter> (Consult section 4.3 for details about paths.)

Results will be generated in the Results_{STUDY}_{POPULATION}/ folder. You can check progress via messages printed to the logfile.

Step 7. Check the execution of the script on your data:

7.1. Examine the end of the logfile: The command in step 6.2 ran the analysis program. The prompt will return in the command window when the script has finished running. If it has completed successfully, within the final 10 lines of the log file (SD_5HTTLPR_script_v8.1.log) you should find the statement:

=== Congratulations. The script has run to completion.

If you do not see this message, please contact the coordinating team with your logfile.

Step 8. Compress your results.

Once the script finishes running, move your logfile and copy your edited SD_5HTTLPR_script_v8.1.r into the Results_{STUDY}_{POPULATION}/ directory.

A. Windows users: Use 7-zip, or the software you used to unpack the script package to make a compressed file (e.g. .zip). 7-zip also allows you to encrypt the file. If you choose to encrypt, please send the key in a separate email to achorton@wustl.edu.

B. Mac or Linux users:

  B.1. Tarball the directory by typing (from the script directory):

      tar -cpvzf {STUDY}_{POPULATION}.tgz Results_{STUDY}_{POPULATION}/

  where {STUDY} is your study acronym and {POPULATION} is the two-letter population code detailed in SD_5HTTLPR_script_v8.1.r section 1 questions 1.2 and 1.3.

  B.2. You may wish to encrypt your results file with 7-zip (Linux) or Keka (free software for the Mac) or possibly the software you used to unpack the script package. If you choose to encrypt, please send the key in a separate email to achorton@wustl.edu.

Step 9. Return results to the coordinating team:

9.1. Go to https://dropbox.wustl.edu//dropoff.php

9.2. Upload the compressed, encrypted Results_{STUDY}_{POPULATION} file by following the instructions on the Dropbox page.

    1. Enter your name and email address.
    2. For Recipient 1, enter
        Name: Amy Horton
        Email: achorton@wustl.edu.
    3. Browse to identify your file for upload.

We will retrieve your results promptly. They will be automatically deleted from the Washington University Dropbox system 7 days after you upload them.

If you encrypt, please send a separate email to achorton@wustl.edu with the key.

Thank you for participating in the 5-HTTLPR, Stress, and Depression meta-analysis project!