Validated Automatic Brain Extraction of Head CT Images

John Muschelli[c,*], Natalie Ullman[d], Daniel F. Hanley[d], Ciprian M. Crainiceanu[c]

[a] *Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA*
[b] *Department of Neurology, Division of Brain Injury Outcomes, Johns Hopkins Medical Institutions, Baltimore, MD, USA*

Abstract

1. Background

Computed Tomography (CT) brain imaging is commonly used in diagnostic settings. Although CT scans are primarily used in a clinical setting, they can be used to answer research hypotheses. Methods for brain extraction in head CT images has been informally proposed, but never formally validated.

2. Aim

To systematically analyze the performance of FSL's brain extraction tool (BET) on head CT images of patients with intracranial hemorrhage by varying its parameters and options after performing CT-specific preprocessing, and quantitatively comparing its results to a manual gold standard.

3. Methods

Nineteen images from 16 patients with intracranial hemorrhage were selected from 13 different MISTIE stroke trial centers. Each image was either not smoothed or smoothed using a 1mm Gaussian smoother and thresholded using a $0 - 100$ Hounsfield units (HU) range. BET was applied using 1 of 3 fractional intensity (FI) thresholds: 0.01, 0.1, 0.35 and holes in the brain mask produced by BET were filled. For validation purposes, intracranial masks were manually created for all image volumes by expert CT readers. The resulting brain tissue masks were quantitatively compared to the manual segmentations using sensitivity, specificity, accuracy, and the Dice Similarity Index (DSI). Brain extraction across smoothing and FI thresholds were compared using the Wilcoxon signed-rank test.

4. Results

Smoothing images improves brain extraction results using BET for all metrics and irrespective of the FI threshold. Using an FI of 0.01 or 0.1 performed better than 0.35. Thus, all reported results refer only to smoothed data using an FI of 0.01 or 0.1. Using an FI of 0.01 had a higher median sensitivity (0.9921) than an FI of 0.1 (0.9905, $p < 0.001$), lower specificity (0.9979 vs. 0.998; $p < 0.001$), with no difference in accuracy (0.9971 vs. 0.9971; $p = 0.134$) or DSI (0.9894 vs. 0.9896; $p = 0.066$). These measures are all very high indicating that a range of FI values may produce visually indistinguishable brain extractions.

5. Conclusion

BET performs well at brain extraction on thresholded, 1mm smoothed CT images with an FI of 0.01 or 0.1. Smoothing before applying BET is an important step not previously discussed. Analysis code is provided.

*Principal Corresponding Author

Email addresses: jmusche1@jhu.edu (John Muschelli), nullman1@jhmi.edu (Natalie Ullman), dhanley@jhmi.edu (Daniel F. Hanley), ccrainic@jhsph.edu (Ciprian M. Crainiceanu)

Validated Automatic Brain Extraction of Head CT Images

John Muschelli[c,*], Natalie Ullman[d], Daniel F. Hanley[d], Ciprian M. Crainiceanu[c]

[c]*Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA*
[d]*Department of Neurology, Division of Brain Injury Outcomes, Johns Hopkins Medical Institutions, Baltimore, MD, USA*

Keywords: CT, skull stripping, brain extraction, validation

6. Introduction

X-ray computed tomography (CT) scanning is widely available and is a commonly used diagnostic tool in clinical settings [1, 2, 3]. Though analysis of CT images is typically done by qualitative visual inspection, detailed quantification of information using neuroimaging tools is of interest. The reason for this interest is that qualitative inspection of CT scans provides limited quantifiable information that can be used in research. A fundamental processing step for producing quantifiable, reproducible, information about the brain is to extract the brain from the CT image. This process is called brain extraction or skull stripping. This step is necessary because CT images often contain non-brain human structures (e.g. skull, eyes, skin) or non-human elements (e.g. pillow, medical devices) that are not pertinent to brain research. We propose a validated automated solution to brain extraction in head CT scans using established neuroimaging software.

In magnetic resonance imaging (MRI), brain extraction has been extensively studied and investigated (see Wang et al. [4] for a good overview of methods). While an extensive literature accompanied by software exist for brain MRI scans, the same is not true for brain CT. Inspired by the brain MRI literature and software tools, we have adapted the Brain Extraction Tool (BET) [5], a function of the FSL [6] neuroimaging software (v5.0.4), to automatically extract the brain from a CT scan. Variations of this pipeline have been presented before in Solomon et al. [7], and have been replicated in more detail in Rorden et al. [8]. Neither presented a formal validation against a set of manually segmented brain images, which is the goal of our study.

7. Methods

7.1. Participants and CT data

We used CT images patients enrolled in the MISTIE (Minimally Invasive Surgery plus recombinant-tissue plasminogen activator for Intracerebral Evacuation) and ICES (Intraoperative CT-Guided Endoscopic Surgery) trials. These patients had an intracranial hemorrhage at time of scanning; for inclusion criteria, see Mould et al. [9]. CT data were collected as part of the Johns Hopkins Medicine IRB-approved MISTIE research studies with written consent from participants.

7.2. Imaging Data

The study protocol was executed with minor, but important, differences across the 13 sites. Slice thickness of the image varied within the scan for 2 scans, referred to as variable slice thickness. For example, a scan may have 10 millimeter (mm) slices at the top and bottom of the brain but with 5mm slices in the middle of the brain. Therefore, the scans analyzed had different voxel (volume element) dimensions and image resolution prior to registration to the template. These conditions represent how scans are presented for evaluation in many diagnostic cases.

*Principal Corresponding Author
Email addresses: jmusche1@jhu.edu (John Muschelli), nullman1@jhmi.edu (Natalie Ullman), dhanley@jhmi.edu (Daniel F. Hanley), ccrainic@jhsph.edu (Ciprian M. Crainiceanu)

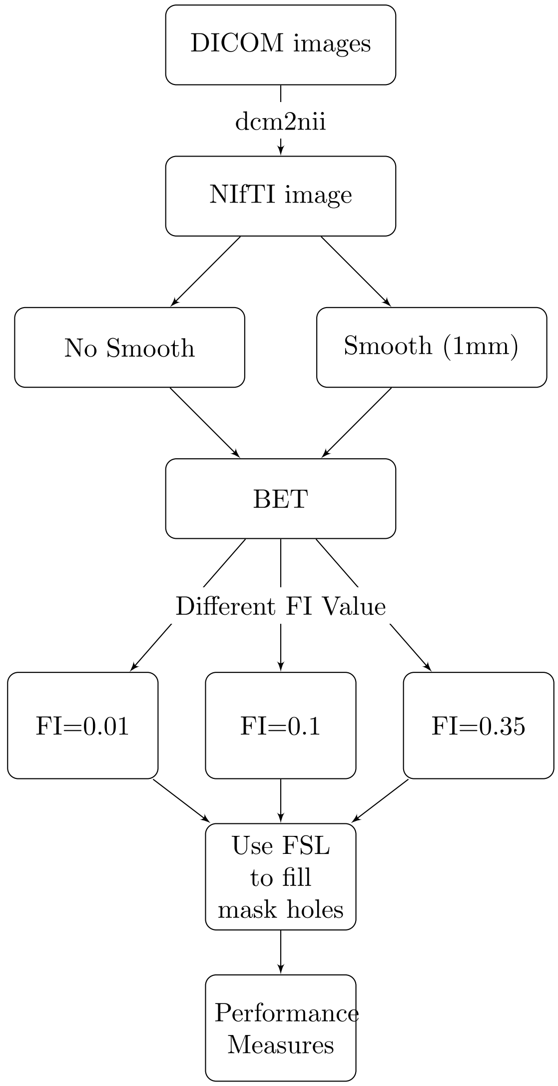

Figure 1: Overall workflow of processing of data. After the raw data has been processed and areas have been extracted, surfaces can be rendered. We are concerned with those steps in orange: creating surfaces, and export to the web. The last branch shows 2 options for export: publishing the figure to the web or enclosing it in a folder with all tools to render it. The second option would allow users to include these zipped directories as supplementary figures until more widely used.

7.3. Manual and Automated Brain Extraction

We analyzed 19 scans, corresponding to 16 unique patients. Brain tissue was manually segmented as a binary mask from DICOM (Digital Imaging and Communications in Medicine) images using the OsiriX imaging software (OsiriX v.4.1, Pixmeo; Geneva, Switzerland) by one expert reader. CT brain images and the binary mask obtained using manual segmentation were exported from OsiriX to DICOM format. Images with gantry tilt were corrected using a customized MATLAB (The Mathworks, Natick, Massachusetts, USA) user-written script (`http://bit.ly/1ltIM8c`). Images were converted to the Neuroimaging Informatics Technology Initiative (NIfTI) data format using `dcm2nii` (2009 version, provided with MRIcro [10]). Images were constrained to values -1024 and 3071 HU to remove potential image rescaling errors and artifacts. No interpolation was done for images with a variable slice thickness. Thickness was determined from the first slice converted and was assumed homogeneous throughout the image. The image processing pipeline can be seen in Figure 1.

Images were thresholded to a brain tissue range (0-100 HU). We applied BET to this raw image and to the smoothed image using a Gaussian kernel (σ = 1mm, using FSL), followed by a re-thresholding of the smooth image using the 0-100 HU range. When BET was applied, we varied fractional intensity (FI) parameter to determine its influence on performance: we used values of 0.35 (as recommended in Rorden et al. [8]), 0.1, 0.01.

We also present one example case which demonstrates that brain extraction performance was acceptable only after smoothing before BET was applied.

7.4. Measuring and Testing Brain Extraction Performance

Five common measurements of performance were calculated for each image: sensitivity, specificity, accuracy, and the Dice Similarity Index (DSI). See Inline Supplementary Methods 1 for the calculation of each measure.

[Insert Supplementary Methods 1 here]

We took the paired difference of each measure using different pipelines (e.g. 0.01 vs. 0.1, smoothed data). We tested these differences using the Wilcoxon signed-rank test.

8. Results

Figure 2A illustrates the performance of each variation of the BET pipeline in Figure 1. The smoothed pipelines (top) perform better than the unsmoothed pipelines (bottom) and BET performs poorly on some scans without smoothing.

Figure 2B displays the performance for brain extraction for the pipelines using smoothed brain data. Because the performance for all metrics was high when using smooth images, it was necessary change the y-axis from $[0, 1]$ to $[0.95, 1]$. Using an FI of 0.01 or 0.1 performed better than 0.35; we will focus and compare these results for these values of FI when BET was applied to smooth images. Using an FI of 0.01 had a higher median sensitivity (0.9921) than an FI of 0.1 (0.9905, $p < 0.001$), lower specificity (0.9979 vs. 0.998; $p < 0.001$), and no difference in accuracy (0.9971 vs. 0.9971; $p = 0.134$) or DSI (0.9894 vs. 0.9896; $p = 0.066$). Overall, regardless of p-values, these measures are all high in practice and largely adequate for brain extraction.

Although Figure 2 displays that using FI of 0.01 or 0.1 provides adequate results of brain extraction for most cases, they perform relatively well regardless of smoothing the data. Figure 3 displays an example where using unsmoothed data performs poorly for these FI, demonstrating why smoothing is essential for a general brain extraction procedure for CT.

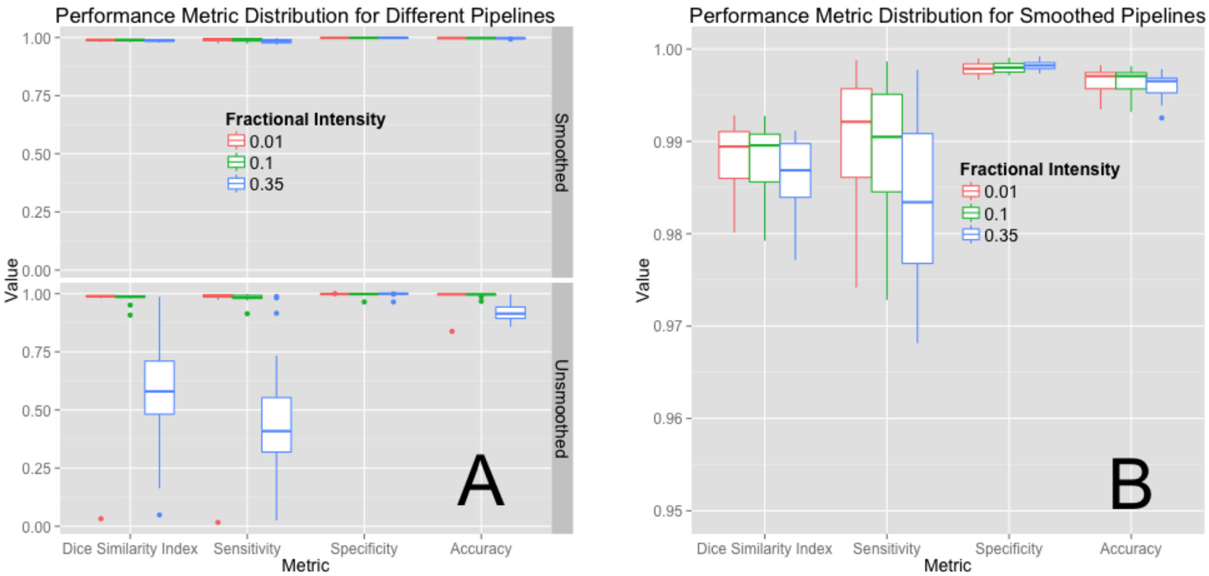

Figure 2: **Performance Metric Distribution for Different Pipelines.** Panel A displays the boxplots for performance measures when running the pipeline with a different fractional intensity (FI), using smoothed data (top) or unsmoothed data (bottom). Panel B presents the smooothed data only, rescaled to show discrimination between the different FI. Overall, FI of 0.01 and 0.1 perform better than 0.35 in all categories other than specificity. Using smoothed data improves performance in all performance metrics, markedly when an FI of 0.35 is used. Panel B demonstrates that using an FI of 0.01 on smoothed data is the best pipeline.

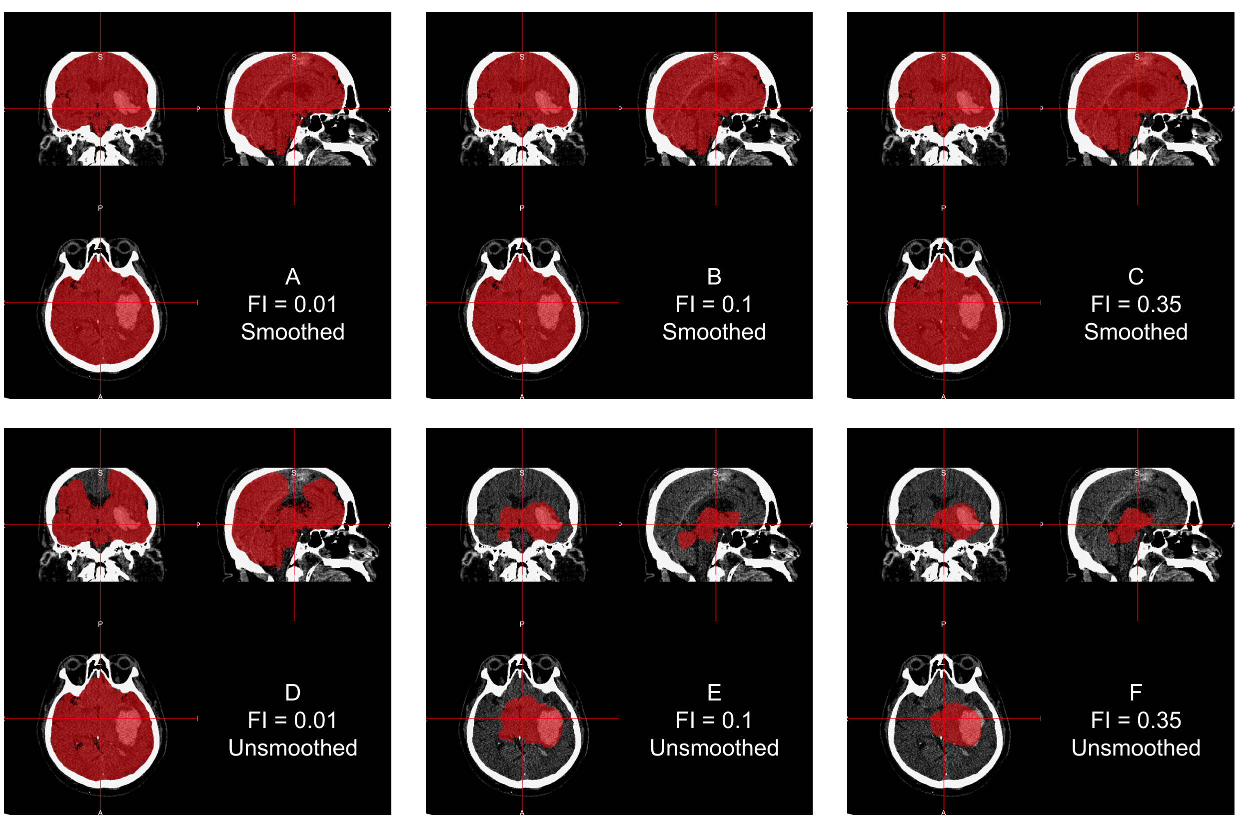

Figure 3: **Example Demonstrating How Smoothing Prior to BET is Essential** For one subject, the CT image is displayed with the brain extracted mask in red after running all pipelines. Panels A, B, and C represent applying BET using FI of 0.01, 0.01, and 0.35, respectivley, to smoothed data. Panels D, E, and F correspond to applying BET using FI 0.01, 0.01, and 0.35 on unsmoothed data. Using smoothed data is required for adequate brain extraction.

9. Conclusions

We present an automated brain extraction pipeline for CT images that was validated using gold-standard manual segmentations of brain tissues. Overall, we found that smoothing the data with a conservative smoother (1mm Gaussian kernel) and using an FI of 0.01 or 0.1 provides good brain extraction for the sample studied. We believe this pipeline is generalized for most brain CT scans. These software (FSL and R) are freely available to download under public licenses are open source. Thus, this pipeline should be widely available.

With the resulting automated brain masks output from this pipeline, researchers can now do within-tissue operations, such as intensity normalization, brain volume estimation, tissue/tumor/hemorrhage segmentation, etc.

The `R` function to perform brain extraction is located at `http://bit.ly/CTBET_RCODE` and example bash script `http://bit.ly/CTBET_BASH`.

Acknowledgements

We thank the patients and families who volunteered for this study and Genentech Inc. for the donation of the study drug (Alteplase).

Sources of Funding

The project described was supported by the NIH grant RO1EB012547 from the National Institute of Biomedical Imaging And Bioengineering, T32AG000247 from the National Institute on Aging, R01NS046309, RO1NS060910, RO1NS085211, R01NS046309, U01NS080824 and U01NS062851 from the National Institute of Neurological Disorders and Stroke, and RO1MH095836 from the National Institute of Mental Health. Minimally Invasive Surgery and rt-PA in ICH Evacuation Phase II (MISTIE II) was supported by grants R01NS046309 and U01 NS062851 awarded to Dr. Daniel Hanley from the National Institutes of Health (NIH)/National Institute of Neurological Disorders and Stroke (NINDS). ICES was led by Co-Principal Investigator Dr. Paul Vespa at the University of California Los Angeles. Minimally Invasive Surgery and rt-PA in ICH Evacuation Phase III (MISTIE III) is supported by the grant U01 NS080824 awarded to Dr. Daniel Hanley from the National Institutes of Health (NIH)/National Institute of Neurological Disorders and Stroke (NINDS). Clot Lysis: Evaluating Accelerated Resolution of Intraventricular Hemorrhage Phase III (CLEAR III) is supported by the grant U01 NS062851 awarded to Dr. Daniel Hanley from the National Institutes of Health (NIH)/National Institute of Neurological Disorders and Stroke (NINDS).

Inline Supplementary Methods 1

Let I_{ia}, I_{im} be the indicators that voxel i is labeled to be in the brain mask for the automatic and manual masks, respectively.

A voxel i is labeled to be a true positive (TP) when $I_{ia} = 1$ and $I_{im} = 1$, false positive (FP) when $I_{ia} = 1$ and $I_{im} = 0$, false negative (FN) when $I_{ia} = 0$ and $I_{im} = 1$, and true negative (TN) when $I_{ia} = 0$ and $I_{im} = 0$. The number of true positive voxels is defined as:

$$\sum_{i=1}^{V} (I_{ia} \times I_{im})$$

Sensitivity is defined as

$$\frac{\#\text{TP}}{\#\text{TP} + \text{FN}} = \frac{\sum_{i=1}^{V} (I_{ia} \times I_{im})}{\sum_{i=1}^{V} I_{im}},$$

specificity is defined as

$$\frac{\#\text{TN}}{\#\text{TN} + \text{FP}} = \frac{\sum_{i=1}^{V} \{(1 - I_{ia}) \times (1 - I_{im})\}}{\sum_{i=1}^{V} (1 - I_{im})},$$

overall accuracy is defined as:

$$\frac{\#\mathrm{TN}+\mathrm{TP}}{\#\mathrm{TN}+\mathrm{FN}+\mathrm{TP}+\mathrm{FP}} = \frac{\sum_{i=1}^{V}\left[(I_{ia}\times I_{im})+\{(1-I_{ia})\times(1-I_{im})\}\right]}{\sum_{i=1}^{V} I_{ia}+\sum_{i=1}^{V} I_{im}},$$

and the Dice Similarity Index (DSI) is defined as

$$\frac{2\times\#\mathrm{TP}}{\#\mathrm{TN}+\mathrm{FN}+\mathrm{TP}+\mathrm{FP}} = \frac{2\times\sum_{i=1}^{V}(I_{ia}\times I_{im})}{\sum_{i=1}^{V} I_{ia}+\sum_{i=1}^{V} I_{im}}.$$

References

[1] R. Sahni, J. Weinberger, Management of intracerebral hemorrhage, Vascular Health and Risk Management 3 (5) (2007) 701–709, ISSN 1176-6344, URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2291314/`.

[2] J. A. Chalela, C. S. Kidwell, L. M. Nentwich, M. Luby, J. A. Butman, A. M. Demchuk, M. D. Hill, N. Patronas, L. Latour, S. Warach, Magnetic resonance imaging and computed tomography in emergency assessment of patients with suspected acute stroke: a prospective comparison, The Lancet 369 (9558) (2007) 293–298.

[3] P. D. Schellinger, O. Jansen, J. B. Fiebach, W. Hacke, K. Sartor, A standardized MRI stroke protocol comparison with CT in hyperacute intracerebral hemorrhage, Stroke 30 (4) (1999) 765–768.

[4] Y. Wang, J. Nie, P.-T. Yap, G. Li, F. Shi, X. Geng, L. Guo, D. Shen, A. D. N. Initiative, et al., Knowledge-guided robust mri brain extraction for diverse large-scale neuroimaging studies on humans and non-human primates, PloS one 9 (1) (2014) e77810.

[5] S. M. Smith, Fast robust automated brain extraction, Human Brain Mapping 17 (3) (2002) 143155, ISSN 1097-0193, doi:\bibinfo{doi}{10.1002/hbm.10062}, URL `http://onlinelibrary.wiley.com/doi/10.1002/hbm.10062/abstract`.

[6] M. Jenkinson, C. F. Beckmann, T. E. J. Behrens, M. W. Woolrich, S. M. Smith, FSL, NeuroImage 62 (2) (2012) 782–790, ISSN 1053-8119, doi:\bibinfo{doi}{10.1016/j.neuroimage.2011.09.015}, URL `http://www.sciencedirect.com/science/article/pii/S1053811911010603`.

[7] J. Solomon, V. Raymont, A. Braun, J. A. Butman, J. Grafman, User-friendly software for the analysis of brain lesions (ABLe), Computer methods and programs in biomedicine 86 (3) (2007) 245–254.

[8] C. Rorden, L. Bonilha, J. Fridriksson, B. Bender, H.-O. Karnath, Age-specific CT and MRI templates for spatial normalization, NeuroImage 61 (4) (2012) 957–965, ISSN 1053-8119, doi:\bibinfo{doi}{10.1016/j.neuroimage.2012.03.020}, URL `http://www.sciencedirect.com/science/article/pii/S1053811912002935`.

[9] W. A. Mould, J. R. Carhuapoma, J. Muschelli, K. Lane, T. C. Morgan, N. A. McBee, A. J. Bistran-Hall, N. L. Ullman, P. Vespa, N. A. Martin, I. Awad, M. Zuccarello, D. F. Hanley, Minimally Invasive Surgery Plus Recombinant Tissue-type Plasminogen Activator for Intracerebral Hemorrhage Evacuation Decreases Perihematomal Edema, Stroke 44 (3) (2013) 627–634, ISSN 0039-2499, 1524-4628, doi:\bibinfo{doi}{10.1161/STROKEAHA.111.000411}, URL `http://stroke.ahajournals.org/content/44/3/627`.

[10] C. Rorden, M. Brett, Stereotaxic Display of Brain Lesions, Behavioural Neurology 12 (4) (2000) 191–200, ISSN 0953-4180, doi:\bibinfo{doi}{10.1155/2000/421719}, URL `http://www.hindawi.com/journals/bn/2000/421719/abs/`.