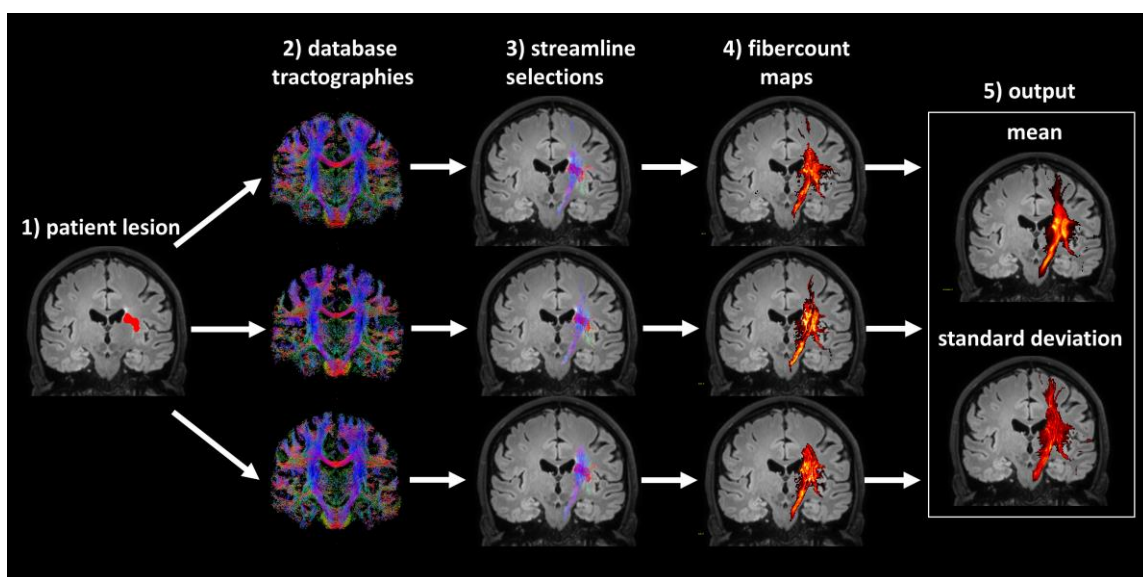


simDTI Manual

1: Introduction:	3
2: Installing the toolbox:	4
3. Getting started	5
4. Preparing the necessary files	6
The lesion mask:	6
The database:	6
5. Running the toolbox	7

1: Introduction:

SimDTI is a toolbox for predicting white matter damage based on lesion segmentation in MNI space. It works by superimposing a lesion volume on the tractographies of many subjects and selecting streamlines that pass the lesion. The selected streamlines are mapped to a fiber-count volume and subsequently various statistics can be derived from that, including most importantly the mean and standard deviation. A schematic of the pipeline of the toolbox can be seen here below:



Schematic of the procedure for creating predictions of damage to the white matter structure. A lesion is manually segmented and normalized to MNI space (1), and subsequently imposed on normalized tractographies of control subjects in a database (2). For each subject in the database all streamlines that touched the lesion area were selected (3) and mapped to a volume representing the fiber count for each voxel (4). The mean and standard deviation of the fiber count maps was calculated across subjects in the database (5).

In addition, the damaged streamlines are mapped to a connectome. The connectome is stored in an ascii table and shows which Regions of Interest (ROIs) the damaged streamlines were connecting.

2: Installing the toolbox:

The toolbox runs under Linux requires several analyses packages to be installed on your system. The toolbox is written in IDL, and calls functions of FSL and MRTrix, so all these three packages need to be installed/initialized on your system for the toolbox to run.

The toolbox is compiled IDL code, so that it can run with the IDL virtual machine without the need for an IDL license. You do, however, need to have IDL 8.3 or higher (which includes the IDL virtual machine) installed on your system to be able to run it. You can download IDL for free from:

<https://www.harrisgeospatial.com/>

Remember that you need to register, and your registration must be verified before you can download, which might take up to 24 hours. Instructions on how to install IDL you can find on:

<https://www.harrisgeospatial.com/Support/SelfHelpTools/HelpArticles/HelpArticles-Detail/TabId/2718/ArtMID/10220/ArticleID/15066/Install-and-License-IDL-86.aspx>

You can disregard the portion covering the license server, as you will be only using the virtual machine.

Detailed instructions on installing FSL can be found on:

<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FslInstallation>

Instructions for installing MRTrix can be found on:

https://mrtrix.readthedocs.io/en/latest/installation/linux_install.html

3. Getting started

After installing FSL, MRTrix, and IDL you can put the simDTI folder somewhere on your system. After that, you can add a single line to our .bashrc file in your home folder using a text editor. The example below is for when the toolbox is installed as /home/user/simdti/simdti.sav

```
alias simdti='idl -vm=/home/user/simdti/simdti.sav'
```

After this you can start the toolbox by opening a console and typing '*simdti*' in the console.

4. Preparing the necessary files

The lesion mask:

The main input is a lesion mask in MNI space and in nifti-format, with ones indicating lesion voxels, and zeros non-lesion voxels. The most straightforward approach is to create these manually using e.g. MRlcron (<https://www.nitrc.org/projects/mricron>) or itk-SNAP (<http://www.itksnap.org/pmwiki/pmwiki.php>). Automatic segmentation is of course more convenient, but often requires multiple input images and human supervision. Also, take into account that especially large lesions can influence the normalization algorithm, requiring a dedicated toolbox (e.g. the Clinical Toolbox for SPM; <https://www.nitrc.org/projects/clinicaltbx/>) to transform the lesion segmentation to MNI-space.

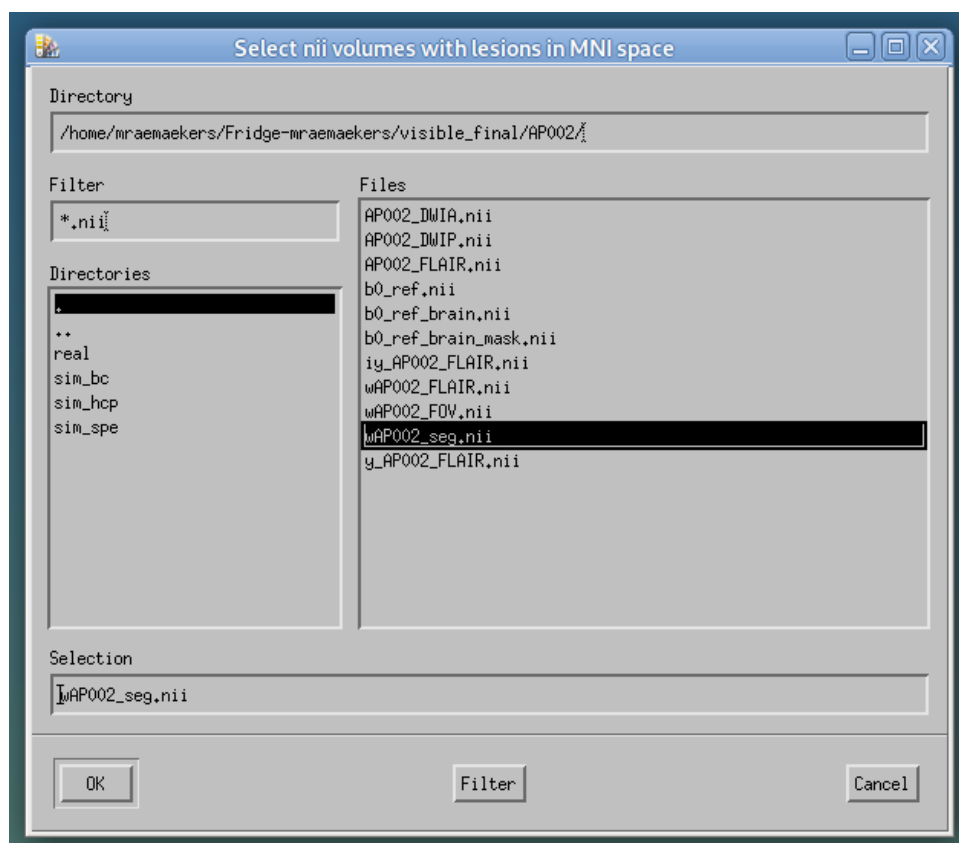
The database:

The database consists of a set of whole-brain tractography-files in MNI space and in .tck-format (compliant with MRtrix; <https://www.mrtrix.org>) and for each tractography an ROI-segmentation labeling only the grey-matter voxels. The creation of a database of considerable size is a computationally intensive process and requires some expert knowledge. If you don't want to go through the hassle of creating a database yourself, please send me an email (m.raemaekers-2@umcutrecht.nl), and I will open a download link for you. Note the size of a database can easily exceed 100 GB, which makes quite a download.

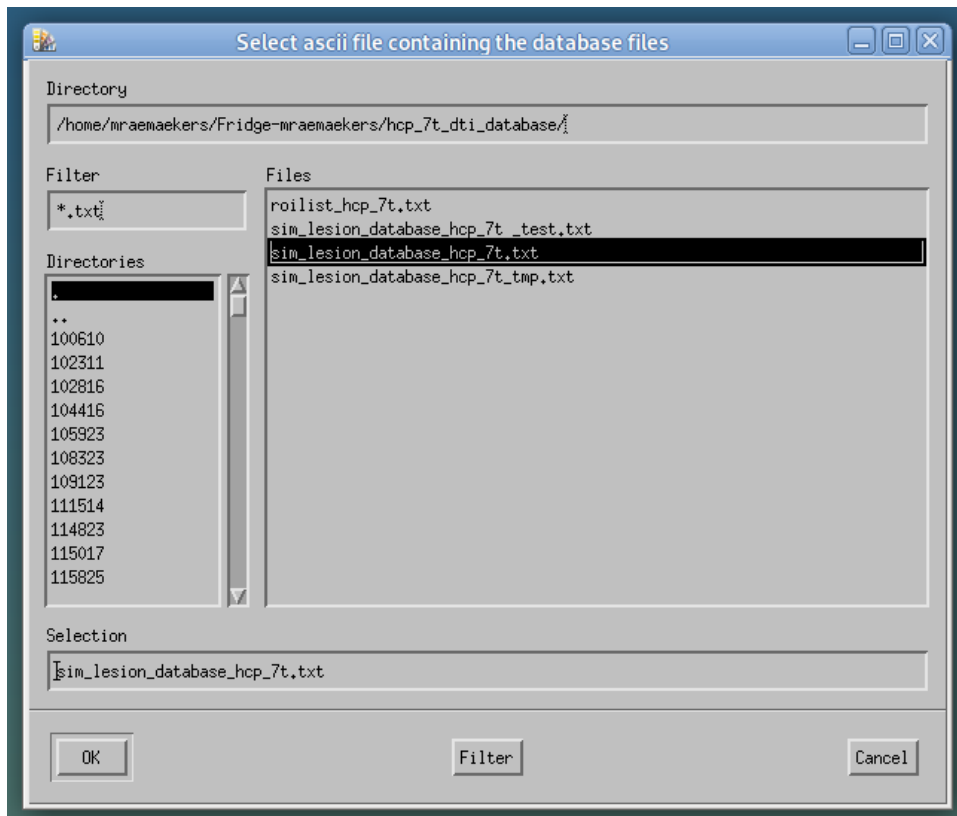
The database can be addressed using a .txt file including a three-column ASCII table, where the first column specifies the subject code, the second column the tractography file, and the third column the nifti-file. An example of a database file has been added to the github of simDTI.

5. Running the toolbox

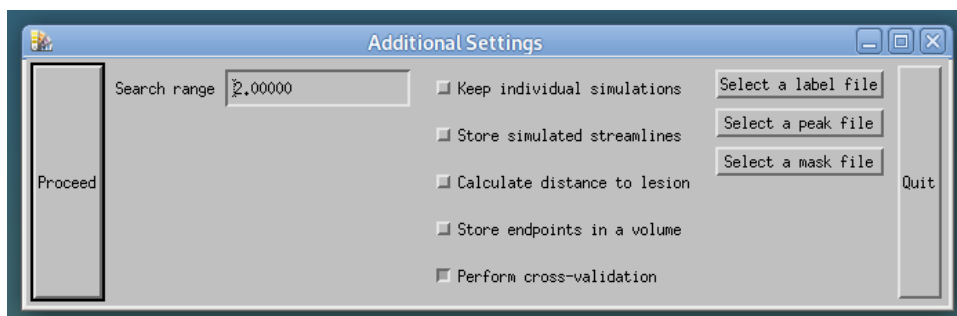
To start making predictions of white matter damage, type 'simdti' in the console. You first get an image that you are running the IDL-virtual machine. You can click on this image to continue. After that, you get a file selection menu, where you can choose one or multiple nifti-files containing the lesion mask. For selecting multiple nifti-files simultaneously, they have to be stored in the same folder. You can select multiple files by holding control or shift when pressing the left mouse button. Also take note that the output volumes will have the same dimensions as the input volumes.



After selecting the input-volume(s), press 'OK', and you will then be required to select the database-file (see section 4)



After selecting the database-file, you get to the menu shown here below:



Here you can set a few additional options for the analysis including:

The **'Search range'** is relevant for the process of creating the connectomes and specifies how far to search for a grey-matter ROI from the start or endpoints of a streamline. A slight search is necessary as most streamlines will terminate before crossing the grey-matter boundary. 2 mm is the default.

'Keep individual simulations' indicates to store the fiber-count maps of all subjects in the database when making predictions. Note that this creates a huge amount of data, so make sure you have plenty of disk space available.

'Calculate distance to lesion' produces an additional volume containing for each voxel with a non-zero prediction the mean distance towards the lesion along the streamlines.

The option **'Store endpoints'** creates an additional volume in which only the counts of the endpoints of the streamlines predicted to be damaged are stored.

'Perform cross-validation performs' a leave-one-out cross-validation on the simulation result. This creates a reliability estimate of the prediction which is stored in a separate text-file. This reliability

estimate calculated both for the connectomes as well as for the volumes. Especially with large databases, this setting can prolong the calculations quite a bit.

'Select a label file' is an optional procedure for including the names corresponding to the ROI as used in the database in the connectome results. These names will then be displayed in the connectome result instead of the numbers. The label file is a two-column ASCII/txt-file, with the first column indicating the ROI-number, and the second column the names of the ROIs

'Select peak file' is an optional parameter to estimate the directional component of the predictions. You do this by selecting a 9-volume .mif file including the vector components of FODs that you want to check the predicted streamlines against. The vector components specify a maximum of 3 directions for each voxel (multiple directions in areas with crossing fibers). If specified, the toolbox will also produce a mean amplitude estimate for each of the specified directions. This setting is not yet fully validated, so use it at your own risk.

'Select a mask file' allows you to select a volumetric mask (in nifti-format) that restricts the number of voxels used in calculating reliability estimates. This option will make reliability estimates quicker and more sensible.

After you are satisfied with the selected options you can press **'Proceed'**, which will start the making of the predictions. Depending on the number of subjects in the database and the quality of your hardware this may take up to several hours for each subject.