

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/308495187>

# NeuroDebian Virtual Machine Deployment Facilitates Trainee-Driven Bedside Neuroimaging Research

Article in *Journal of Child Neurology* · September 2016

DOI: 10.1177/0883073816668113

CITATIONS

0

READS

59

4 authors, including:



Alexander Li Cohen

Boston Children's Hospital

53 PUBLICATIONS 12,614 CITATIONS

SEE PROFILE




Hugo Botha

Mayo Foundation for Medical Education and Research

75 PUBLICATIONS 565 CITATIONS

SEE PROFILE

# NeuroDebian Virtual Machine Deployment Facilitates Trainee-Driven Bedside Neuroimaging Research

Journal of Child Neurology  
1-6  
© The Author(s) 2016  
Reprints and permission:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/0883073816668113  
jcn.sagepub.com  


Alexander Cohen, MD, PhD<sup>1</sup>, Daniel Kenney-Jung, MD<sup>1</sup>,  
Hugo Botha, MB, ChB<sup>2</sup>, and Jan-Mendelt Tillema, MD<sup>1</sup>

## Abstract

Freely available software, derived from the past 2 decades of neuroimaging research, is significantly more flexible for research purposes than presently available clinical tools. Here, we describe and demonstrate the utility of rapidly deployable analysis software to facilitate trainee-driven translational neuroimaging research. A recipe and video tutorial were created to guide the creation of a NeuroDebian-based virtual computer that conforms to current neuroimaging research standards and can exist within a HIPAA-compliant system. This allows for retrieval of clinical imaging data, conversion to standard file formats, and rapid visualization and quantification of individual patients' cortical and subcortical anatomy. As an example, we apply this pipeline to a pediatric patient's data to illustrate the advantages of research-derived neuroimaging tools in asking quantitative questions "at the bedside." Our goal is to provide a path of entry for trainees to become familiar with common neuroimaging tools and foster an increased interest in translational research.

## Keywords

education, cortical localization, epilepsy, quantitative MRI, translational research, virtual machine

Received June 6, 2016. Received revised August 4, 2016. Accepted for publication August 10, 2016.

The past 2 decades have seen tremendous advances in neuroimaging capabilities, with ever-increasing clinical utility. In parallel, systems neuroscience has seen an even larger expansion of novel techniques that, while commonplace in the research laboratory, are still relatively unknown in the clinical setting. Several publicly funded and freely available software packages<sup>1-7</sup> have been developed for neuroimaging research to perform sophisticated structural and functional analyses. Several of these software packages have broad utility with hundreds to thousands of publications each and active development efforts at their home institutions. Although these programs do not have FDA approval for routine clinical use, the capabilities available in these applications are often significantly more flexible than those typically available in day-to-day clinical practice and can be of significant value to the interested clinician/researcher.

Most child neurology trainees are not exposed to these research tools in their training curricula. As a result, many interesting, appropriate, and potentially easily implemented research questions are often deemed "too difficult" to answer and hence avoided, even given a higher burden in the pediatric population to minimize sedation and maximize the utility of each neuroimaging session. Conversely, given the ever-growing demands on the time and energy of clinical trainees

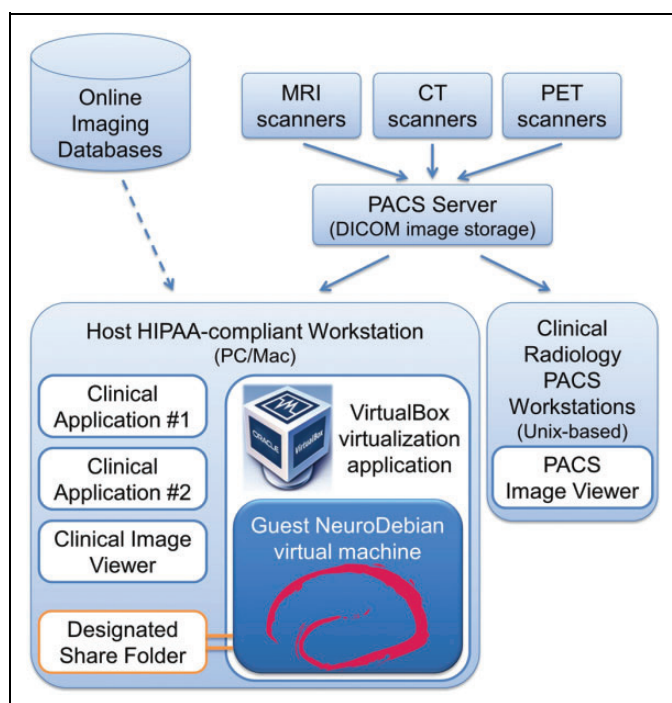
and faculty members, learning how to apply well-established automated analysis pipelines, such as the FreeSurfer segmentation and surface reconstruction tools,<sup>4,5</sup> can seem daunting. While constructing a basic neuroimaging curriculum to introduce contemporary software packages to Neurology trainees at Mayo Clinic, we quickly recognized that the largest and most immediate barriers to entry include (1) access to preconfigured neuroimaging laboratory computers, (2) a basic to moderate level of Unix/command line interface knowledge, and (3) a lack of either onsite resources in the form of research mentorship and technical support, time and means to travel to (often internationally) scheduled workshops, or lack of dedication, time commitment, and interest to teach oneself the relevant techniques. There are several robust online tutorials that could address this last difficulty; however, these are often aimed at

<sup>1</sup> Department of Child and Adolescent Neurology, Mayo Clinic, Rochester, MN, USA

<sup>2</sup> Department of Neurology, Mayo Clinic, Rochester, MN, USA

## Corresponding Author:

Alexander Cohen, MD, PhD, Department of Child and Adolescent Neurology, Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA.  
Email: alexander.cohen2@childrens.harvard.edu



**Figure 1.** Schematic of a NeuroDebian Virtual Machine within a HIPAA-compliant clinical network. Shown is a schematic representation of a NeuroDebian virtual machine hosted on a HIPAA-compliant clinical workstation. The guest virtual machine, that is, a “computer within a computer,” can run a distinct operating system and software suite and does not interfere with the host computer’s clinical applications. The guest NeuroDebian virtual machine (lower left panel) can be allowed to share the network connection of the host computer to interact with available resources, such as Radiology Picture Archive and Communication System (PACS) servers or online imaging databases. This occurs alongside routine clinical evaluation of images on Radiology PACS Workstations (lower right panel). Additionally, a folder/directory can be designated on the host computer (yellow box) that both the host and guest systems can access for controlled transfer of files to and from the virtual machine.

the experienced technical computer user. Here, we present our efforts to address the first and second of these challenges.

## Methods

Our goal was the creation of a standardized and rapidly deployable computer environment within the confines of a preexisting HIPAA-compliant clinical computer network, regardless of variable computer configurations (Figure 1). This computer environment should be as similar as possible to research/industry standards, allowing users to ask for troubleshooting advice from the robust online forums and mailing lists without having to explain or describe any custom local implementations. In addition, there cannot be any interference with the host computer’s clinical applications. Our solution was to create a recipe-based approach to combine and configure existing tools, in particular the free and open-source NeuroDebian Linux distribution<sup>8</sup> and Oracle’s VirtualBox, to create a standardized Virtual Machine, that is, a computer within a computer, tailored for clinical neuroimaging applications. Although a natively configured computer would be ideal, this is not a practical endeavor for trainees unfamiliar

with Unix-based systems. Virtualization software has also advanced to the point where there is only a minor drop in performance from performing analyses in this virtual environment.

We have included here our full “recipe” used to download, install, and configure the relevant components of a standard NeuroDebian environment with particular emphasis on the FMRIB Software Library (FSL), developed by the Oxford Centre for Functional Magnetic Resonance Imaging of the Brain,<sup>2</sup> FreeSurfer, developed by the Martinos Center for Biomedical Imaging at Massachusetts General Hospital,<sup>4,5</sup> and the Human Connectome Project’s Connectome Workbench, developed at Washington University in St Louis<sup>6</sup> (Table 1). Following the steps contained in this recipe, the reader can create a neuroimaging-tailored NeuroDebian virtual environment within 4 to 6 hours, with the largest portion of time allotted for downloading the various tools. If multiple host computers are available with identical hardware, a virtual machine created on one host can also be easily copied and/or accessed on the other machines as well, decreasing deployment time as well. We also generated a complete walk-through video tutorial that leads the reader through each step from start to finish (Video 1).

Clinical magnetic resonance imaging (MRI) protocols are typically tailored toward answering specific clinical questions with the fewest number of scans and often do not contain specialized sequences used in rigorous research projects, such as bias fieldmaps or high-resolution 3-dimensional T2 imaging. Fortunately, there are many clinical research questions that can still be addressed using existing data and modern tools for registration of multimodal data, surface reconstruction, and volumetric segmentation/labeling. To facilitate this approach, we also compiled a set of “connector” commands/scripts (Video 2) that allow for rapid conversion from clinical Digital Imaging and Communications in Medicine (DICOM) files to visualization and analysis of individual patient’s cortical surfaces and labeled subcortical segmentations using the recently created Human Connectome Project’s Minimal Preprocessing Pipelines<sup>9</sup> and Connectome Workbench,<sup>6</sup> the latter of which is lightweight and portable to Windows, Mac, and Linux environments. As with the NeuroDebian “recipe” above, we generated a walk-through video tutorial that demonstrates the use of the NeuroDebian environment and these scripts to analyze a single patient’s clinical data (Video 2).

## Results

A standardized computing environment can be easily created within a HIPAA-compliant system without compromising network security or requiring new hardware (Figure 1). Clinically acquired data can be rapidly processed, that is, within a single day, to ask quantitative questions “at the bedside” that can lead to productive research avenues or as exploratory clinical assessments alongside the radiologic visual interpretation of the source images. This system allows for secure retrieval of clinical imaging data from DICOM nodes with Gingko CADx (where permitted), conversion to Neuroimaging Informatics Technology Initiative (NIFTI) file formats with dcm2nii, multimodal data registration with FSL or Advanced Normalization Tools (ANTs),<sup>10</sup> surface reconstruction and volumetric segmentation with FreeSurfer, and data visualization and exploration using the Connectome Workbench. Other commonly used software packages, such as Statistical Parametric Mapping

**Table 1.** Recipe for Creating a NeuroDebian-Based Virtual Machine.<sup>a</sup>

1. Download and install the latest version (5.0.16) of VirtualBox:
  - a. Go to: <https://www.virtualbox.org/wiki/Downloads>
  - b. NOTE: This will require administrative rights for your host computer; however, this is the *only* time you will need it throughout this recipe.
2. Download and import the latest version (8.0.0) of the NeuroDebian Virtual Machine that is appropriate for your “host” computer:
  - a. Go to: <http://neuro.debian.net/index.html>
  - b. Step-by-step instructions: <http://neuro.debian.net/vm.html>
  - c. NOTE: Choose the 32-bit version for best compatibility
  - d. Designate a “Shared Folder”, eg, a folder called “host” on the desktop.
3. Register for a FreeSurfer license (free):
  - a. Go to: <https://surfer.nmr.mgh.harvard.edu/registration.html>
  - b. Open the subsequent email and copy the designated 4 lines of text; this is your FreeSurfer license information
  - c. Save these 4 lines of text to a plain text file called `license.txt` in the “shared” folder created above.
4. Start the NeuroDebian Virtual Machine and once it is running, perform the following steps WITHIN the new windowed environment:
  - a. Open the Synaptic Package Manager (Applications Menu→System→Synaptic Package Manager):
  - b. NOTE: Password is “neurodebian”
  - c. Click on “Mark All Upgrades”, and
  - d. Search for and install the following packages:
    1. `ginkgocadx`
    2. `dcmstk`
    3. `mricron`
    4. `fsl-complete`
    5. `ants`
    6. `connectome-workbench`
  - e. Click on “Apply”
5. Open a terminal window and run the following commands in order:
  1. `wget -c ftp://surfer.nmr.mgh.harvard.edu/pub/dist/freesurfer/5.3.0-HCP/freesurfer-Linux-centos4-stable-pub-v5.3.0-HCP.tar.gz`
  2. `cd /opt`
  3. `sudo tar -xvf ~/freesurfer-Linux-centos4-stable-pub-v5.3.0-HCP.tar.gz`
  4. `sudo git clone https://github.com/Washington-University/Pipelines.git`
  5. `sudo git clone https://github.com/alexlicohen/script-toolbox.git`
  6. `cp /opt/script-toolbox/env_vm/* ~`
  7. `echo 'source ~/alexrc' >> ~/.bashrc`
  8. `sudo cp ~/host/license.txt /opt/freesurfer/license.txt`
  9. `sudo chmod a+r /opt/freesurfer/license.txt`

<sup>a</sup>Prerequisites for following this recipe include a relatively modern host computer, ie, one with a x86-based CPU and running any common operating system, and at least 50 GB of free hard drive space that you can use. Access to patient or research subject DICOM data via Q/R access to a PACS Server or via CD/USB drive will also be ideal; however, large data sets are available online for research and education. Note: There is a video tutorial associated with this manuscript that demonstrates the implementation of this recipe.

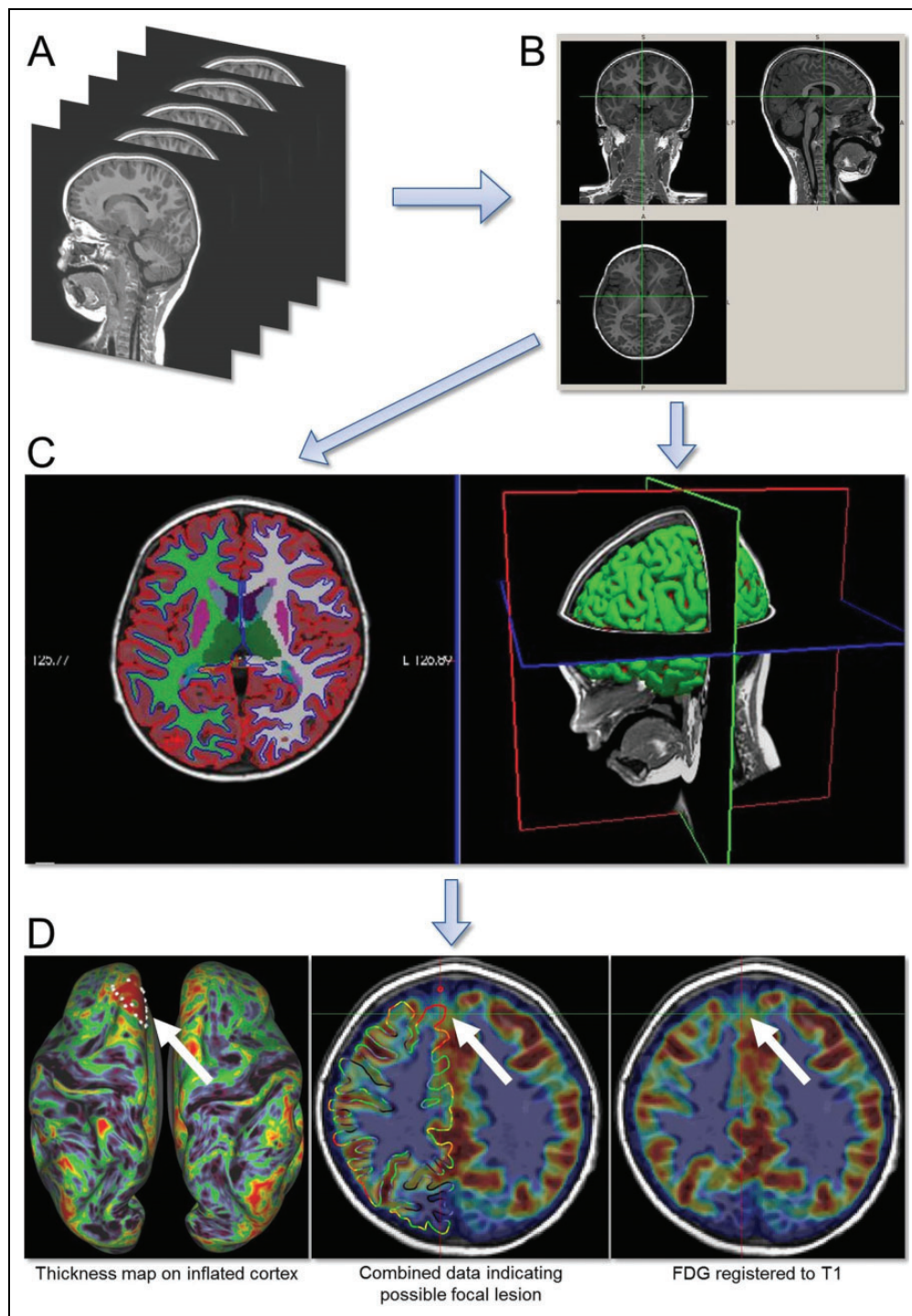
(SPM), AFNI, and Nipype are also compatible with this computing environment and can be installed easily.<sup>1,3,11</sup>

Here, we give an example of using modern neuroimaging research tools to explore clinical data (Figure 2). It should be noted, however, that while the demonstrated analysis stream can be very helpful to point out abnormalities, the results should continue to be considered investigational and when possible, compared against clinically validated tools. We present the clinical scenario of a child who has been having left frontal onset seizures, with multiple MRIs that are normal on visual inspection. A subsequent 18-F fluorodeoxyglucose–positron emission tomography (FDG-PET) scan revealed left frontal hypometabolism raising the possibility of candidacy for epilepsy surgery. In an exploratory fashion, the patient’s data were assessed for subtle cortical thickening, and if present, whether this overlapped with the subject’s FDG-PET findings. Shown are depictions of the data at multiple stages of analysis. Our analysis revealed a possible focal area of cortical

thickening within the left medial frontal lobe, measured at 4.65 mm, which correlated with the area of FDG hypometabolism. It is important to note, however, that the results of this demonstrated analysis stream are investigational and should always be compared against clinically validated tools and neuroradiology expertise.

## Discussion

Although there have been efforts to generate completely pre-configured Linux distributions for neuroimaging, for example, Lin4Neuro,<sup>12</sup> these still require ongoing maintenance to stay current and have a smaller support base than NeuroDebian. Our approach creates an environment as close as possible to a “native” neuroimaging laboratory system, building on existing efforts to standardize neuroimaging research. This approach also maintains compatibility with online tutorials and course materials allowing for transfer of learned



**Figure 2.** Case example of using modern research tools to explore clinical data. In this case, a child who has been having left frontal-onset seizures with multiple magnetic resonance images (MRIs) that are normal on visual inspection is being evaluated for possible epilepsy surgery candidacy. The patient has undergone an FDG-PET scan that revealed left frontal hypometabolism (D, right panel and backdrop of middle panel). In an exploratory fashion, the patient's data were also assessed for subtle changes in cortical thickness measures using the tools described here. The patient's clinical images (A) were transferred into a NeuroDebian virtual machine, where they could be reviewed using research tools such as FSL (B) or further processed into volumetric segmentations and surface reconstructions using Freesurfer (C). This process also produces multiple measures of cortical geometry, such as cortical thickness, which can be viewed as false color maps in the Connectome Workbench (D, left panel and ribbon in middle panel). Our analysis revealed possible focal cortical thickening within the left (shown in neuroimaging convention of Left=Left) medial frontal lobe, measured at 4.65 mm. Interestingly, this correlated with the area of FDG hypometabolism (D, white arrows), but as shown, there are similarly other regions of thickening in multiple areas throughout the cortex (red areas in left and middle panels of D). The results of the demonstrated analysis stream should be considered investigational and compared against clinically validated tools and neuroradiology expertise. NOTE: There is a video associated with this manuscript that demonstrates the analysis stream shown here.



**Table 2.** Online Resources and Support Infrastructure for the Described Tools.**Available Tutorials and Workshops:**

- FSL: <http://fsl.fmrib.ox.ac.uk/fslcourse/>
- FreeSurfer: <http://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial>
- Connectome Workbench: [http://www.humanconnectome.org/documentation/tutorials/Connectome\\_WB\\_Tutorial\\_v1.0.pdf](http://www.humanconnectome.org/documentation/tutorials/Connectome_WB_Tutorial_v1.0.pdf)
- NOTE: Several of these tutorials require downloading additional data sets, which can be quite large. Consider deleting these files when you are done with each tutorial.

**Support Forums and Mailing Lists:**

- VirtualBox: <https://forums.virtualbox.org/>
- NeuroDebian: <http://neuro.debian.net/about.html>
- FSL: <http://fsl.fmrib.ox.ac.uk/fslwiki/Support>
- FreeSurfer: <https://surfer.nmr.mgh.harvard.edu/fswiki/FreeSurferSupport>
- Connectome Workbench: <http://www.humanconnectome.org/contact/index.php#subscribe>

Video 1 (online only). Recorded process of creating and configuring a NeuroDebian virtual machine. This video portrays the entire process of creating and configuring a NeuroDebian virtual machine following the recipe contained in Table 1.

Video 2 (online only). Recorded example of using modern research tools to explore clinical data quantitatively. This video portrays the entire analysis stream described in the text and also shown in Figure 2 to assess for possible cortical dysplasia in a patient with magnetic resonance imaging (MRI)-negative focal-onset seizures.

skills for the computationally naive audience. Alternative solutions, such as the LONI Pipeline<sup>7</sup> from UCLA (<http://pipeline.loni.usc.edu/>) and the NITRC-Computing Environment (<http://www.nitrc.org/ce/>), were primarily designed to be run on cloud computing platforms and are thus not natively appropriate for use with HIPAA-protected data. The system we describe here is entirely self-contained and can perform analyses without transmitting data outside of the clinical network.

We hope that this manuscript helps address the first of the described major obstacles to learning how to perform clinical neuroimaging research. The second prerequisite, knowledge of the Unix/Linux filesystem and the command line interface, remains for most modern neuroimaging research tools. Fortunately, understanding the command line interface and the Linux filesystem are requirements across a broad spectrum of scientific and industrial endeavors. As such, there are many free high-quality tutorials available in print and online, many of which are interactive and confer a sufficient level of competence within 1 to 2 hours to start utilizing neuroimaging specific tutorials and tools (Table 2). The final challenge of learning how to use these tools to ask research questions in the clinical setting will always be unique to each reader's situation. As a starting point, we direct the reader to the software-specific online tutorials and communities that are available, especially in regard to interpretation of the eventual results. Variation in these continually evolving software packages have been described,<sup>13</sup> and careful

interpretation of one's own data is important, but falls outside of the scope of this article. Similarly, changes affecting the appearance of the brain across development, and in disease states, can alter the performance of image-processing algorithms and requires additional consideration. We anticipate, however, that application of these software tools to pertinent clinical questions will breed more sustainable learning than preplanned tutorials and we encourage the reader to explore and experiment with available data.

**Acknowledgments**

This work was performed at the Mayo Clinic in Rochester, MN, USA. This work was presented at the 68th American Academy of Neurology Meeting in Vancouver, British Columbia, Canada on April 17, 2016.

**Author Contributions**

AC, DKJ, HB, and JMT substantially contributed to study conception and design. AC and DKJ contributed to data acquisition, analysis, and interpretation of data. AC drafted the manuscript. DKJ, HB, and JMT critically revised the manuscript for important intellectual content.

**Declaration of Conflicting Interests**

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The authors received no financial support for the research, authorship, and/or publication of this article.

**Ethical Approval**

No ethics review board approval was needed for this study. Patient data utilized as example data were obtained with documented parental informed consent.

**Supplemental Material**

A Video Supplement for this article is available at <http://jcn.sagepub.com/supplemental>.

**References**

1. Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res.* 1996;29:162-173.
2. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. Fsl. *Neuroimage.* 2012;62:782-790.
3. Friston KJ. *Statistical Parametric Mapping: The Analysis of Functional Brain Images.* 1st ed. Amsterdam: Elsevier/Academic Press; 2007.
4. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage.* 1999;9:179-194.
5. Fischl B, Salat DH, Busa E, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron.* 2002;33:341-355.

6. Marcus D, Harwell J, Olsen T, et al. Informatics and data mining tools and strategies for the human connectome project. *Front Neuroinform.* 2011;5:4.
7. Dinov ID, Van Horn JD, Lozev KM, et al. Efficient, distributed and interactive neuroimaging data analysis using the LONI pipeline. *Front Neuroinform.* 2009;3:22.
8. Halchenko YO, Hanke M. Open is not enough. Let's take the next step: an integrated, community-driven computing platform for neuroscience. *Front Neuroinform.* 2012;6:22.
9. Glasser MF, Sotiropoulos SN, Wilson JA, et al. The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage.* 2013;80:105-124.
10. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage.* 2011;54:2033-2044.
11. Gorgolewski K, Burns CD, Madison C, et al. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework. *Front Neuroinform.* 2011;5:13.
12. Nemoto K, Dan I, Rorden C, et al. Lin4Neuro: a customized Linux distribution ready for neuroimaging analysis. *BMC Med Imaging.* 2011;11:3.
13. Popescu V, Schoonheim MM, Versteeg A, et al. Grey matter atrophy in multiple sclerosis: clinical interpretation depends on choice of analysis method. *PLoS One.* 2016;11:e0143942.