

Computational Methods for Neuroimaging in R: Stroke Hemorrhage in X-ray Computed Tomography Scanning

by

John Muschelli III

**A dissertation submitted to The Johns Hopkins University
in conformity with the requirements for the degree of
Doctor of Philosophy**

Baltimore, Maryland

May, 2016

© 2016 by John Muschelli III

All rights reserved

Chapter 1

Acknowledgments

First and foremost, I would like to thank my parents for always pushing me. I would not have achieved anything close to a PhD without them. To my mother, for making me call her and discuss my day, even when I wanted to. You always made me laugh and picked me up when I questioned continuing. Your life lessons and advice have been invaluable. To my father, you are the most hard-working person I have ever met and hope to even become $\frac{1}{10}$ as much as you are. When you paid me to study for my AP tests instead of stressing about working at a crappy job stayed with me to make the right things a priority. A thank you to my sisters, who pushed me in different ways, but with equal effectiveness. Your support can never be measured. I love you all deeply.

I would like to thank Stephanie Michele Zamora for helping me in every way, including with this thesis. Without your love and support, I would not be here and cannot ask for a better partner.

Andrew Jaffe, thank you for our many consultations and discussions, mostly over beers.

A crucial and special thank you to my friend and college Elizabeth Margaret Sweeney (my “work wife”), with whom I have gone over all projects in excruciating detail. I have found your comments to be invaluable and with you I have had the important discussions during many projects. Thank you for telling me to keep writing, at times to take my own advice, and hundreds of lunches (including work lunches) that made going into the office worthwhile. I will never find a similar colleague and friend, one forged through the depths of probability theory, bounded in work in neuroimaging, and solidified with our statistical lineage and shared moments throughout our work. It is much easier to leave Hopkins knowing you are moving on as well.

I would like to thank the Department of Biostatistics and the Epidemiology and Biostatistics training grant from the NIH (T32AG021334), which gave me the required support to work on projects I felt relevant. Thank you to Dr. Michelle Carlson and Dr. Ravi Varadhan for your mentorship.

Thank you Dr. Taki Shinohara for the collaboration and discussion. You have always given thoughtful and criticism and advice. and I wish you the best in your career.

Thank you Dr. Daniel Hanley for all the close collaborations over the years. The topics of the thesis are a testament to how closely I have worked with you and your team. I have always valued your frank criticism and your challenge to remember the clinical implications of my work.

Thank you Dr. Brian Caffo for encouraging me to come back to the PhD program after advising me in my Master’s degree. It has been the best and most transformative decisions of my life. I will always remember you having

time to chat, whether about statistics, the department, or general fun tech stuff.

Thank you Dr. Ciprian Crainiceanu for all your advising, about my degree, job search, and life. Your help on all my projects has been crucial. Your help on the stuff “not in the curriculum”, in some ways, has been even more useful . You taught me how to create a large network of connections, reminded me that of the many things important outside school, and have been a strong advocate for me in every way. You have shown the dedication of a true advisor.

Table of Contents

1 Acknowledgments	ii
Table of Contents	v
List of Tables	x
List of Figures	xi
2 Introduction	2
2.1 Creating New Tools: R Packages	2
2.2 Computed Tomography: Trials and Successes	4
2.3 Computed Tomography: Moving to Biomarkers	6
3 fslr: Connecting the FSL Software with R	9
3.1 Introduction	9
3.2 fslr workflow	10
3.2.1 fslr setup	11
3.2.2 Image preprocessing with fslr	12
3.2.3 Bias-field correction	13

3.3	Within-visit co-registration	15
3.4	Between-visit co-registration	17
3.5	Brain extraction	20
3.6	Registration to the MNI template	23
3.6.1	Applying transformations to co-registered data	24
3.7	Conclusion	26
3.8	Supplemental material	28
3.8.1	Smoothing images	28
3.8.2	Thresholding images	28
3.8.3	Eroding and dilating images	29
3.8.4	Extracting image information	30
3.8.5	Additional <code>fslr</code> functionality	30
4	PItcHPERFeCT: Primary Intracranial Hemorrhage Probability Estimation using Random Forests on CT	38
4.1	Introduction	39
4.2	Methods	41
4.2.1	Data	41
4.2.2	Participants and Imaging Data	41
4.2.3	Hemorrhage Segmentation and Location Identification	42
4.2.4	Image Processing: Brain Extraction, Registration	43
4.2.5	Image Registration	44
4.2.6	Brain Mask Erosion	46

4.2.7	Imaging Predictors	46
4.2.7.1	CT voxel intensity information	46
4.2.7.2	Local Moment Information	46
4.2.7.3	Within-plane Standard Scores	48
4.2.7.4	First-pass Segmentation	49
4.2.7.5	Contralateral Difference Images	49
4.2.7.6	Global Head Information	50
4.2.7.7	Standardized-to-template Intensity	50
4.2.8	Voxel Selection Procedure	51
4.2.9	Models	53
4.2.10	Estimating a Cutoff for Model Probability	54
4.2.11	Measuring and Testing ICH Prediction Performance . .	55
4.3	Results	56
4.3.1	Dice Similarity Index	56
4.3.2	ICH Volume Estimation	57
4.3.3	Model Choice	61
4.4	Discussion	61
4.4.1	Conclusions	65
4.5	Supplemental Material	79
4.5.1	Examples of Dice Similarity Index in Test Scans . . .	79
4.5.2	Model Specification	79

5 Quantitative Localization of Supratentorial Intracranial Hemorrhage	89
5.1 Introduction	89
5.2 Methods	90
5.2.1 Subjects and Demographics	90
5.2.2 Imaging Data	90
5.2.3 Hemorrhage Segmentation and Location Identification	91
5.2.4 Image Registration	91
5.2.5 ICH Localization and Engagement	92
5.3 Results	93
5.3.1 Prevalence of ICH Engagement in the Brain	93
5.3.1.1 ICH Localization and Engagement	94
5.3.1.2 Highest Predictive Region Analysis	95
5.4 Discussion	96
5.4.1 Summary	96
5.5 Sources of Funding	96
5.6 Disclosures	97
5.7 Tables	98
5.8 Figures	99
6 Validated Automatic Brain Extraction of Head CT Images	103
6.1 Introduction	103
6.2 Methods	105

6.2.1	Participants and CT data	105
6.2.2	Imaging Data	106
6.2.2.1	Validation of Automated Head Segmentation	106
6.2.3	Manual and Automated Brain Extraction	106
6.2.4	Image Processing	106
6.2.5	Measuring and Testing Brain Extraction Performance .	109
6.2.6	Consistency of Manual Brain Extraction	110
6.2.7	Failure Rate and Intraclass Correlation Estimate	111
6.3	Results	112
6.3.1	Manual and Automated Brain Extraction	112
6.3.2	Consistency of Manual Brain Extraction	113
6.3.3	Failure Rate and Intraclass Correlation Estimate	116
6.4	Discussion	118
7	Discussion and Conclusion	129
7.1	Neuroimaging and R	129
7.2	CT Analysis	130
7.2.1	ICH Segmentation	130
7.2.2	Predictive Regions of Stroke Severity	131
7.3	Conclusion	132

List of Tables

5.1	Distribution of the Top 10 Areas of Engagement of Highest Predictive Regions.	98
6.1	Failure Rates for each Processing Pipeline of Brain Extraction of the 1095 Scans Analyzed.	117
6.2	Failure Rates for Different Scanner Types using Smoothed Data and an FI of 0.01 Processing Pipeline of Brain Extraction of the 1095 Scans Analyzed.	117

List of Figures

3.1	Results of Inhomogeneity Correction.	14
3.2	Results of within-visit co-registration.	17
3.3	Between-visit registration process.	18
3.4	Results from FLIRT.	19
3.5	Between-visit registration results.	21
3.6	Results from BET.	22
3.7	Results from FNIRT.	25
3.8	Results of fslr functions.	29

58

4.5 Comparison of Estimated and Manual Intracerebral Hemor-

60

- 4.6 **Patient with Lowest Dice Similarity Index.** We present the patient with the lowest Dice Similarity Index (DSI), a measure of spatial overlap, from the chosen predictor model fit with a random forest. The lowest DSI was 0. The green indicates a correct classification of ICH from the model, blue indicates a false negative, where the manual segmentation denoted the area to be ICH but the predicted one did not, and red indicates a false positive, where the predicted segmentation denoted the area to be ICH but the manual one did not.

4.7 **Patient with 25th Quantile Dice Similarity Index.** We present the patient with the 25th quantile Dice Similarity Index (DSI), a measure of spatial overlap, from the chosen predictor model fit with a random forest. The 25th quantile DSI was 0.872. The green indicates a correct classification of ICH from the model, blue indicates a false negative, where the manual segmentation denoted the area to be ICH but the predicted one did not, and red indicates a false positive, where the predicted segmentation denoted the area to be ICH but the manual one did not. . . .

4.8	Patient with 75th Quantile Dice Similarity Index. We present the patient with the 75 th quantile Dice Similarity Index (DSI), a measure of spatial overlap, from the chosen predictor model fit with a random forest. The 75 th quantile DSI was 0.928. The green indicates a correct classification of ICH from the model, blue indicates a false negative, where the manual segmentation denoted the area to be ICH but the predicted one did not, and red indicates a false positive, where the predicted segmentation denoted the area to be ICH but the manual one did not.	77
4.9	Patient with Highest Dice Similarity Index. We present the patient with the highest Dice Similarity Index (DSI), a measure of spatial overlap, from the chosen predictor model fit with a random forest. The highest DSI was 0.956. The green indicates a correct classification of ICH from the model, blue indicates a false negative, where the manual segmentation denoted the area to be ICH but the predicted one did not, and red indicates a false positive, where the predicted segmentation denoted the area to be ICH but the manual one did not.	78
5.1	ICH Engagement Prevalence.	99
5.2	Highest Predictive Region (HPR) Analysis.	100
6.1	Brain Extraction Processing Pipeline.	108
6.2	Performance Metric Distribution for Different Pipelines. . .	114
6.3	Example Case where Smoothing before BET is Required. . .	115

6.4 Intracranial Volume (ICV) Estimates for Scans Less than 10 Days Post-Baseline.	118
6.5 Comparison of Performance Measures from 2 Manual Brain Segmentations.	124
6.6 Comparison of Intracranial Volume (ICV) Estimates from 2 Manual Brain Segmentations.	125

Chapter 2

Introduction

The following thesis is laid out in separate chapters. The chapters will include: 1) `fslr`: connecting the FSL software with R, 2) Segmentation of intracerebral hemorrhage in computed tomography scans using random forests, and 3) Quantitative localization of intracerebral hemorrhages.

2.1 Creating New Tools: R Packages

One of the overall goals of my work during my time as a student has been to build tools for analysis, with a specific target population of statisticians as users. In chapter 3, we discuss the `fslr` package we develop for the R programming language (R Core Team, 2015). This package interfaces with FSL (FMRIB Software Library), a commonly-used software for processing and analyzing neuroimaging data (Jenkinson et al., 2012). Although this package has essentially the same functionality of FSL, it allows R users to interface with tested software without learning specific syntax. Mainly, we want to allow statisticians to process and analyze neuroimaging data in the same software

they do their analysis, which is becoming more commonly R.

I have created a series of other R packages and will discuss them here. Much of neuroimaging analysis is done in MATLAB (The Mathworks, Natick, Massachusetts, USA), especially as the neuroimaging field has had a root in engineering and MATLAB is a commonly-used software in engineering. Moreover, the largely popular statistical parametric mapping software (SPM) (Wellcome Trust Centre for Neuroimaging, London, UK) has been implemented in MATLAB. Although MATLAB has a similar syntax to R in some cases, many statisticians do not use MATLAB. Thus, statisticians who primarily or exclusively use R for analysis do not have access to the large suite of neuroimaging tools.

As such, I created the `matlabr` package to easily call MATLAB code and scripts from R and return the results if necessary. Building on top of this package, I created the `spm12r` package which downloads the SPM (version 12, SPM12) software and provides R functions to call relevant SPM12 functions and processing pipelines. Although most of my research here will be on structural neuroimaging, these pipelines can be applied to functional magnetic resonance imaging (fMRI), a highly popular research tool for analyzing the brain.

Conveying the results of these analyses is crucial for collaboration, so we created the `brainR` package (Muschelli, Sweeney, and Crainiceanu, 2014). This package relies on `rgl`, an R implementation of the graphics programming interface Open Graphics Library (OpenGL) (Shreiner et al., 2009), that allows for 3D rendering. These rendered objects can be interactively explored on

screen. **brainR** uses **rgl** to generate the figures and export these figures as webpages using WebGL, the web implementation of OpenGL. Using these objects and the X Toolkit (XTK) (<https://github.com/xtk/X>), one can create customizable webpages with versatile user-interaction capabilities.

Overall, one of the large goals of my work has been to implement analyses within R and port the functionality of other neuroimaging software into R for statisticians to use. This has allowed for the implementation of full pipelines in R and allows a large community to use tools without learning specific syntax for that software. Overall, I believe these tools will allow statisticians to more easily perform imaging-specific processes for analysis and give them greater control of the complete analysis pipeline.

2.2 Computed Tomography: Trials and Successes

Research in neuroimaging is dominated by magnetic resonance imaging (MRI), with over 20,000 publications per year (Vlaardingerbroek and Boer, 2013). X-ray computed tomography (CT) is used extensively for diagnosis and prognosis, but is primarily focused on clinical applications. After working with the aforementioned MRI software tools, I found that the direct application of these to CT were not always sufficient. Therefore, I adapt some of the tools for MRI for CT processing and the same analytic framework from MRI for CT analysis.

In Muschelli et al. (2015), I create an analysis pipeline to take a raw CT image and perform brain extraction. This operation determines which areas of the image are located within the skull, including brain tissue, ventricles,

the brain stem, and in the cases of patients with stroke, the stroke hemorrhage. We use the brain extraction tool (BET), introduced and validated by **smith_fast_2002** a function of FSL (Jenkinson et al., 2012).

A similar method had been proposed previously (**solomon_user-friendly_2007**; Rorden et al., 2012) but never formally validated or performed on a large number of scans. In Muschelli et al. (2015), we validate 36 images from 36 patients with intracranial hemorrhage using our segmentation method compared to those done manually by one expert CT reader and had a high level of overlap. We further analyze 1095 longitudinal scans from 129 patients and found the failure rates are low (5.2%) and the intracranial volume is estimated consistently in the same patient (intraclass correlation coefficient: 0.929, 95% CI: 0.91, 0.945).

Though brain segmentation can give you an good estimate of intracranial volume, it can also be seen as a crucial preprocessing step towards doing analysis on the brain. In chapter 4, we employ this skull stripping as a preliminary and crucial step in the analysis. After we obtain a brain-only image, we create a series of derivative images from the single CT scan by using morphological operations, such as smoothing, taking local neighborhood moments, such as standard deviation and skew, and normalizing images with respect to intensity or space.

After creating these predictors, we create a voxel (3-dimensional pixel or “volume element”) selection procedure to filter the areas likely to be hemorrhage. We then create statistical models for the probability a given voxel represents a hemorrhage in the brain. We use the binary mask given by manual

segmentation of the hemorrhage as the outcome in these models. We employ logistic regression, logistic regression with a LASSO penalty, a generalized additive model, and a random forest algorithm to predict the voxel probability. We use a small training set of images ($N = 10$) and predict on a set of test images ($N = 102$) and estimated high overlap between the predicted binary hemorrhage mask and the mask done by manual segmentation. Overall, we chose the random forest algorithm because it had the highest overlap and the highest Pearson correlation between the estimated volume and the manual segmentation volume 0.932 (95% CI: 0.901, 0.954).

2.3 Computed Tomography: Moving to Biomarkers

In chapter 5, we look at population level analyses of CT scans of patients with hemorrhage. In this chapter we registered the CT scans to a CT template (Rorden et al., 2012), which is located in “template space”, which can be considered a population-level space. Each scan is registered to this image, and a transformation is estimated. We then apply this transformation to the manually-segmented hemorrhage masks to put the hemorrhage masks in the template space. Although we use the hemorrhage masks from manual segmentation as they are the gold standard, we could just have easily used the predicted segmentation from the method described in chapter 4.

Now that scans and masks from different patients are in the same space, the voxels should be comparable across patients. We first average the hemorrhage masks to create a 3-dimensional map of where hemorrhages occur in this

population. As this template is in the same space as analyses commonly done with MRI, there are whole-brain segmentations, referred to as atlases, which assign each voxel to a specific structure of the brain. Using an atlas, we can then describe the areas of the brain where hemorrhages are engaged in this population, giving a more quantitative description of location. For example, a reader may assign a patient to having a hemorrhage in the putamen, whereas our method could estimate that the hemorrhage is 67% in the putamen and describe the other areas the hemorrhage engages. Therefore, we believe this process allows us to get an automated measure of engagement and one that is more quantitative than the categorical classification currently done by readers.

We also perform an exploratory analysis that does testing a each voxel to see whether the presence of a hemorrhage is associated with a stroke severity score, the National Institute of Health Stroke Scale (NIHSS). At each voxel, we perform a t-test, which yields a p-value. We threshold this p-value, which yields a “high predictive region” for NIHSS. For each patient, we create a one-dimensional measure describing how much a the hemorrhage from that patient engages with this region. We show that this measure is more predictive of NIHSS compared to the categorical classification currently done, even after adjusting for other known variables associated with NIHSS. Further validation of these regions is needed.

References

- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.R-project.org/>.
- Jenkinson, Mark, Christian F. Beckmann, Timothy E. J. Behrens, Mark W. Woolrich, and Stephen M. Smith (2012). “FSL”. In: *NeuroImage* 62.2, pp. 782–790. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2011.09.015](https://doi.org/10.1016/j.neuroimage.2011.09.015). URL: <http://www.sciencedirect.com/science/article/pii/S1053811911010603> (visited on 01/15/2014).
- Muschelli, John, Elizabeth Sweeney, and Ciprian Crainiceanu (2014). “brainR: Interactive 3 and 4D Images of High Resolution Neuroimage Data”. In: *R Journal* 6.1, pp. 41–48.
- Shreiner, Dave et al. (2009). *OpenGL programming guide: the official guide to learning OpenGL, versions 3.0 and 3.1*. Addison-Wesley Professional.
- Vlaardingerbroek, Marinus T and Jacques A Boer (2013). *Magnetic resonance imaging: theory and practice*. Springer Science & Business Media.
- Muschelli, John, Natalie L. Ullman, W. Andrew Mould, Paul Vespa, Daniel F. Hanley, and Ciprian M. Crainiceanu (2015). “Validated automatic brain extraction of head CT images”. In: *NeuroImage*. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2015.03.074](https://doi.org/10.1016/j.neuroimage.2015.03.074). URL: <http://www.sciencedirect.com/science/article/pii/S1053811915002700> (visited on 04/20/2015).
- Rorden, Christopher, Leonardo Bonilha, Julius Fridriksson, Benjamin Bender, and Hans-Otto Karnath (2012). “Age-specific CT and MRI templates for spatial normalization”. In: *NeuroImage* 61.4, pp. 957–965. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2012.03.020](https://doi.org/10.1016/j.neuroimage.2012.03.020). URL: <http://www.sciencedirect.com/science/article/pii/S1053811912002935> (visited on 01/15/2014).

Chapter 3

fslr: Connecting the FSL Software with R

3.1 Introduction

FSL (FMRIB Software Library) is a commonly-used software for processing and analyzing neuroimaging data (Jenkinson et al., 2012). This software provides open-source, command-line tools and a graphical user interface (GUI) for image processing tasks such as image smoothing, brain extraction (Smith, 2002), bias-field correction, segmentation (Zhang, Brady, and Smith, 2001), and registration (Jenkinson and Smith, 2001; Jenkinson et al., 2002). Many of these functions are used extensively in medical imaging pipelines. According to a recent survey paper by Carp (2012), 13.9% of published neuroimaging studies used FSL.

There exist a number of R packages for reading and manipulating image data, including **AnalyzeFMRI** (Bordier et al., 2011), **RNiftyReg** (Clayden, 2015), and **fmri** (Tabelow and Polzehl, 2011) (see the Medical Imaging CRAN task view <http://cran.r-project.org/web/views/MedicalImaging>.

[html](#) for more information). Although these packages are useful for performing image analysis, much of the fundamental functionality that FSL and other imaging software provide is not currently implemented in R. In particular, this includes algorithms for performing slice-time correction, motion correction, brain extraction, tissue-class segmentation, bias-field correction, co-registration, and normalization. This lack of functionality is currently hindering R users from performing complete analysis of image data within R. Instead of re-implementing FSL functions in R, we propose a user-friendly interface between R and FSL that preserves all the functionality of FSL, while retaining the advantages of using R. This will allow R users to implement complete imaging pipelines without necessarily learning software-specific syntax.

The **fslr** package relies heavily on the **oro.nifti** (Whitcher, Schmid, and Thornton, 2011) package implementation of images (referred to as `nifti` objects) that are in the Neuroimaging Informatics Technology Initiative (NIfTI) format, as well as other common image formats such as ANALYZE. **oro.nifti** also provides useful functions for plotting and manipulating images. **fslr** expands on the **oro.nifti** package by providing additional functions for manipulation of `nifti` objects.

3.2 fslr workflow

The general workflow for most **fslr** functions that interface with FSL is as follows:

1. Filename or `nifti` object is passed to **fslr** function.

2. FSL command is created within `fslr` function and executed using the `system` command.
3. Output is written to disk and/or read into R and returned from the function.

From the user's perspective, the input/output process is all within R. The advantage of this approach is that the user can read in an image, do manipulations of the `nifti` object using standard syntax for arrays, and pass this object into the `fslr` function without using FSL-specific syntax written in a shell language. Also, one can perform image operations using FSL, perform operations on the `nifti` object in R that would be more difficult using FSL, and then perform additional operations using FSL by passing that object to another `fslr` command. Thus, users can create complete pipelines for the analysis of imaging data by accessing FSL through `fslr`.

3.2.1 `fslr` setup

To use `fslr`, a working installation of FSL is required. The following code was run using FSL version 5.0.0. `fslr` must also have the path of FSL specified. If using R from a shell environment, and the `FSLDIR` environment variable is set (which can be done when installing FSL), `fslr` will use this as the path to FSL. If using R through a graphical user interface (GUI) such as RStudio (RStudio, Boston, MA), environmental variables and paths are not explicitly exported. Therefore, `FSLDIR` is not set, and the path to FSL can be specified using `options(fsl.path="/path/to/fsl")`.

fslr also requires an output type for the format of images returned from FSL. Some **fslr** functions produce intermediate files that the user may want removed after the command is executed. When the filename argument is passed to a **fslr** command, the extension of the file does not need to be specified, but simply the prefix. When the command is executed, the FSL command appends an extension, and to remove this file, using the R command `file.remove`, the extension for the file is required. If working in a shell environment, **fslr** will use the environment variable for output type: `FSLOUTPUTTYPE`. If working in a GUI, the default is given by `NIFTI_GZ`, which returns compressed NIfTI images, ending in “`.nii.gz`”. This can be changed by setting the `fsl.outputtype` option. See <http://fsl.fmrib.ox.ac.uk/fsl/fsl-4.1.9/fsl/formats.html> for a description of FSL output types.

The R code below is all that is needed to load the **fslr** package, set the path to FSL, and specify the output type for files, respectively.

```
library(fslr)
options(fsl.path="/usr/local/fsl")
options(fsl.outputtype = "NIFTI_GZ")
```

3.2.2 Image preprocessing with **fslr**

We present a complete analysis of structural magnetic resonance imaging (MRI) data performed using **fslr** and R. Images were obtained from a patient with multiple sclerosis (MS) at 2 different visits (Sweeney et al., 2013), located at bit.ly/FSL_Data. At each visit, the image modalities obtained were T1-weighted (T1), T2-weighted (T2), fluid-attenuated inversion recovery (FLAIR), and proton density (PD). In this example we will perform a MRI

bias-field correction using FAST (FMRIB’s Automated Segmentation Tool) (Zhang, Brady, and Smith, 2001), co-register scans within visits to the T1 image of that visit, and register T1 images between visits. Once these operations have been performed, one can take within-modality difference images to see the changes between visits. We will also register all images to a common stereotaxic template, as this is common in population-based analyses.

3.2.3 Bias-field correction

MRI images typically exhibit good contrast between soft tissue classes, but intensity inhomogeneities in the radio frequency (RF) field can cause differences in the ranges of tissue types at different spatial locations. These inhomogeneities can cause problems with algorithms based on histograms, quantiles, or raw intensities (Zhang, Brady, and Smith, 2001). Therefore, correction for image inhomogeneities is a crucial step in many analyses. FSL implements the bias-field correction from Guillemaud and Brady (1997) in its FAST segmentation pipeline (Zhang, Brady, and Smith, 2001).

The `fsl_biascorrect` command from `fslr` will create the corrected images. We pass in the filename in the `file` argument, any additional options for FSL in the `opts` argument, such as `-v` for verbose diagnostic outputs, and the `output filename` in the `outfile` argument, which is our inhomogeneity-corrected image.

```
fsl_biascorrect(file = "01-Baseline_T1.nii.gz",
                 outfile= "01-Baseline_T1_FSL_BiasCorrect",
                 opts= "-v")
```

We can observe the difference in voxel values from the baseline T1 image

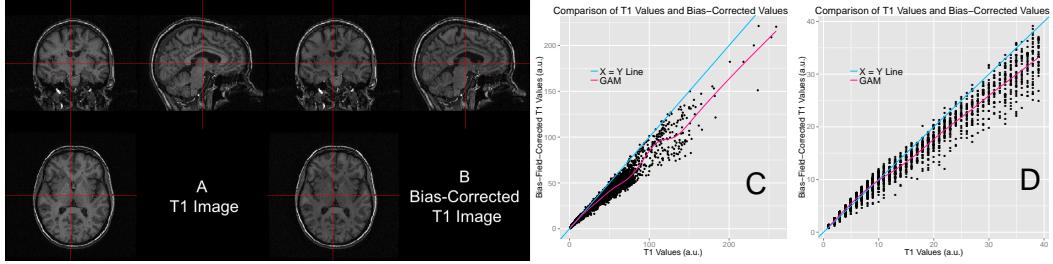


Figure 3.1: Results of Inhomogeneity Correction. We present the original T1 image (A), bias-corrected T1 image (B), and the scatterplot of the sampled values comparing the values from the T1 image to the bias-corrected values (C). We see in panel C for values in the low range of the data (< 10), the T1 values and bias-corrected T1 values, on average, fall along the diagonal (blue line), which is further illustrated in panel D, which plots values < 40 . Values > 10 for the original T1 image are lower than the bias-corrected T1 values shown by a generalized additive model (GAM) smoother (pink line, panel C).

compared to the bias-corrected version in Figure 3.1. In panel A we display the T1 image, and in panel B we display the bias-corrected T1 image. The T1 image is brighter in the middle of the image, while the bias-corrected image is more uniform in white matter (brighter regions). As this difference may be hard to distinguish visually, we present the scatterplot of these images in Figure 3.1C, using the `ggplot2` package (Wickham, 2009). Note, both scales are in arbitrary units (a.u.).

The blue line in Figure 3.1C represents the 45° diagonal line, where the original and bias-corrected image intensities are equal ($X = Y$), and the pink line represents a generalized additive model (GAM) (Hastie and Tibshirani, 1990) scatterplot smoother estimate obtained using the `mgcv` package (Wood, 2011). We see that for values in the low range of the data (< 10), the T1 values and bias-corrected T1 values, on average, fall along the diagonal, but in the higher range the bias-corrected values are lower.

3.3 Within-visit co-registration

All subsequent steps will be performed on the bias-corrected images. We will first co-register the images within each separate visit to the T1 image from that visit. This operation overlays the images on one another and allows us to investigate joint distributions of voxel intensities from different image modalities. This is performed using FMRIB’s Linear Image Registration Tool (FLIRT) (Jenkinson and Smith, 2001; Jenkinson et al., 2002). As the images are from the same individual, we may assume that the overall shape of the brain has not changed, but each scan may have undergone a translation and/or rotation in space. Therefore, we will use a rigid-body transformation, with 6 degrees of freedom (dof).

The `fslr` command `flirt` calls the FSL command `flirt`, taking the input image (`infile`) and the reference image that serves as a template (`reffile`). Any additional options for FLIRT can be passed using the `opts` argument. We will use the defaults (i.e. trilinear interpolation) and the `-v` option for diagnostic messages to be printed. Since we are doing a rigid-body transformation, we set the degrees of freedom (dof) to 6. Here we present the code for registering the baseline T2 image to the baseline T1 image; we will subsequently repeat this process for the baseline FLAIR and PD images and for the follow-up scans.

```
flirt(reffile = "01-Baseline_T1_FSL_BiasCorrect",
      infile = "01-Baseline_T2_FSL_BiasCorrect",
      omat = "01-Baseline_T2_FSL_BiasCorrect_rigid_to_T1.mat",
      dof = 6,
      outfile = "01-Baseline_T2_FSL_BiasCorrect_rigid_to_T1",
      opts = "-v")
```

The resulting image transformation is stored using the file name passed to the `omat` (**output matrix**) argument. This matrix can be used to transform other images, that were in the same space as the input image, to the reference image space. The `fslr` package will currently only return `nifti` objects, and not a list of objects, such as the output image, transformation matrix, etc. Thus, any transformation files that are needed after the command is executed must be specified.

After co-registration, one could compare images of different modalities at the same voxels, such as T1 versus FLAIR images, which is presented in Figure 3.2. The images are presented at the same cross section for the baseline T1 (Panel 3.2A) and FLAIR (Panel 3.2B) images. The same brain areas are presented in each modality, indicating adequate registration.

In the previous example, we presented a rigid-body transformation, using the default parameters. `flirt` has options for different cost functions to optimize over, interpolation operators to estimate voxel intensity, and additional degrees of freedom for performing affine transformations. These options can be passed to the FSL `flirt` command using the `opts` argument in the `fslr` `flirt` function.

Note that each `fslr` function has a corresponding help function, which is the `fslr` command appended with `.help()`, which prints out the FSL help page for that function. For example, users can see which options can be changed in the FSL `flirt` command by executing the `flirt.help()` function. Additional non-linear registration techniques are presented in the section “Registration to the MNI Template”.

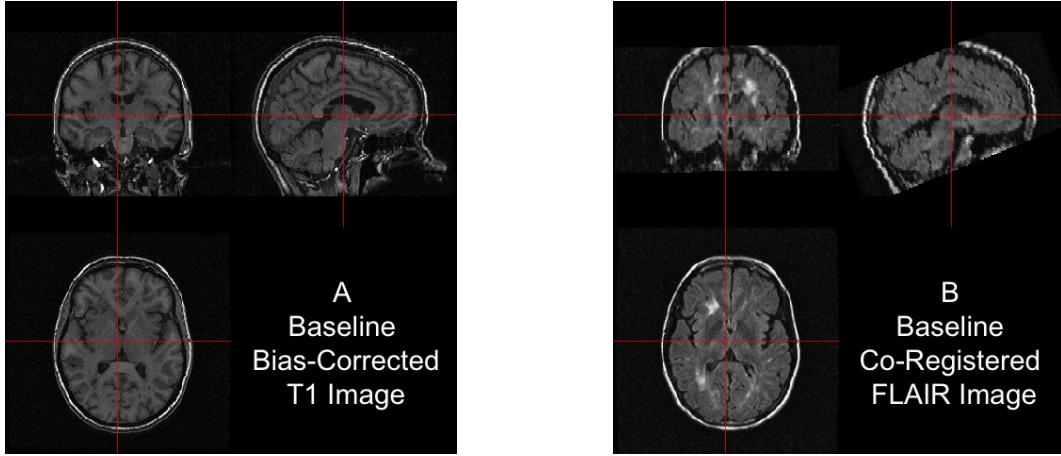


Figure 3.2: Results of within-visit co-registration. We present the bias-corrected T1 image (A) and the co-registered bias-corrected FLAIR image (B).

3.4 Between-visit co-registration

Though across-modality comparisons can be achieved by performing within-visit co-registration, across-visit registration is required for assessing within-modality differences between longitudinal scans. To compute difference images, we co-register follow-up images to the baseline images within each modality. Similar to the within-visit co-registration, we use a rigid-body transformation. We will register the T1 images from baseline and follow-up, and apply this transformation to the co-registered-to-T1 images from above (see Figure 3.3 for illustration).

Though this registration involves two interpolations of the data and may not be optimal for within-modality comparisons, we have already obtained the co-registered-to-T1 images in the section “Within-Visit Co-registration” and must perform only one additional registration. This operation also demonstrates how to apply transformation matrices in `fslr`. Here we register the

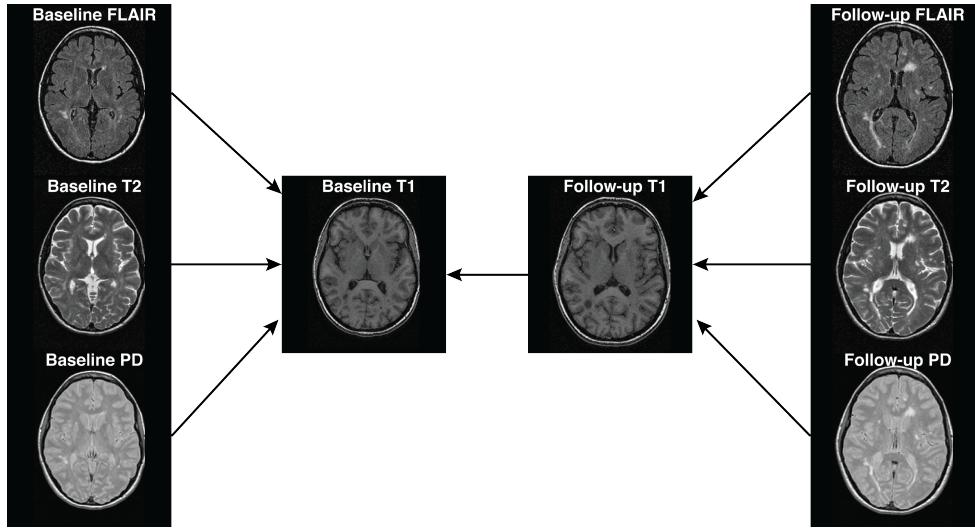


Figure 3.3: Between-visit registration process. First, we registered all scans within a visit (baseline or follow-up) to the T1 image. We then registered the follow-up T1 image to the baseline T1 image and applied the transformation to the follow-up T2, FLAIR, and PD images previously co-registered to the follow-up T1 image.

follow-up T1 image to the baseline T1 image, again using a rigid-body transformation (6 dof):

```
flirt(reffile = "01-Baseline_T1_FSL_BiasCorrect",
      infile = "01-Followup_T1_FSL_BiasCorrect",
      omat = "01-Followup_T1_FSL_BiasCorrect_rigid_to_BaseT1.mat",
      dof = 6,
      outfile = "01-Followup_T1_FSL_BiasCorrect_rigid_to_BaseT1",
      opts = "-v")
```

Now, both T1 images are aligned in the space of the baseline T1 image. We present the results in Figure 3.4: the bias-corrected baseline T1 image in A and the co-registered bias-corrected follow-up T1 in B. The images displayed at the same cross section correspond to the same brain area, indicating a good registration.

Using the `flirt_apply` function from `fslr`, we can apply the transformation

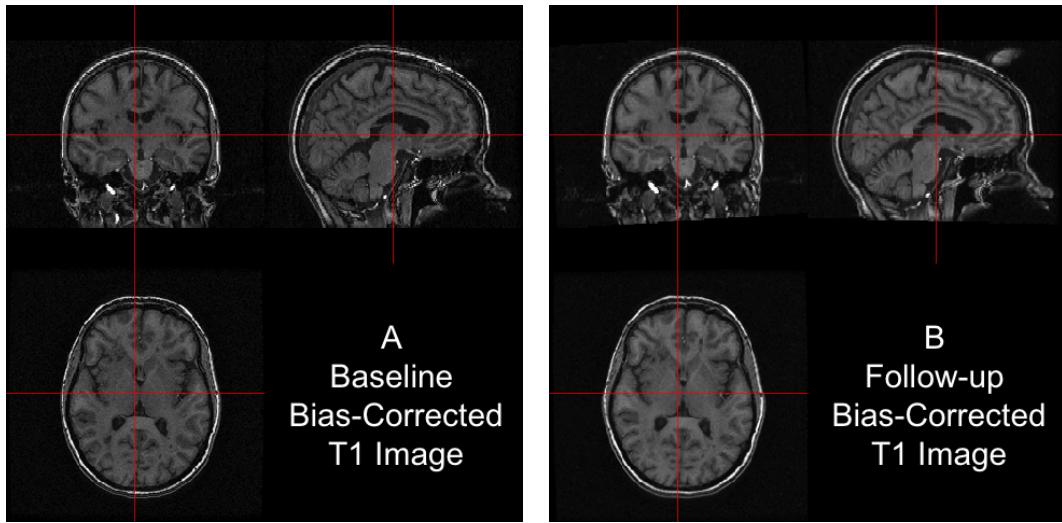


Figure 3.4: Results from FLIRT. The bias-corrected baseline T1 is presented in (A) and the registered bias-corrected follow-up T1 is presented in (B), each displayed at the same intersection. We observe that the observed images correspond to the same brain area, indicating a good registration.

matrix to the T2, PD, and FLAIR images from the follow-up visit, previously co-registered to the T1 from follow-up, to align them to the baseline T1 image space. The code below aligns the follow-up T2 image, previously registered to the follow-up T1 image, to the baseline T1 image:

```

flirt_apply(reffile = "01-Baseline_T1_FSL_BiasCorrect",
            # register to this
            infile = "01-Followup_T2_FSL_BiasCorrect_rigid_to_T1",
            # reg to Followup T1
            initmat = "01-Followup_T1_FSL_BiasCorrect_rigid_to_BaseT1.mat",
            #transform
            outfile = "01-Followup_T2_FSL_BiasCorrect_rigid_to_BaseT1"
            # output file
)

```

In Figure 3.5, we display each image after FLIRT has been applied. Each image is in the baseline T1 image space, displayed at the same cross section. Each panel shows the same brain areas across modalities, indicating adequate registration. We see that some areas of the brain are cropped from the field of view, which may be problematic if relevant brain areas are removed. We have registered all images with the skull and extracranial tissue included. A better method may be to perform registration on brain tissues only, in which case we must perform brain extraction.

3.5 Brain extraction

The process of extracting brain tissue from the acquired image, referred to as brain extraction or skull stripping, is a crucial step in many analyses. We will perform brain extraction using FSL's brain extraction tool (BET) (Smith, 2002; Jenkinson, Pechaud, and Smith, 2005) using parameters recommended by Popescu et al. (2012), which were derived from patients with MS. No other published R package on CRAN (e.g [AnalyzeFMRI](#), [RNiftyReg](#), or [fmri](#)) has brain extraction functionality for brain imaging. Other neuroimaging software provide brain extraction, such as AFNI (Cox, 1996), SPM (Ashburner

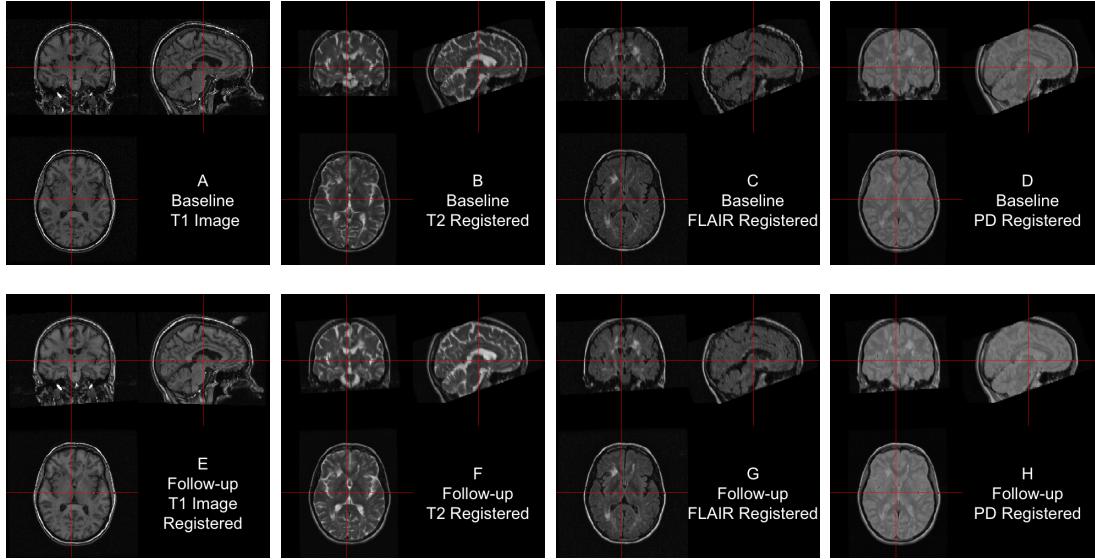


Figure 3.5: Between-visit registration results. The complete set of acquired images, first co-registered within visit to the T1 image of that visit, then registered to the baseline T1 image using the follow-up T1 to baseline T1 transformation matrix. All registrations performed rigid-body transformations.

and Friston, 2005), and Freesurfer (Fischl, 2012), and multi-atlas label fusion techniques (Doshi et al., 2013).

```
fslbet(infile = '01-Baseline_T1',
       outfile = "01-Baseline_T1_FSL_BiasCorrect_Brain",
       opts = "-B -f 0.1 -v", # from Popescu et al.
       betcmd = "bet",
       intern=FALSE)
```

We ran BET on the non-corrected T1 image as the `-B` option performs inhomogeneity correction from FAST as part of the procedure. The option `-f 0.1` denotes the fractional intensity (FI) parameter in BET: it varies between 0 and 1 and determines the location of the edge of the segmented brain image; smaller values correspond to larger brain masks. In Figure 3.6, the bias-corrected T1 image is shown with the brain mask overlaid in red (panel A)

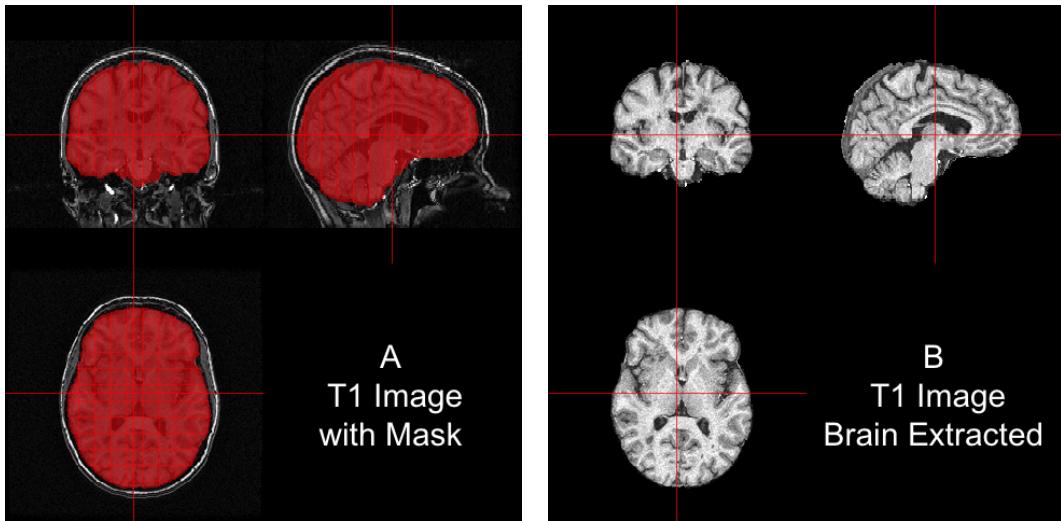


Figure 3.6: Results from BET. In (A), we show the bias-corrected T1 image with the mask from BET overlaid in red. In (B), we display the extracted brain. We see that the brain extraction performed well, not including any areas of the skull or the neck while not discarding large areas of the brain.

and the resulting masked brain (panel B). We see that the brain extraction performed well, not including any areas of the skull or the neck while not discarding brain tissue. Towards the back of the brain, some areas of the subarachnoid space remain, which may be unacceptable for certain analyses, such as estimation of the volume of brain tissue.

Note that `fslbet` writes both a file containing the brain-extracted image and another image containing the binary brain mask. As all other images are registered to the baseline T1 space, we can use this mask to extract the brain from other images, such as the baseline T2 image, using the `fslr` function `fslmask`. In this example, we mask the registered-to-T1, bias-corrected T2 image with the binary brain mask, save the image to the filename specified in `outfile`, and also set the `retimg` option to `TRUE`, indicating the `fslmask`

command to return the **image**. The returned object is a **nifti** object, assigned to the R object **mask**:

```
mask <- fslmask(file="01-Baseline_T2_FSL_BiasCorrect_rigid_to_T1",
                  mask = "01-Baseline_T1_FSL_BiasCorrect_Brain_mask",
                  outfile = "01-Baseline_T2_FSL_BiasCorrect_rigid_to_T1_Brain",
                  retimg = TRUE)
```

We now have all images in the same stereotaxic space with the brain extracted.

3.6 Registration to the MNI template

In many studies, information is aggregated across a population of images from different participants. For the information to have the same interpretation spatially across participants, images from all participants need to be aligned in the same stereotaxic space (“template” space), requiring registration to a standard template image. A frequently used set of templates are provided by MNI (Montreal Neurological Institute). We have registered the baseline T1 image to the MNI T1 template (Grabner et al., 2006), included with FSL. As an individual’s brain does not necessarily have the same size as the template, it is not appropriate to use rigid-body transformations. Instead, non-linear transformations are needed. As these are patients with MS and have lesions, different non-linear registrations to template space can have a large impact on the outcome of analysis (see Eloyan et al. (2014) for discussion).

We will first register the baseline T1 image to the T1 template using an affine registration, which can perform scaling and shearing operations in addition to translation and rotation. Although an affine transformation has more

degrees of freedom than a rigid transformation, it may not provide a registration sufficient for analysis. We will then use FNIRT (FMRIB’s Nonlinear Image Registration Tool) to achieve better overlap of local brain structures (Jenkinson et al., 2012; Andersson et al., 2007). As we are concerned with good overlap only in brain structures, and not in areas such as the skull, we will register the brain-extracted brain images to the brain-only template. The **fslr** function **fnirt_with_affine** will register using **flirt** with an affine transformation and then non-linearly register this image to the template using **fnirt**. If this affine transformation is not executed before **fnirt**, the image will not achieve the necessary scaling into the template space.

```
fnirt_with_affine(infile = "01-Baseline_T1_FSL_BiasCorrect_Brain",
                  reffile = file.path(fsldir(), "data",
                                      "standard", "MNI152_T1_1mm_brain"),
                  flirt.omat =
                      "01-Baseline_T1_FSL_BiasCorrect_Brain_affine_toMNI.mat",
                  flirt.outfile =
                      "01-Baseline_T1_FSL_BiasCorrect_Brain_affine_toMNI",
                  outfile = "01-Baseline_T1_FSL_BiasCorrect_Brain_toMNI")
```

The results of the registration can be seen in Figure 3.7. Each panel represents a different axial slice ($z = 25, 45, 92$, or 137) in the template space of the template image (A, C, E, G) or the registered T1 image (B, D, F, H). Each slice shows the registered T1 image has similar brain structure represented in the same area as the template image, indicating good registration.

3.6.1 Applying transformations to co-registered data

Since all the data is represented in the same image space, we can apply the estimated affine transformation and non-linear warping coefficient field to

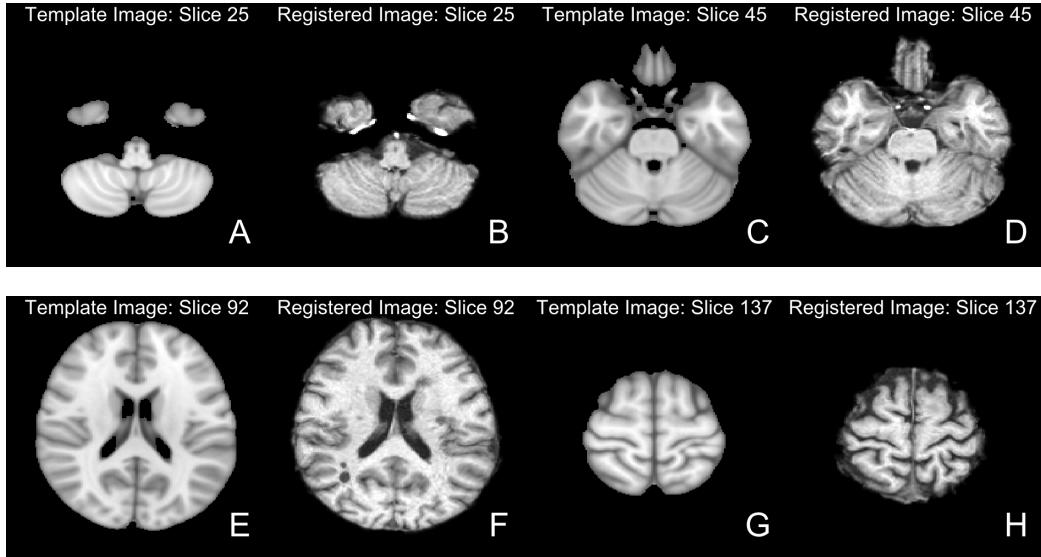


Figure 3.7: Results from FNIRT. We present different axial slices of the template (**A**, **C**, **E**, **G**) and the registered T1 image (**B**, **D**, **F**, **H**). The slices represented are 25 (**A**, **B**), 45 (**C**, **D**), 92 (**E**, **F**) and 137 (**G**, **H**). We note that areas of the brain coincide between the template and registered image.

each image to represent that image in template space. The affine transformation must be applied with `flirt_apply` and the warping coefficient using `fsl_applywarp`, which calls the FSL `applywarp` function.

Here we present the application of the transformations to the baseline T2 image, previously registered to the baseline T1.

```
flirt_apply(infile = "01-Baseline_T2_FSL_BiasCorrect_rigid_to_T1_Brain",
           reffile = file.path(fsldir(), "data", "standard", "MNI152_T1_1mm_brain"),
           initmat = "01-Baseline_T1_FSL_BiasCorrect_Brain_affine_toMNI.mat",
           outfile = "01-Baseline_T2_FSL_BiasCorrect_rigid_to_T1_Brain_toMNI")
fsl_applywarp(infile = "01-Baseline_T2_FSL_BiasCorrect_rigid_to_T1_Brain_toMNI",
              reffile = file.path(fsldir(), "data", "standard", "MNI152_T1_1mm_brain"),
              warpfile = "01-Baseline_T1_FSL_BiasCorrect_Brain_affine_toMNI_warpcoef",
              outfile = "01-Baseline_T2_FSL_BiasCorrect_rigid_to_T1_Brain_toMNI")
```

These two operations can also be performed in a single call to the `fslr`

`fnirt_with_affine_apply` function.

With multiple participants, this process yields a multi-person, multi-modal, longitudinal imaging dataset that can be used for analyses.

3.7 Conclusion

The neuroimaging community has developed a large collection of tools for image processing and analysis. R has a number of packages to perform operations on images; [EBImage](#) is one good example (Pau et al., 2010). Much of the fundamental functionality of neuroimage processing is not currently available in R, such as brain extraction and tissue segmentation. We present **fslr** to provide R users functions for image processing and analysis that are based on FSL, an established image processing and analysis software suite. Interfacing R with existing, powerful software provides users with thoroughly-tested software and an additional community of users, which would not be available if the functions were rewritten in R. **fslr** should be easy to use for any standard R user; the workflow allows R users to manipulate array-like `nifti` objects, pass them to **fslr** functions, which return `nifti` objects. Moreover, as FSL and R are open source and free, this software is readily available to all users.

There has been an increasing popularity of similar interfacing of tools within the Python community such as Nipype (Gorgolewski et al., 2011) (<https://qa.debian.org/popcon.php?package=nipype>). As many users of R may not have experience with Python or bash scripting, we believe **fslr** provides a lower threshold for use in the R community. Other packages

provide R users additional neuroimaging processing functionality such as **AnalyzeFMRI** (Bordier et al., 2011), **RNiftyReg** (Clayden, 2015), and **fmri** (Tabelow and Polzehl, 2011).

For example, other inhomogeneity correction methods exist, such as the popular N3 (Sled, Zijdenbos, and Evans, 1998) and N4 (Tustison et al., 2010), methods which are not implemented in **fslr**. **ANTsR** (<http://stnava.github.io/ANTsR/index.html>) is a currently unpublished R package that interfaces with the ANTs (advanced normalization tools) software suite (Avants et al., 2011). ANTs has implementations of these correction methods, an increased set of registration techniques, and other methods for image processing. Other packages such as this, along with **fslr**, can create a diverse set of tools for neuroimaging within R, building on preexisting and widely-accepted software.

Most importantly, as **fslr** is based on the R framework, all the benefits of using R are available, such as dynamic documents, reproducible reports, customized figures, and state-of-the-art statistical methods. These benefits provide unique functionality compared to other software packages for neuroimaging.

All data and code processed here is located at https://github.com/muschellij2/FSLR_Data.

3.8 Supplemental material

3.8.1 Smoothing images

Let us show how to pass a Gaussian smoother over an image using `fslsmooth` (FSL `fslmaths -s` function). First we will read in the registered-to-template baseline T1 brain image:

```
t1_to_temp = readNIfTI("01-Baseline_T1_FSL_BiasCorrect_Brain_toMNI", reorient=FALSE)
```

We will smooth the image using a Gaussian smoother with $\sigma = 3\text{mm}^3$:

```
smooth = fslsmooth(t1_to_temp, sigma = 3, retimg=TRUE)
```

The result is presented in Figure 3.8A.

3.8.2 Thresholding images

The `fslr fslbin` function will binarize a `nifti` image object: all values ≤ 0 are set to 0, and set to 1 otherwise. Let us binarize the registered image:

```
binned = fslbin(t1_to_temp, retimg=TRUE)
```

The result is presented in Figure 3.8B.

The `fslr fslthresh` function provides more control for thresholding by setting a lower threshold (`thresh` argument) and upper threshold (`uthresh` argument). These thresholds are inclusive, and will set values less than (but not equal to) `thresh` or greater than (but not equal to) `uthresh` to 0. Voxels with values between `thresh` and `uthresh` (inclusive) will be returned as their original value. Let us threshold the smoothed image between 30 and 50:

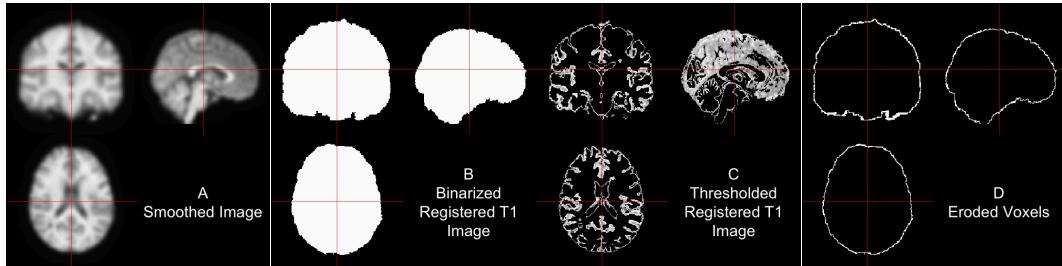


Figure 3.8: Results of fslr functions. We present the smoothed registered-to-template T1 image (A), binarized image (B), thresholded image (C), and image of the eroded voxels after eroding the binarized image (D).

```
thresh = fslthresh(t1_to_temp, thresh = 30, uthresh = 50, retimg=TRUE)
```

The result is presented in Figure 3.8C.

3.8.3 Eroding and dilating images

In many applications, one wants to erode (i.e. shrink) an image mask. The **fslr** function **fslerode** performs this operations. Note, if no options are specified for the kernel (in the **kopts** argument), the default $3 \times 3 \times 3$ voxel box kernel is used. Here we erode the binarized image from above and plot the voxels eroded:

```
eroded = fslerode(binned, retimg=TRUE)
```

The result is presented in Figure 3.8D. If one inverts the binary mask, performs erosion, and then inverts the resulting erosion mask, this procedure is equivalent to dilation. The **fslr** function **fsldilate** (version 1.4 and above) will also perform image dilation.

3.8.4 Extracting image information

Although most **fslr** functions provide an image as the designated output, one may wish to extract image information from a NIfTI file on disk, without reading it into R. The **fslr** `fslstats`, `fslval`, and `fslhd` functions are particularly useful and call functions of the same name from FSL.

For example, we can extract the number of slices in of the third dimension of the bias-corrected T1 image using `fslval`:

```
fslval("01-Baseline_T1_FSL_BiasCorrect_Brain", keyword = "dim3")  
[1] "124"
```

We could also extract the entire header using `fslhd` and assign it to an object:

```
img_hdr = fslhd("01-Baseline_T1_FSL_BiasCorrect_Brain")
```

We can extract the mean of the image of non-zero voxels:

```
fslstats("01-Baseline_T1_FSL_BiasCorrect_Brain", opts = "-M")  
[1] "51.264791"
```

Overall, there are many functions and options that allow the user to compute statistics or obtain header information from an image on disk without having to load it into R, which can reduce computation time.

3.8.5 Additional **fslr** functionality

Although the main goal of **fslr** is to interface R and FSL, there is a set of functions in **fslr** that are not designed to interface with FSL, but rather provide

helper functions for `nifti` objects from the **oro.nifti** package. We will display 2 example functions: `cal_img` and `niftiarr`. The `cal_img` function resets the `cal_min` and `cal_max` slots on a `nifti` object, which are used to determine colors when plotting. The `niftiarr` function copies a `nifti` object and replaces the `.Data` slot, which contains the actual image intensity values, with a provided array.

Let us illustrate by discussing 2 ways to mask an image. We will read in the bias-corrected T1 image and the mask from BET:

```
base_t1 = readNIfTI("01-Baseline_T1_FSL_BiasCorrect", reorient=FALSE)
base_t1_mask = readNIfTI("01-Baseline_T1_FSL_BiasCorrect_Brain_mask", reorient=FALSE)
```

One way to mask the T1 image is to multiply the image by the binary mask:

```
base_t1_1 = base_t1 * base_t1_mask
class(base_t1_1)

[1] "array"
```

We see that the resulting object is an array and not a `nifti` object (as of **oro.nifti** version 0.4.3). This may be a problem when trying to plot or manipulate this object using methods for `nifti` objects. To address this problem, the `niftiarr` function in **fslr** inputs a `nifti` object and an array, and returns a `nifti` object with the provided array in the `.Data` slot, copying over the image information from the input `nifti` object. As of **oro.nifti** version 0.5.0, the output from above will be a `nifti` object, but the function explained below, `niftiarr`, is still of use in cases when creating a new `nifti` object.

```
base_t1_1 = niftiarr(base_t1, base_t1_1)
class(base_t1_1)
```

```
[1] "nifti"
attr(,"package")
[1] "oro.nifti"
```

Another way of masking the image is to subset the values of the image that are not in the mask and setting those values to 0 (or some other value).

```
base_t1_2 = base_t1
base_t1_2[base_t1_mask == 0] = 0
class(base_t1_2)
```

```
[1] "nifti"
attr(,"package")
[1] "oro.nifti"
```

We see that this correctly returns an object of class `nifti`. One problem is that we have changed the data in the `nifti` object `base_t1_2` but did not reset the other slots in this object to reflect this change.

In a `nifti` object, the `cal_min` and `cal_max` slots equal the minimum and maximum values, respectively, of the data. The `orthographic` function (from `oro.nifti`) uses these values for plotting; also, if these slots do not equal the minimum and maximum, the `writeNIFTI` function (from `oro.nifti`) will fail. The `cal_img` is a simple helper function that will set the `cal_min` and `cal_max` slots to the correct values. Let us look at the range of the data and the `cal_min` and `cal_max` slots:

```
range(base_t1_2)

[1] 0.0000 409.3908

c(base_t1_2@cal_min, base_t1_2@cal_max)

[1] 0 0
```

An issue with `readNIfTI` function from `oro.nifti` is that the `cal_min` and `cal_max` slots may be both read as zero. Let us set these to the range using the `cal_img` command from `fslr`:

```
base_t1_2 = cal_img(base_t1_2)
c(base_t1_2@cal_min, base_t1_2@cal_max)
```

```
[1] 0.0000 409.3908
```

We see that after these operations are done 2 in different ways, the resulting `nifti` objects are equivalent.

```
identical(base_t1_1, base_t1_2)

[1] TRUE
```

Additional helper functions such as these are included in `fslr` for plotting and image manipulation.

References

- Jenkinson, Mark, Christian F. Beckmann, Timothy E. J. Behrens, Mark W. Woolrich, and Stephen M. Smith (2012). "FSL". In: *NeuroImage* 62.2, pp. 782–790. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2011.09.015](https://doi.org/10.1016/j.neuroimage.2011.09.015). URL: <http://www.sciencedirect.com/science/article/pii/S1053811911010603> (visited on 01/15/2014).
- Smith, Stephen M (2002). "Fast robust automated brain extraction". eng. In: *Human brain mapping* 17.3, pp. 143–155. ISSN: 1065-9471. DOI: [10.1002/hbm.10062](https://doi.org/10.1002/hbm.10062).
- Zhang, Yongyue, Michael Brady, and Stephen Smith (2001). "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm". In: *Medical Imaging, IEEE Transactions on* 20.1, pp. 45–57. URL: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=906424 (visited on 09/09/2014).
- Jenkinson, Mark and Stephen Smith (2001). "A global optimisation method for robust affine registration of brain images". In: *Medical image analysis* 5.2, pp. 143–156. URL: <http://www.sciencedirect.com/science/article/pii/S1361841501000366> (visited on 09/09/2014).
- Jenkinson, Mark, Peter Bannister, Michael Brady, and Stephen Smith (2002). "Improved optimization for the robust and accurate linear registration and motion correction of brain images". In: *Neuroimage* 17.2, pp. 825–841. URL: <http://www.sciencedirect.com/science/article/pii/S1053811902911328> (visited on 09/09/2014).
- Carp, Joshua (2012). "The secret lives of experiments: Methods reporting in the fMRI literature". In: *NeuroImage* 63.1, pp. 289–300. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2012.07.004](https://doi.org/10.1016/j.neuroimage.2012.07.004). URL: <http://www.sciencedirect.com/science/article/pii/S1053811912007057> (visited on 11/11/2014).
- Bordier, Cécile, Michel Dojat, Pierre Lafaye de Micheaux, and others (2011). "Temporal and spatial independent component analysis for fMRI data sets embedded in the AnalyzeFMRI R package". In: *Journal of Statistical Software*

- 44.9, pp. 1–24. URL: <http://www.hal.inserm.fr/inserm-00659425/> (visited on 09/09/2014).
- Clayden, Jon (2015). *RNiftyReg: Medical Image Registration Using the NiftyReg Library*, Jon Clayden and based on original code by Marc Modat and Pankaj Daga. URL: <http://CRAN.R-project.org/package=RNiftyReg>.
- Tabelow, Karsten and Jörg Polzehl (2011). “Statistical parametric maps for functional MRI experiments in R: The package fmri”. In: *Journal of Statistical Software* 44.11, pp. 1–21. URL: <http://www.jstatsoft.org/v44/i11/paper> (visited on 09/09/2014).
- Whitcher, Brandon, Volker J. Schmid, and Andrew Thornton (2011). “Working with the DICOM and NIfTI Data Standards in R”. In: *Journal of Statistical Software* 44.6, pp. 1–28. URL: <http://www.jstatsoft.org/v44/i06/paper> (visited on 09/12/2014).
- Sweeney, E. M., R. T. Shinohara, C. D. Shea, D. S. Reich, and C. M. Crainiceanu (2013). “Automatic Lesion Incidence Estimation and Detection in Multiple Sclerosis Using Multisequence Longitudinal MRI”. en. In: *American Journal of Neuroradiology* 34.1, pp. 68–73. ISSN: 0195-6108, 1936-959X. DOI: [10.3174/ajnr.A3172](https://doi.org/10.3174/ajnr.A3172). URL: <http://www.ajnr.org/content/34/1/68> (visited on 11/14/2014).
- Guillemaud, Régis and Michael Brady (1997). “Estimating the bias field of MR images”. In: *Medical Imaging, IEEE Transactions on* 16.3, pp. 238–251. URL: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=585758 (visited on 09/11/2014).
- Wickham, Hadley (2009). *ggplot2: elegant graphics for data analysis*. Springer. URL: http://books.google.com/books?hl=en&lr=&id=bes-AAAAQBAJ&oi=fnd&pg=PR5&dq=ggplot2:+elegant+graphics+for+data+analysis&ots=SA8_QA6N00&sig=iAb1q3tX5ZkAh01UH12uPiySFdI (visited on 09/12/2014).
- Hastie, Trevor J. and Robert J. Tibshirani (1990). *Generalized additive models*. Vol. 43. CRC Press. URL: http://books.google.com/books?hl=en&lr=&id=qa29r1Ze1coC&oi=fnd&pg=PR13&dq=generalized+additive+models&ots=j3YSfqBWoR&sig=9WE9tUp7uoXJtjwCmcSLtvSr_g (visited on 11/04/2014).
- Wood, Simon N. (2011). “Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73.1, pp. 3–36. URL: <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9868.2010.00749.x/full> (visited on 11/04/2014).

- Jenkinson, Mark, Mickael Pechaud, and Stephen Smith (2005). "BET2: MR-based estimation of brain, skull and scalp surfaces". In: *Eleventh annual meeting of the organization for human brain mapping*. Vol. 17. URL: <http://mickaelpechaud.free.fr/these/HBM05.pdf> (visited on 09/12/2014).
- Popescu, V., M. Battaglini, W. S. Hoogstrate, S. C. J. Verfaillie, I. C. Sluimer, R. A. van Schijndel, B. W. van Dijk, K. S. Cover, D. L. Knol, M. Jenkinson, F. Barkhof, N. de Stefano, and H. Vrenken (2012). "Optimizing parameter choice for FSL-Brain Extraction Tool (BET) on 3D T1 images in multiple sclerosis". In: *NeuroImage* 61.4, pp. 1484–1494. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2012.03.074](https://doi.org/10.1016/j.neuroimage.2012.03.074). URL: <http://www.sciencedirect.com/science/article/pii/S1053811912003552> (visited on 09/12/2014).
- Cox, Robert W. (1996). "AFNI: software for analysis and visualization of functional magnetic resonance neuroimages". In: *Computers and Biomedical research* 29.3, pp. 162–173. URL: <http://www.sciencedirect.com/science/article/pii/S0010480996900142> (visited on 03/30/2015).
- Ashburner, John and Karl J. Friston (2005). "Unified segmentation". In: *Neuroimage* 26.3, pp. 839–851. URL: <http://www.sciencedirect.com/science/article/pii/S1053811905001102> (visited on 03/30/2015).
- Fischl, Bruce (2012). "FreeSurfer". In: *Neuroimage* 62.2, pp. 774–781. URL: <http://www.sciencedirect.com/science/article/pii/S1053811912000389> (visited on 03/30/2015).
- Doshi, Jimit, Guray Erus, Yangming Ou, Bilwaj Gaonkar, and Christos Davatzikos (2013). "Multi-Atlas Skull-Stripping". In: *Academic Radiology* 20.12, pp. 1566–1576. ISSN: 1076-6332. DOI: [10.1016/j.acra.2013.09.010](https://doi.org/10.1016/j.acra.2013.09.010). URL: <http://www.sciencedirect.com/science/article/pii/S1076633213004182> (visited on 03/30/2015).
- Grabner, Günther, Andrew L. Janke, Marc M. Budge, David Smith, Jens Pruessner, and D. Louis Collins (2006). "Symmetric Atlasing and Model Based Segmentation: An Application to the Hippocampus in Older Adults". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006*. Ed. by David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Dough Tygar, Moshe Y. Vardi, Gerhard Weikum, Rasmus Larsen, Mads Nielsen, and Jon Sporring. Vol. 4191. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 58–66. ISBN: 978-3-540-44727-6, 978-3-540-44728-3. URL: http://link.springer.com/10.1007/11866763_8 (visited on 09/22/2014).

- Eloyan, Ani, Haochang Shou, Russell T. Shinohara, Elizabeth M. Sweeney, Mary Beth Nebel, Jennifer L. Cuzzocreo, Peter A. Calabresi, Daniel S. Reich, Martin A. Lindquist, and Ciprian M. Crainiceanu (2014). "Health Effects of Lesion Localization in Multiple Sclerosis: Spatial Registration and Confounding Adjustment". In: *PLoS ONE* 9.9, e107263. DOI: [10.1371/journal.pone.0107263](https://doi.org/10.1371/journal.pone.0107263). URL: <http://dx.doi.org/10.1371/journal.pone.0107263> (visited on 11/14/2014).
- Andersson, Jesper LR, Mark Jenkinson, Stephen Smith, and others (2007). "Non-linear registration, aka Spatial normalisation FMRIB technical report TR07JA2". In: *FMRIB Analysis Group of the University of Oxford*. URL: <http://fmrifl.medsci.ox.ac.uk/analysis/techrep/tr07ja2/tr07ja2.pdf> (visited on 11/04/2014).
- Pau, Gregoire, Florian Fuchs, Oleg Sklyar, Michael Boutros, and Wolfgang Huber (2010). "EBImage—an R package for image processing with applications to cellular phenotypes". In: *Bioinformatics* 26.7, pp. 979–981. DOI: [10.1093/bioinformatics/btq046](https://doi.org/10.1093/bioinformatics/btq046).
- Gorgolewski, Krzysztof, Christopher D. Burns, Cindee Madison, Dav Clark, Yaroslav O. Halchenko, Michael L. Waskom, and Satrajit S. Ghosh (2011). "Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python". In: *Frontiers in neuroinformatics* 5. URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/pmc3159964/> (visited on 09/11/2014).
- Sled, John G., Alex P. Zijdenbos, and Alan C. Evans (1998). "A nonparametric method for automatic correction of intensity nonuniformity in MRI data". In: *Medical Imaging, IEEE Transactions on* 17.1, pp. 87–97. URL: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=668698 (visited on 09/18/2014).
- Tustison, Nicholas J., Brian B. Avants, Philip A. Cook, Yuanjie Zheng, Alexander Egan, Paul A. Yushkevich, and James C. Gee (2010). "N4ITK: improved N3 bias correction". In: *Medical Imaging, IEEE Transactions on* 29.6, pp. 1310–1320. URL: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5445030 (visited on 11/04/2014).
- Avants, Brian B., Nicholas J. Tustison, Gang Song, Philip A. Cook, Arno Klein, and James C. Gee (2011). "A reproducible evaluation of ANTs similarity metric performance in brain image registration". In: *NeuroImage* 54.3, pp. 2033–2044. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2010.09.025](https://doi.org/10.1016/j.neuroimage.2010.09.025). URL: <http://www.sciencedirect.com/science/article/pii/S1053811910012061> (visited on 11/04/2014).

Chapter 4

PItCHPERFeCT: Primary Intracranial Hemorrhage Probability Estimation using Random Forests on CT

4.1 Introduction

Intracerebral hemorrhage (ICH) is a neurological condition that results from a blood vessel rupturing into the tissue and possibly the ventricles of the brain. The use of X-ray computed tomography (CT) scans allows clinicians and researchers to qualitatively and quantitatively describe the characteristics of a hemorrhage to guide interventions and treatments. CT scanning is widely available and is the most commonly used diagnostic tool in patients with ICH (Sahni and Weinberger, 2007). The volume of ICH has been consistently demonstrated to be an important diagnostic predictor of stroke severity, long-term functional outcome, and mortality (Broderick et al., 1993; Hemphill et al., 2001; Tuhrim et al., 1999). ICH volume change is also common primary outcome (Anderson et al., 2008; Anderson et al., 2010; Qureshi et al., 2011; Mayer et al., 2005) and secondary outcome (Morgan et al., 2008a; Anderson et al., 2008; Morgan et al., 2008b) in clinical trials. Moreover, the location of the ICH has been shown to affect functional outcome in patients with stroke (Rost et al., 2008; Castellanos et al., 2005).

ICH volume can be rapidly measured using techniques such as the ABC/2 method (Broderick et al., 1993). In this method, a reader chooses which slice has the largest area of hemorrhage, draws a line along the longest axis of the hemorrhage (denoted A) and the orthogonal line that bisects the hemorrhage (B). The reader then counts the number of slices where hemorrhage is present (C). The volume estimate is $\frac{A \times B \times C}{2}$, which is an approximation of an ellipsoid (Kothari et al., 1996). As this method only requires 3 measurements, this method can be done rapidly.

Although ABC/2 is widely used, Divani et al. (2011) found that volume measurement errors using ABC/2 were significantly greater than those using planimetry measurements at measuring the true volume of a hemorrhage, especially for irregularly shaped ICH and for smaller thickness (i.e. higher resolution) scans. Recently, Webb et al. (2015) found that ABC/2 measurements at a clinical site, 81% of the 4,369 scans were within 5 milliliters (mL) of ICH volume compared to planimetry methods, but only 41% were within 20%. Moreover, ICH may initially have a regular shape where ABC/2 performs well, but many surgical intervention and procedure targets the removal of ICH, which changes its shape or cause re-bleeding and additional ICH. ABC/2 does not perform well in these cases. Moreover, ABC/2 also does not take into account any intraventricular hemorrhage (IVH) present within the image, which has been shown to be prognostic of 30-day mortality (Hemphill et al., 2001; Tuhrim et al., 1999). ABC/2 also been shown to consistently over-estimate infarct volume (Pedraza et al., 2012), and can have significant inter-rater variability (Hussein et al., 2013). Therefore, we believe a rapid, automated method for estimating hemorrhage from CT scans has diagnostic and prognostic value.

Methods have been proposed for segmentation of ICH on magnetic resonance images (MRI) (Wang et al., 2013; Carhuapoma et al., 2003). In our current study and most clinical settings, protocol are not standardized for MRI scans as these are not uniform across clinical sites for which sequences are taken and which scans are standard of care for ICH. Therefore, we required a method that relied solely on CT scan information.

We wish to create an algorithm that can estimate the probability of ICH at a voxel-level, a binary ICH image, and the volume of ICH. We will compare our predicted ICH maps to the gold standard – manual segmentation. Other methods have been presented for automated methods for estimating ICH from CT scans (Gillebert, Humphreys, and Mantini, 2014; Prakash et al., 2012; Loncaric, Cosic, and Dhawan, 1996; Loncaric et al., 1999; Pérez et al., 2007). These methods include fuzzy clustering (Prakash et al., 2012; Loncaric, Cosic, and Dhawan, 1996), simulated annealing (Loncaric et al., 1999), 3-dimensional (3D) mathematical morphology operations (Pérez et al., 2007), and template-based comparisons (Gillebert, Humphreys, and Mantini, 2014). With the exception of Gillebert, Humphreys, and Mantini (2014) (with request), no software for ICH segmentation is publicly available. In addition, this software requires varied level of technical knowledge. Therefore, we provide a completely automated pipeline of analysis from raw images to binary hemorrhage masks and volume estimates, and provide a webpage to test the software.

4.2 Methods

4.2.1 Data

4.2.2 Participants and Imaging Data

We used CT images from patients enrolled in the MISTIE (Minimally Invasive Surgery plus recombinant-tissue plasminogen activator for Intracerebral Evacuation) stroke trial (Morgan et al., 2008a). We analyzed 112 scans taken prior to randomization and treatment, corresponding to the first scan acquired

post-stroke for 112 unique patients. Inclusion criteria into the study included: 18 to 80 years of age, spontaneous supratentorial intracerebral hemorrhage above 20 milliliters (mL) in size (for full criteria, see Mould et al. (2013)). The population analyzed here had a mean (SD) age was 60.7 (11.2) years, was 68.8% male, and was 53.6% Caucasian, 31.2% African American, 10.7% Hispanic, and 4.5% Asian or Pacific islander. CT data were collected as part of the Johns Hopkins Medicine IRB-approved MISTIE research studies with written consent from participants.

The study protocol was executed with minor, but important, differences across the 26 sites. Scans were acquired using 4 scanner manufacturers: GE ($N = 46$), Siemens ($N = 38$), Philips ($N = 20$), and Toshiba ($N = 8$). In head CT scanning, the gantry may be tilted so that sensitive organs, such as the eyes, are not exposed to X-ray radiation. This causes scan slices to be acquired at an oblique angle with respect to the patient. This gantry tilt was observed in 88 scans. Slice thickness of the image varied within the scan for 14 scans. For example, a scan may have 10 millimeter (mm) slices at the top and bottom of the brain and 5mm slices in the middle of the brain. Therefore, the original scans analyzed had different voxel (volume element) dimensions. These conditions represent how scans are presented for evaluation in many diagnostic cases.

4.2.3 Hemorrhage Segmentation and Location Identification

ICH was manually segmented on CT scans using the OsiriX imaging software by expert readers (OsiriX v. 4.1, Pixmeo; Geneva, Switzerland). Readers

employed a semiautomated threshold-based approach using a Hounsfield unit (HU) range of 40 to 80 to select potential regions of ICH (Bergström et al., 1977; Smith, Rosand, and Greenberg, 2006); these regions were then further quality controlled and refined by readers using direct inspection of images. Binary hemorrhage masks were created by setting voxel intensity to 1 if the voxel was classified as hemorrhage, regardless of location, and 0 otherwise.

4.2.4 Image Processing: Brain Extraction, Registration

CT images and binary hemorrhage masks were exported from OsiriX to DICOM (Digital Imaging and Communications in Medicine) format. The image processing pipeline can be seen in Figure 6.1. 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` (provided with MRIcro (Rorden and Brett, 2000)). 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 converted slice and the NIfTI format assumes homogeneous thickness throughout the image.

All image analysis was done in the R statistical software (R Core Team, 2015), using the `fslr` (Muschelli et al., 2015a) package to call functions from the FSL (Jenkinson et al., 2012) neuroimaging software (version 5.0.4), and the `ANTsR` package to call functions from the ANTs (Advanced Normalization Tools) neuroimaging software (Avants et al., 2011a).

Brains were extracted to remove skull, eyes, facial and nasal features, extracranial skin, and more importantly non-human elements of the image captured by the CT scanner, such as the gantry, pillows, or medical devices. Removal of these elements was performed using the brain extraction tool (BET) (Smith, 2002), a function of FSL, using a previously published validated CT-specific brain extraction protocol (Muschelli et al., 2015c).

4.2.5 Image Registration

Rorden et al. (2012) introduced a CT template based on 35 individuals who presented with specific neurological deficits that were suspected to be caused by a stroke, but were later found to be due to a metabolic abnormality. This CT template is represented in MNI (Montreal Neurological Institute) space and brain-extraction was performed on the template. Prior to image processing, brain-extracted images were registered to this brain-extracted template using a rigid-body (6 degrees of freedom) and linearly interpolated to a $1 \times 1 \times 1$ mm voxel resolution. Transformed hemorrhage masks and brain masks were thresholded using a value of 0.5 to preserve mask volume *Frequently Asked Questions for FLIRT*. This operation reoriented the image, ensured isotropic voxel sizes for smoothing and other operations below, and preserved the relative volume of the ICH. All image preprocessing and analysis are done in MNI space, described as template space, unless otherwise specified.

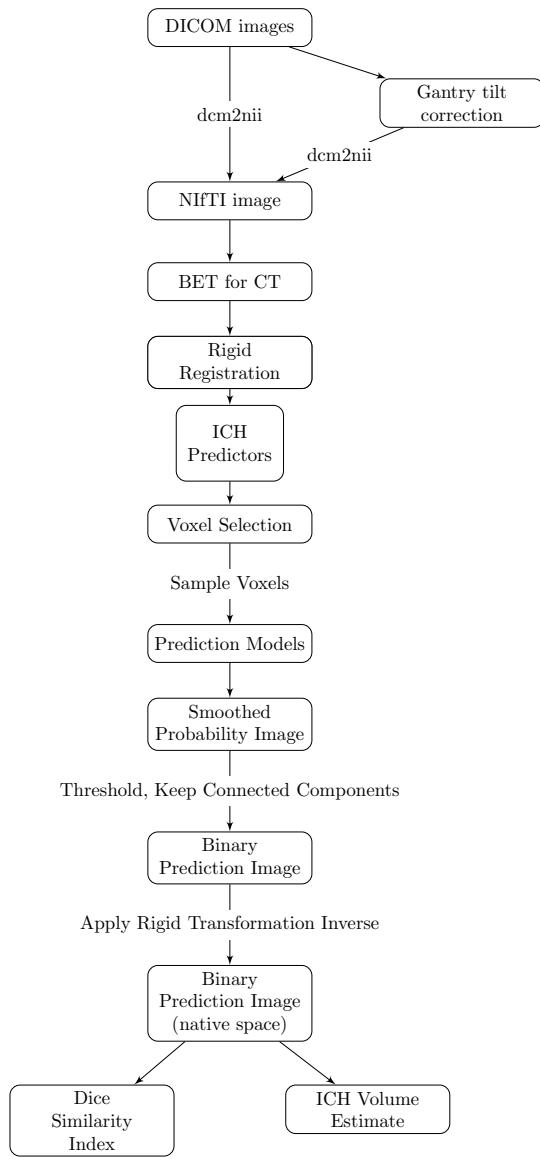


Figure 4.1: Processing Pipeline. Images in DICOM (Digital Imaging and Communications in Medicine) format were gantry tilt corrected if necessary and converted to NIFTI (Neuroimaging Informatics Technology Initiative) format using dcm2nii. After NIFTI conversion, the brain extraction tool (BET) was applied to the image using a previously published protocol. The image was rigidly registered to a brain CT template. We estimated imaging predictors and used these predictors to estimate the probability of ICH in a prediction model. The probability of ICH was thresholded, connected component below 100 voxels (0.1mL) were discarded, and the image was transformed back into original space of the patient. The ICH volume and the Dice Similarity Index, an overlap measure, were calculated compared to the true estimate from the manual segmentation.

4.2.6 Brain Mask Erosion

Each brain mask was eroded by a box kernel ($3 \times 3 \times 1\text{mm}$). Though this erosion may exclude voxels from superficial bleeds towards the cortical surface, it excludes voxels with similar ranges as ICH voxels, caused by 1) incomplete skull stripping or 2) partial voluming effects with the skull. If any voxels from the hemorrhage mask was removed due to brain extraction or brain mask erosion, these voxels were included in estimating model performance but their predicted probability of ICH was set to 0. Therefore, these deleted ICH voxels will always be incorrectly predicted as not ICH.

4.2.7 Imaging Predictors

We derived a set of imaging predictors from each CT scan. We will describe each here with their rationale for use. These features make up the potential set of predictors for image segmentation.

4.2.7.1 CT voxel intensity information

The raw voxel intensity value in HU was included $x(v)$, as it is the main predictor used in visual inspection; high HU values are indicative of hemorrhage. We created an indicator if the value was greater than 40 and less than 80 HU, similar to the criteria used for manual segmentation, denoted $I_{\text{thresh}}(v)$.

4.2.7.2 Local Moment Information

For each voxel, we extracted a neighborhood, denoted $N(v)$, of all adjacent neighboring voxels in 3 dimensions and the voxel itself. Let $x_k(v)$ denote the

voxel intensity in HU for voxel neighbor k , where $k = 1, \dots, 27$. We created the voxel neighborhood mean intensity ($\bar{x}(v)$):

$$\bar{x}(v) = \frac{1}{N(v)} \sum_{k \in N(v)} x_k(v) \quad (4.1)$$

We calculated the voxel neighborhood standard deviation (SD), skew, and kurtosis using the following method of moments estimators:

$$\begin{aligned} \text{SD}(v) &= \sqrt{\frac{1}{N(v)} \sum_{k \in N(v)} (x_k(v) - \bar{x}(v))^2} \\ \text{Skew}(v) &= \frac{\frac{1}{N(v)} \sum_{k \in N(v)} (x_k(v) - \bar{x}(v))^3}{\left[\frac{1}{N(v)} \sum_{k \in N(v)} (x_k(v) - \bar{x}(v))^2 \right]^{3/2}} \\ \text{Kurtosis}(v) &= \frac{\frac{1}{N(v)} \sum_{k \in N(v)} (x_k(v) - \bar{x}(v))^4}{\left(\frac{1}{N(v)} \sum_{k \in N(v)} (x_k(v) - \bar{x}(v))^2 \right)^2} \end{aligned}$$

We acknowledge that we did not divide by $N(v) - 1$ for standard deviation and skew, nor did we subtract by 3 for kurtosis. As $N(v)$ should be the same per voxel, this should not affect the estimates for prediction: it will be accounted for in any generalized linear model by the intercept and scaling of the beta coefficients and a tree-based decision algorithm will simply have a different cutoff value for binning.

Voxels higher in their local mean correspond to voxels adjacent to higher HU voxels on average, which are more likely to be ICH. The higher order

moments can provide information about how homogeneous the intensities in the neighborhood are and where edges occur. We also calculate the percentage of voxels in each neighborhood ($p_{\text{thresh}}(v)$) that have HU values between 40 and 80:

$$p_{\text{thresh}}(v) = \frac{1}{N(v)} \sum_{k \in N(v)} I\{40 \leq x_k(v) \leq 80\} \quad (4.2)$$

which should be higher for ICH voxels as they are surrounded by high HU values.

Voxels that are on the surface or surrounded by non-brain tissue are less likely to be ICH. Voxels not within the eroded mask are set to 0; the following predictors assume that voxels close to many voxels with intensity 0 are less likely to be ICH. We calculated the percentage of voxels that have neighbors of value of 0:

$$p_0(v) = \frac{1}{N(v)} \sum_{k \in N(v)} I\{x_k(v) = 0\} \quad (4.3)$$

and an indicator of whether any voxels in the neighborhood had a value of 0:

$$\bar{I}_0(v) = I\{p_0(v) > 0\} \quad (4.4)$$

4.2.7.3 Within-plane Standard Scores

Some brain structures have high HU values but are not ICH, such as the falx cerebri, which lies largely on the mid-sagittal plane. Moreover, tissues in the top of the brain may have a higher average HU than those in the middle or bottom of the brain. Thus, if values are standardized within each plane (axial, sagittal, coronal), these standard-plane scores may distinguish high values within a plane regardless of a mean shift, which may indicate ICH voxels.

We created standard-plane scores for each voxel on a slice-based level for axial, sagittal, and coronal planes. For each plane $o \in \{\text{axial, sagittal, and coronal}\}$, we calculated the standard-plane score as follows:

$$z_o(v) = \frac{x(v) - \bar{x}(v, o)}{\sigma(v, o)} \quad (4.5)$$

where $\bar{x}(v, o)$ and $\sigma(v, o)$ denote the mean and standard deviation of plane o which contains voxel v , excluding voxels outside the brain mask. In addition to the standardized images within plane we created a Winsorized standardized image, using the Winsorized mean and standard deviation, with a 20% trimming of the distribution, which may allow for better standardization with artifacts of the distribution (such as gross hyperintensities).

4.2.7.4 First-pass Segmentation

We used a previously-published, open-source, general segmentation tool based on Markov random fields for image segmentation, called Atropos (Avants et al., 2011b). We used a 4-tissue class segmentation, and added the top probability values from the top 2 classes to combine them into one class. We used this probability image as a predictor, denoted by $\text{prior}(v)$. Although this tool has been shown to perform well in other studies for tissue-class segmentation, it did not perform adequately as a standalone segmentation tool for ICH in CT.

4.2.7.5 Contralateral Difference Images

As most hemorrhages are constrained to one side of the brain, the contralateral side should have lower HU values. For non-hemorrhage voxels, however, the

contralateral voxels would have similar HU values, and thus their difference would be small or negative. Thus, we right-left flipped the image, and took a difference image:

$$f(v) = x(v) - x(v^*) \quad (4.6)$$

where v^* is the voxel from the contralateral side.

4.2.7.6 Global Head Information

We used the distance to the brain centroid ($d(v)$) to potentially down-weight voxels that are very far from the brain center, which may be artifacts. We also created 3 images that were obtained by smoothing the original image using large Gaussian kernels ($\sigma = 5mm^3, 10mm^3, 20mm^3$), which can capture any potential homogeneity throughout the scan, denoted by $s_5(v)$, $s_{10}(v)$ and $s_{20}(v)$, respectively.

4.2.7.7 Standardized-to-template Intensity

Although standardizing voxels compared to within-scan measurements can detect anomalous tissue, one powerful tool is to use a measure how different a voxel is compared to that voxel in a person from a non-stroke population. We registered the brain-extracted image to the brain-extracted CT template using an affine transformation, followed by a non-linear transformation estimated using Symmetric Normalization (SyN) (Avants et al., 2008). From 30 CT images from non-stroke patients from Dr. Rorden (personal communication), we registered these non-stroke, brain-extracted scans to the CT template, and created a voxel-wise mean image M and voxel-wise standard deviation S

image in template space. For each scan in our study, we created a standardized voxel intensity with respect to this population (z_{template}) using the following equation:

$$z_{\text{template}}(v) = \frac{x(v) - M(v)}{S(v)}$$

This image is warped back into the rigid-body-registered space so that these standardized voxels are aligned with other predictors. This predictor is similar to that used in Gillebert, Humphreys, and Mantini (2014).

4.2.8 Voxel Selection Procedure

We chose 10 scans from 10 patients to perform exploratory data analysis, model fitting, and estimation of model cutoffs, denoted as training data. We used the 102 remaining scans as test data to estimate performance.

From the training data we estimated the 0.5% and 99.5% quantiles for all ICH voxels in the predictors. Voxels that were jointly within these quantiles for z_{axial} , z_{coronal} , p_{thresh} , and within HU between 30 and 100 were considered candidate voxels of ICH; voxels outside of the intersection of these ranges were given a 0 probability of ICH. These cutoffs empirically found to be robust in the test scans; they excluded a mean 63.6 (min: 37.1, max: 89.8) percentage of non-ICH voxels, while including a mean 97.9 (min: 91.6, max: 99.9) percentage of ICH voxels. This voxel selection procedure should improve the specificity of model predictions, and improve computational speed.

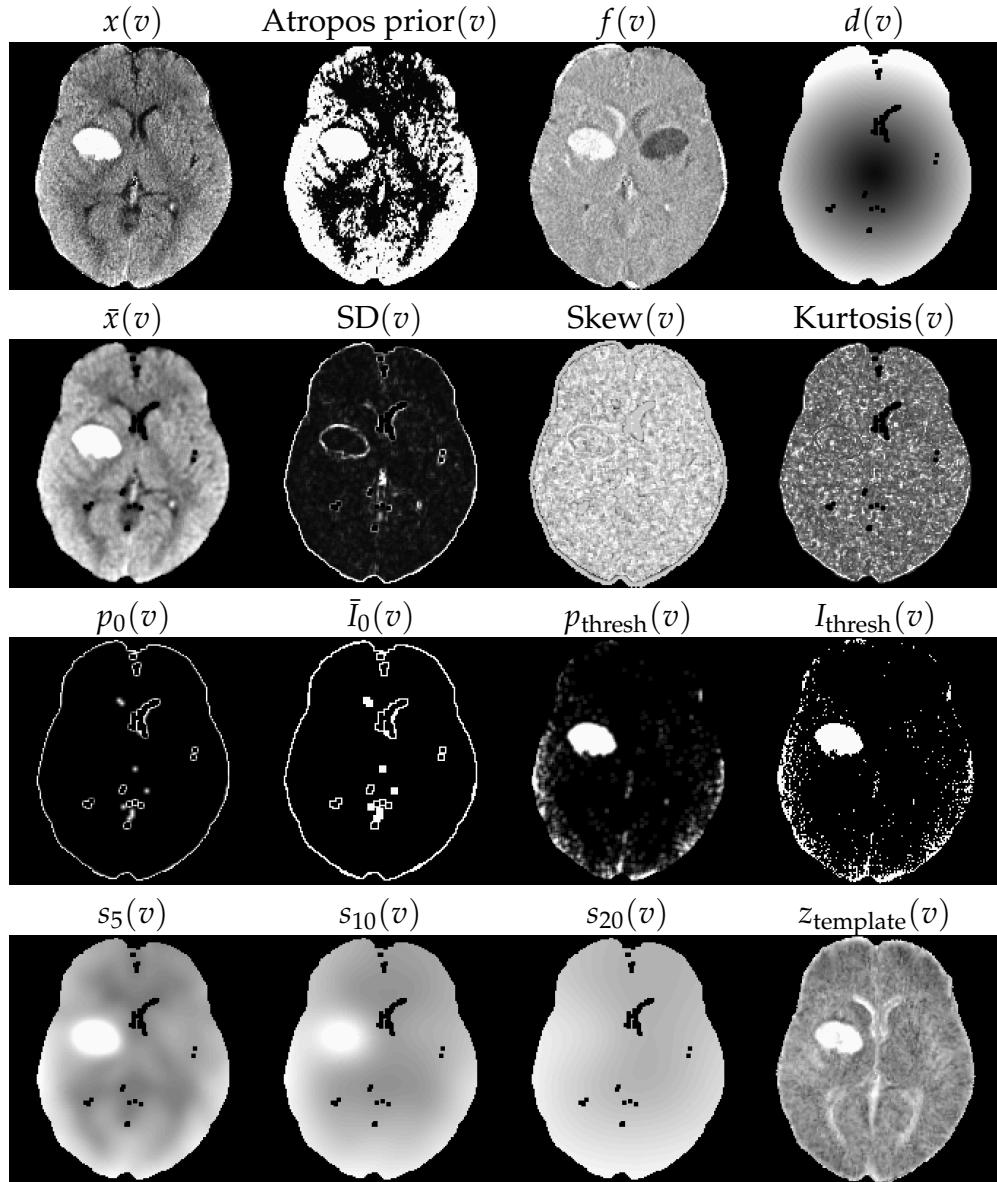


Figure 4.2: Predictor Images. Here we display one slice from one patient for predictor images. The within-plane and Winsorized predictor images were not shown as they are within-subject scaled version of the image $x(v)$ and would appear very similar. Although they appear similar at a subject level, the distribution of these predictors is different across patients. The images that visually separate the areas of ICH compared to the rest of the image are likely to be better predictors.

4.2.9 Models

From these 10 scans, we collapsed all the voxels passing the voxel selection procedure. From these voxels, we randomly sampled 100,000 to fit the models and do data exploration, and used the remaining voxels to estimate the model cutoffs of the probability of ICH. All models were fit with all predictors.

Using the sampled training voxels, we created predictions for the probability of ICH using a 1) logistic regression model, 2) logistic regression model with a penalty, 3) generalized additive model (GAM), and 4) classifier fit with the random forest algorithm.

For the standard and penalized logistic regression model, we used all predictors. The penalized model was fit using the LASSO (Least Absolute Shrinkage and Selection Operator) penalty (Tibshirani, 1996) using the **glmnet** package (Friedman, Hastie, and Tibshirani, 2010). The tuning parameter (λ) for the penalization was chosen using 10-fold cross-validation (of only the training voxels in the model), with the cost function of misclassification rate. The parameter was chosen using the largest value of λ such that the error is within 1 standard error of the minimum, chosen for out-of-sample performance stability.

The generalized additive model (GAM) (Hastie and Tibshirani, 1986; Hastie and Tibshirani, 1990) was also created using indicator variables for binary variables and thin-plate splines for all continuous measures, fit with fast-estimation of restricted maximum likelihood (fREML) using the **mgcv** package (Wood, 2011; Wood, Goude, and Shaw, 2015). Model specifications can be seen in Section 4.5.2.

We fit a random forest (Breiman, 2001) using the **randomForest** package in R (Liaw and Wiener, 2002), using the default pruning parameters and number of trees (`ntree=500, mtry=4`).

4.2.10 Estimating a Cutoff for Model Probability

Each model gives the estimate of the probability a voxel is ICH. In the end, however, we want a predicted binary hemorrhage mask to compare to the manual segmentation. Using the voxels from the training data, we estimated the probability of ICH from each model. For each model probability, we smoothed this probability image using the neighborhood voxels (1 voxel in every direction). To choose a probability cutoff to make a binary image, we used the voxels in the training data that were not sampled for estimating the model. Using these voxels in the smoothed images, we estimated the probability cutoff that minimized the Dice Similarity Index (DSI) (Dice, 1945) between the prediction and true value. The DSI is measure of overlap insensitive to values where neither the true segmentation nor predicted segmentation were considered ICH, and will be used as a performance measure when comparing models. DSI for scan i is calculated by

$$DSI_i = \frac{2 \times TP}{2 \times TP + FN + FP}$$

where TP denotes the number of “true positives”, where the manual and predicted segmentation are both 1, FP denotes the number of “false positives”, where the manual segmentation is 0 and predicted segmentation is 1, and FN denotes the number of “false negatives”, where the manual segmentation is 1

and predicted segmentation is 0. DSI ranges from 0 to 1; 0 indicates no overlap and 1 denotes perfect overlap.

After thresholding the smoothed image using this probability cutoff, we discarded regions with fewer than 100 (0.1mL) connected voxels. We then transformed this binary mask back to original (i.e. native) space using the inverse from the previously estimated rigid-body transformation. As the linear interpolation results in a non-binary mask, we thresholded this image at 0.5 to preserve volume *Frequently Asked Questions for FLIRT*.

For each scan in the test data, this voxel selection and prediction process was performed and each scan has a corresponding binary prediction image. For evaluation of all measures, the comparison was done in the native space of the patient, not in the template space, as this was where the manual segmentation was done.

4.2.11 Measuring and Testing ICH Prediction Performance

For all measures, we used the binary prediction masks from the test scans. We measured performance for each model using the DSI and also estimated the ICH volume of the prediction. The number of true negatives inflates specificity and overall accuracy, since most of the voxels within the brain are non-ICH, and are not reported. We tested the difference between median DSI across models using a Kruskal-Wallis test. If a difference was present, we tested each combination of models (6 combinations) using a Wilcoxon signed rank test, and corrected the p-value using a Bonferroni correction.

We also calculated the relationship of estimated volume of ICH compared

to the manual segmentation using the Pearson correlation and root mean squared (RMSE) between volumes measures. Similarly, we did performed a Kruskal-Wallis test on the medians of the absolute value of the difference in estimated volume from each model and the true volume and performed pairwise tests if necessary. For DSI and correlation, higher values indicate better agreement with the manual segmentation. For RMSE, lower indicates better agreement.

4.3 Results

4.3.1 Dice Similarity Index

In Figure 4.3, we show the DSI distributions from the test data for each model. We see that DSI is high on average for all models, with a few scans having a very small DSI (i.e. failures). The median DSI for each model was: 0.89 (logistic), 0.885 (LASSO), 0.88 (GAM), and 0.899 (random forest). We also note that using the random forest results in a slightly higher median DSI compared to the other models, and there was a difference across medians ($\chi^2(3) = 13.49$, $p < 0.05$). Indeed, after Bonferroni correction, the only comparisons that were significant were comparing the random forest DSI to the DSI from the logistic ($p < 0.001$), LASSO ($p < 0.001$), or GAM ($p < 0.001$) models.

In Figure 4.4, we display the CT scan of the patient in the test data that has the median DSI in the test scans. The image depicts the brain-extracted CT scan and then the CT scan with overlaid colors. The green indicates a correct classification of ICH from the model (true positive), blue indicates a false negative, and red indicates a false positive. Note, the image has finer resolution in

the axial slice (0.5mm by 0.5mm) than in the inferior/superior planes (5mm), as is commonly used for radiological evaluation of hemorrhages. Patients with the lowest, 25th, 75th, and highest DSI are shown in Supplemental Figures 4.6, 4.7, 4.8, and 4.9, respectively.

4.3.2 ICH Volume Estimation

In Figure 4.5, we show the estimated ICH volume versus that from the manual segmentation. The pink line represents the $X = Y$ line, where the estimated and true volume are identical. The blue line represents the linear fit; the line equation and correlation are printed on the plot. The farther away the slope of the equation is from 1 represents a multiplicative bias, where values greater than 1 represents larger estimated volumes. The farther the intercept is from 0 represents and additive bias in the estimated volume, where values greater than 0 again represent larger estimated volumes. The correlation (95% confidence interval (CI)) between the true volume and the volume predicted volume were 0.92 (95% CI: 0.884, 0.945) for the logistic model, 0.916 (0.878, 0.942) for the LASSO, 0.908 (95% CI: 0.866, 0.937) for the GAM, and 0.932 (95% CI: 0.901, 0.954) for the random forest. The RMSE for all the logistic (RMSE: 10.67 mL), LASSO (10.83 mL), and random forest (10.27 mL) models, but was slightly higher for the GAM model (11.36 mL). The Kruskal-Wallis test indicated no significant difference in the median absolute value of the difference in estimated versus true volume over models ($\chi^2(3) = 2.3$, $p = 0.51$).

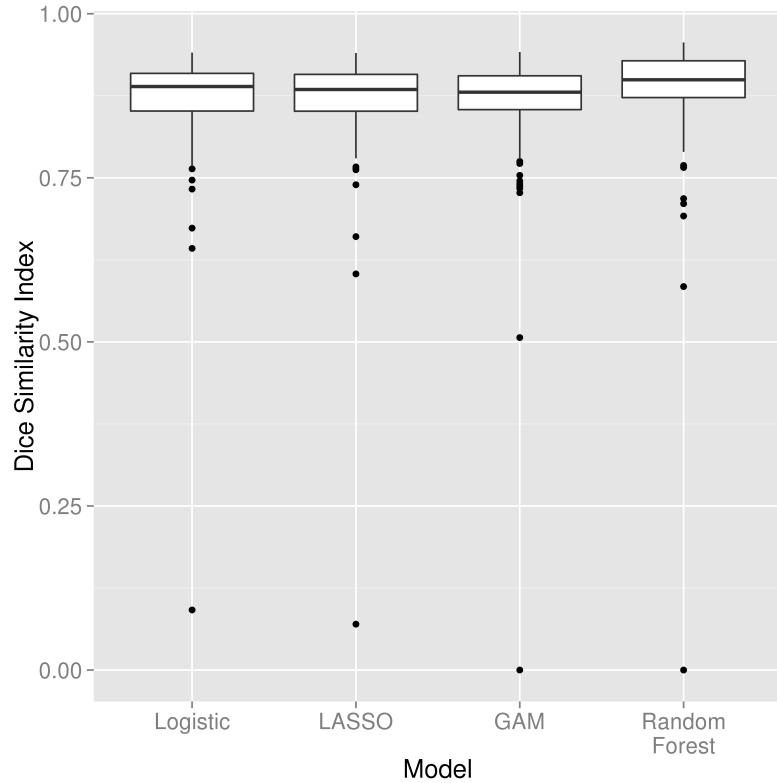


Figure 4.3: Distribution of Dice Similarity in Test Scans. Here we display the boxplot of the Dice Similarity Index (DSI), a measure of spatial overlap between the estimated hemorrhage mask and the manually-delineated hemorrhage mask, in the 102 test scans. We present the DSI distribution for each model fit: a logistic regression, a logistic model penalized with the Least Absolute Shrinkage and Selection Operator penalty (LASSO), a generalized additive model (GAM), and a random forest algorithm. Overall, we see high agreement between the manual and estimated hemorrhage masks with the median of 0.89 for the logistic model, 0.885 for the LASSO, 0.88 for the GAM, and 0.899 for the random forest. The median DSI for the random forest was significantly higher than those of the other 3 models, after adjusting for multiplicity using a Bonferroni correction (all $p < 0.05$).

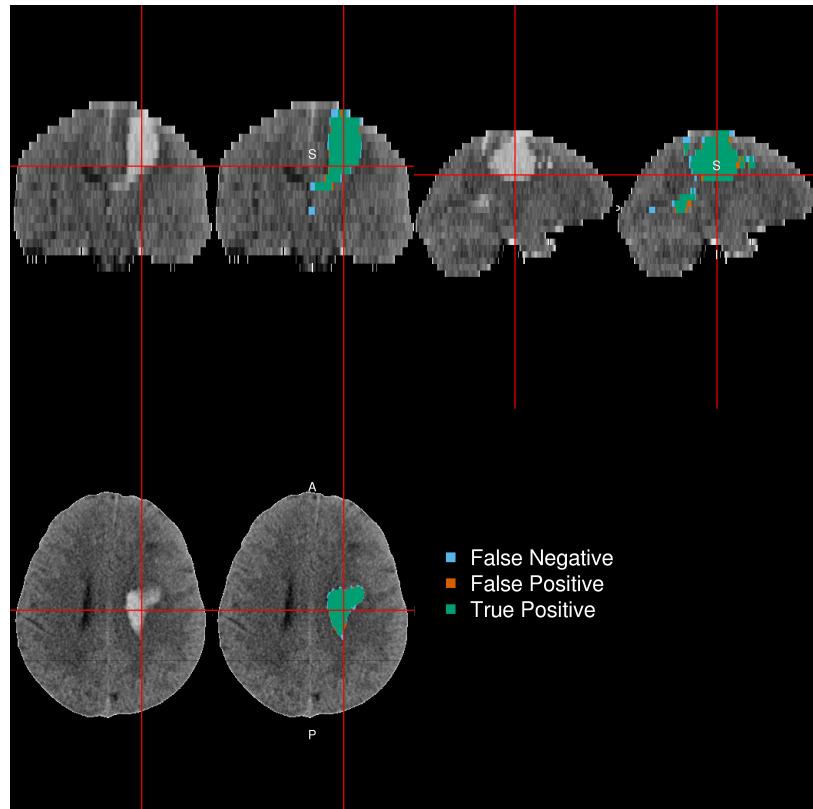


Figure 4.4: Patient with Median Dice Similarity Index. We present the patient with the median Dice Similarity Index (DSI), a measure of spatial overlap, from the chosen predictor model fit with a random forest. The median DSI was 0.899, which indicates high spatial overlap. The green indicates a correct classification of ICH from the model, blue indicates a false negative, where the manual segmentation denoted the area to be ICH but the predicted one did not, and red indicates a false positive, where the predicted segmentation denoted the area to be ICH but the manual one did not.

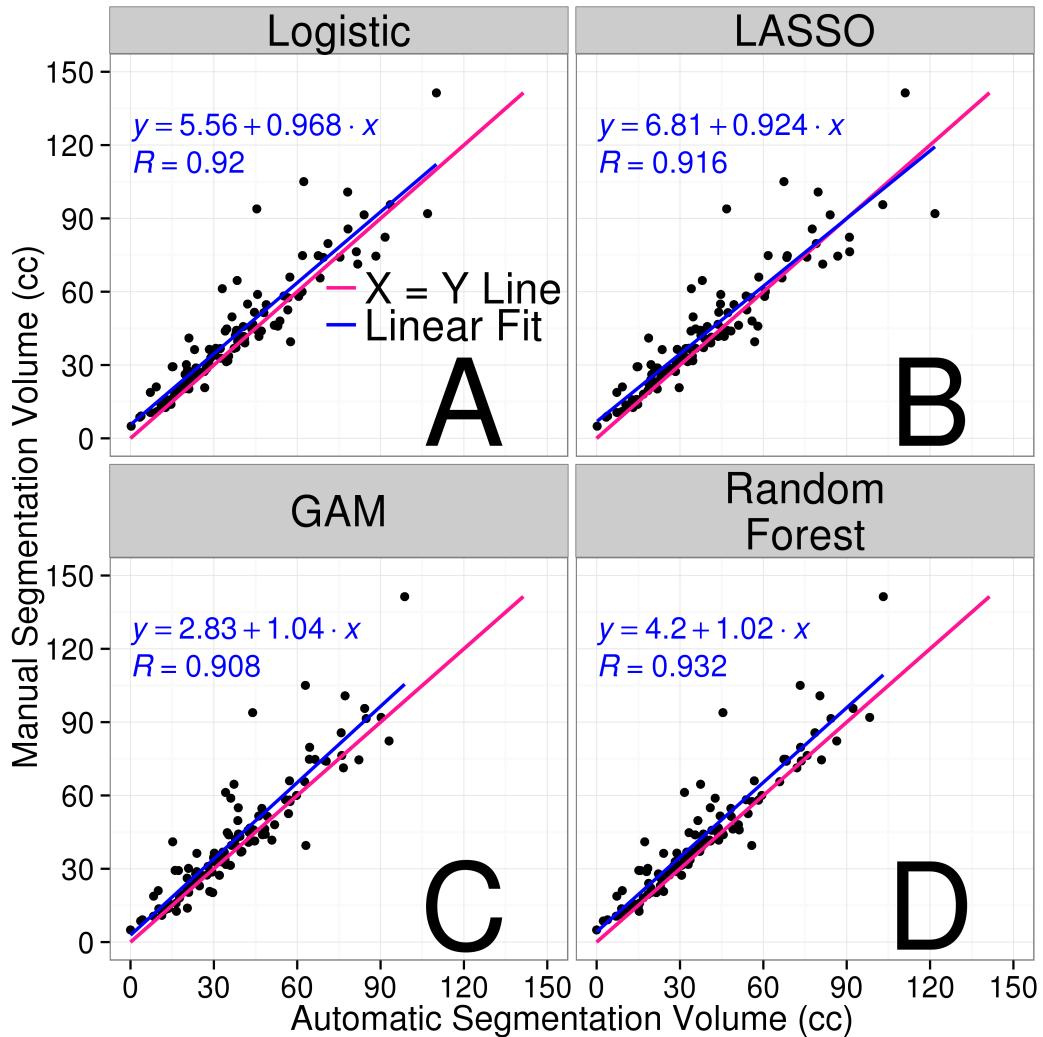


Figure 4.5: Comparison of Estimated and Manual Intracerebral Hemorrhage Volume for Each Model. In each panel, we show the volume of intracerebral hemorrhage (ICH) estimated from each model (x-axis) versus that from the gold-standard manual segmentation (y-axis) in the 102 test scans. The pink line represents the $X = Y$ line, which represents perfect agreement. The blue line represents a linear fit of the data, and the estimated slope equation is displayed along with the Pearson correlation. Panel A represents the volume from the logistic regression model, B represents that from logistic model penalized with the LASSO, C represents that from a generalized additive model (GAM), and D represents that from a random forest algorithm. Overall, we see high agreement between the estimated volume from automated segmentation from each model as all correlations are above 0.9. The farther away the slope of the equation is from 1 represents a multiplicative bias, where values greater than 1 represent larger estimated volumes. The farther the intercept is from 0 represents an additive bias in the estimated volume, where values greater than 0 again represent larger estimated volumes.

4.3.3 Model Choice

Overall, all models perform well for ICH segmentation. Some failures exist, but the algorithm using the random forest to fit the probability of ICH had higher median DSI, the lowest RMSE, and the highest correlation compared to the other models. Therefore, when implementing the algorithm, the model used will be the random forest.

4.4 Discussion

We have presented a novel, fully automated method for segmentation of ICH from CT scans. Our method uses only CT scans from patients with stroke as additional information from MRI was not standardized nor is the standard of care with the disease. We validated this method against the manual segmentation of ICH as a gold standard. We used the Dice Similarity Index and correlation between the volume of ICH from manual and automatic segmentation as measures to determine how well an algorithm performed.

We create a set of predictors that attempt to capture the features relevant to distinguishing ICH versus non-ICH areas and described the rationale for each predictor. We have used these predictors in a series of models, including those fit with logistic regression, logistic regression penalized with the LASSO, generalized additive models, and random forests. All models result in high values for the DSI and high correlations for the volume of ICH from manual and automatic segmentation. Estimating the probability of ICH using random forests appeared was chosen as the final algorithm based on DSI;

that algorithm also had high correlations of automated versus manual ICH volume.

Comparing the manual to the automatic segmentation in each quantile of DSI showed that some cases may fail ($N = 1$ with DSI below 0.5), but much of the discrepancy occurs at the edges of the hemorrhage. This may be due to the smoothing operations performed in the model, which are isotropic and do not preserve edges. Anisotropic smoothing, such as those presented in Perona, Shiota, and Malik (1994), may help preserve edges and result in better segmentation.

Other methods have been presented for segmentation of ICH from CT scans (Gillebert, Humphreys, and Mantini, 2014; Prakash et al., 2012; Loncaric, Cosic, and Dhawan, 1996; Loncaric et al., 1999; Pérez et al., 2007). Loncaric, Cosic, and Dhawan (1996) performed an analysis on only one 2-dimensional scan and could not be compared. The method in Pérez et al. (2007) was semiautomated and was validated by visual inspection, which did not succeed in 6 out of 36 scans (16.7%). If we use the criteria of a DSI below 0.5 as a failure, we had 1 fail out of 102 test scans (1%). We have higher median DSI (0.899) than those reported in Gillebert, Humphreys, and Mantini (2014) (approximately 0.62 and 0.78 for hemorrhagic strokes from graphs) and were comparable to (Prakash et al., 2012) (0.897, 0.858, and 0.9173 for different groups with hemorrhage). Loncaric et al. (1999) did not compare hemorrhage masks to the manual segmentation, but compared ICH volumes on the manual and automatic segmentation on 5 subjects measured 3 times, which had a Pearson correlation of 0.917 on all 15 scans, similar to that reported from the random

forest model selected above ($R = 0.932$).

Only Gillebert, Humphreys, and Mantini (2014) responded with any requests for software to perform the segmentation. Although we believe that the method from Gillebert, Humphreys, and Mantini (2014) is comparable, it has not been packaged for general use. As the analysis above was done in R, we have released a package that can perform ICH segmentation (<https://github.com/muschellij2/ichseg>), including the models for prediction, CT template from Rorden et al. (2012), standardized mean and standard deviation images, and functions to register the images, create the predictors, predict from the models, and return a binary hemorrhage mask. Although a package is ideal for prediction on a large number of images, we believe that this limits those who can test or try the software. Therefore, we have released a Shiny (Chang et al., 2015) R application online (http://bit.ly/ICH_SEG) for researchers to input CT scans and the application will output ICH segmentation masks, while giving an image depicting each processing step.

One of the potential concerns with image registration with patients with ICH is image registration. Although this method uses non-linear registration for one predictor, the standardized-to-template intensity, this predictor is transformed back to the rigid-to-template space after standardization. The method generally relies only rigid-body registration to a template. This registration is largely done for head reorientation and resampling to isotropic voxel sizes across patients. Thus, when morphological operations are performed, such as smoothing using millimeter specifications, they do not depend on the original voxel sizes. Although the models were fit in the rigid-to-template

space, voxels from the same space in template space are not compared across patients in the models. Also, the method returns hemorrhage masks in the native space of the patient, so they are easily comparable to any segmentation method performed on the original scan, such as clustering. Thus, we use non-linear registration in one step, but do not rely on highly accurate image registration for comparability of voxels across patients, which can be problematic with patients with large hemorrhages.

Another concern from the process above is that the algorithm was only fit on 10 patient scans. Moreover, it was only fit with the sampled 100,000 voxels that passed the voxel selection procedure from these scans. Although this is a small sample, the models have shown to have high out-of-sample accuracy. As this was the original procedure done, the training and test sets were not changed after this analysis to help avoid overfitting the models to the test set. If a cross-validation approach was done across the entire set of scans, the models must also be combined in some way to give a final prediction. Validation of this method on additional data is required, but the test scans used are from multiple sites, different scanners, and have patients with heterogeneous hemorrhages both heterogeneous in size and location. Thus, we believe this method to have good out-of-sample accuracy in a similar population.

As the process results in binary hemorrhage masks, one can use methods described in Muschelli et al. (2015b) to estimate quantitative measures of hemorrhage location. Other automated analysis of hemorrhages can be done, such as shape analysis, which could not be done without a binary mask.

Overall, this process allows for researchers to use these hemorrhage masks for other voxel-based analyses that can yield novel insights into the relationship between hemorrhage characteristics and patient outcomes.

4.4.1 Conclusions

We have implemented and validated a fully automated segmentation algorithm of ICH in CT scans. The method relies on series of processing steps and creating a set of relevant predictors based on morphological operations such as smoothing, registration, and intensity normalization. This method has been shown to agree with the gold standard of manual delineation of hemorrhages. As an automated process, it is faster, does not require extensive radiologic image experience, is scalable to thousands of images, and does not have inter-reader variability. As the process results in binary hemorrhage masks, one can then use these to define or test characteristics of ICH compared to patient-level information, such as those described in Muschelli et al. (2015b)) to estimate quantitative measures of hemorrhage location. Most importantly, the volume of ICH is automatically calculated and can be used as a covariate in analysis, as this has been shown to be associated with long-term functional outcome (Broderick et al., 1993; Jordan, Kleinman, and Hillis, 2009; Tuhrim et al., 1999).

Acknowledgments

We would like to thank the patients and families who volunteered for this study, Genentech Inc. for the donation of the study drug (Alteplase), and the readers who manually segmented the ICH. Dr. Chris Rorden was also

extremely helpful in adapting his dcm2ni i software to some issues specific to CT scans.

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 U01NS062851 awarded to Dr. Daniel Hanley from the National Institutes of Health (NIH)/National Institute of Neurological Disorders and Stroke (NINDS). 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).

Bibliography

- Sahni, Ramandeep and Jesse Weinberger (2007). "Management of intracerebral hemorrhage". In: *Vascular Health and Risk Management* 3.5, pp. 701–709. ISSN: 1176-6344. URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2291314/> (visited on 04/04/2014).
- Broderick, J. P., T. G. Brott, J. E. Duldner, T. Tomsick, and G. Huster (1993). "Volume of intracerebral hemorrhage. A powerful and easy-to-use predictor of 30-day mortality." In: *Stroke* 24.7, pp. 987–993. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/01.STR.24.7.987](https://doi.org/10.1161/01.STR.24.7.987). URL: <http://stroke.ahajournals.org/content/24/7/987> (visited on 04/24/2014).
- Hemphill, J. Claude, David C. Bonovich, Lavrentios Besmertis, Geoffrey T. Manley, and S. Claiborne Johnston (2001). "The ICH Score A Simple, Reliable Grading Scale for Intracerebral Hemorrhage". In: *Stroke* 32.4, pp. 891–897. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/01.STR.32.4.891](https://doi.org/10.1161/01.STR.32.4.891). URL: <http://stroke.ahajournals.org/content/32/4/891> (visited on 03/25/2014).
- Tuhrim, Stanley, Deborah R. Horowitz, Michael Sacher, and James H. Godbold (1999). "Volume of ventricular blood is an important determinant of outcome in supratentorial intracerebral hemorrhage". In: *Critical care medicine* 27.3, pp. 617–621. URL: http://journals.lww.com/ccmjournal/Abstract/1999/03000/Volume_of_ventricular_blood_is_anImportant.45.aspx (visited on 12/09/2014).
- Anderson, Craig S., Yining Huang, Ji Guang Wang, Hisatomi Arima, Bruce Neal, Bin Peng, Emma Heeley, Christian Skulina, Mark W. Parsons, Jong Sung Kim, and others (2008). "Intensive blood pressure reduction in acute cerebral haemorrhage trial (INTERACT): a randomised pilot trial". In: *The Lancet Neurology* 7.5, pp. 391–399. URL: <http://www.sciencedirect.com/science/article/pii/S1474442208700693> (visited on 07/08/2014).
- Anderson, Craig S., Yining Huang, Hisatomi Arima, Emma Heeley, Christian Skulina, Mark W. Parsons, Bin Peng, Qiang Li, Steve Su, Qing Ling Tao,

- Yue Chun Li, Jian Dong Jiang, Li Wen Tai, Jin Li Zhang, En Xu, Yan Cheng, Lewis B. Morgenstern, John Chalmers, and Ji Guang Wang (2010). "Effects of Early Intensive Blood Pressure-Lowering Treatment on the Growth of Hematoma and Perihematomal Edema in Acute Intracerebral Hemorrhage The Intensive Blood Pressure Reduction in Acute Cerebral Haemorrhage Trial (INTERACT)". In: *Stroke* 41.2, pp. 307–312. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/STROK.0.109.561795](https://doi.org/10.1161/STROK.0.109.561795). URL: <http://stroke.ahajournals.org/content/41/2/307> (visited on 12/09/2014).
- Qureshi, Adnan I., Yuko Y. Palesch, Renee Martin, Jill Novitzke, Salvador Cruz-Flores, Asad Ehtisham, Mustapha A. Ezzeddine, Joshua N. Goldstein, Jawad F. Kirmani, Haitham M. Hussein, M. Fareed K. Suri, Nauman Tariq, Yuan Liu, and Atach Investigators (2011). "Association of Serum Glucose Concentrations During Acute Hospitalization with Hematoma Expansion, Perihematomal Edema, and Three Month Outcome Among Patients with Intracerebral Hemorrhage". In: *Neurocritical Care* 15.3, pp. 428–435. ISSN: 1541-6933, 1556-0961. DOI: [10.1007/s12028-011-9541-8](https://doi.org/10.1007/s12028-011-9541-8). URL: <http://link.springer.com/article/10.1007/s12028-011-9541-8> (visited on 12/09/2014).
- Mayer, Stephan A., Nikolai C. Brun, Kamilla Begtrup, Joseph Broderick, Stephen Davis, Michael N. Diringer, Brett E. Skolnick, and Thorsten Steiner (2005). "Recombinant Activated Factor VII for Acute Intracerebral Hemorrhage". In: *New England Journal of Medicine* 352.8, pp. 777–785. ISSN: 0028-4793. DOI: [10.1056/NEJMoa042991](https://doi.org/10.1056/NEJMoa042991). URL: <http://www.nejm.org/doi/full/10.1056/NEJMoa042991> (visited on 12/09/2014).
- Morgan, T., M. Zuccarello, R. Narayan, P. Keyl, K. Lane, and D. Hanley (2008a). "Preliminary findings of the minimally-invasive surgery plus rtPA for intracerebral hemorrhage evacuation (MISTIE) clinical trial". In: *Cerebral Hemorrhage*. Springer, pp. 147–151. URL: http://link.springer.com/chapter/10.1007/978-3-211-09469-3_30 (visited on 09/09/2014).
- Morgan, T., I. Awad, P. Keyl, K. Lane, and D. Hanley (2008b). "Preliminary report of the clot lysis evaluating accelerated resolution of intraventricular hemorrhage (CLEAR-IVH) clinical trial". In: *Cerebral Hemorrhage*. Springer, pp. 217–220.
- Rost, Natalia S., Eric E. Smith, Yuchiao Chang, Ryan W. Snider, Rishi Chanderraj, Kristin Schwab, Emily FitzMaurice, Lauren Wendell, Joshua N. Goldstein, Steven M. Greenberg, and Jonathan Rosand (2008). "Prediction of Functional Outcome in Patients With Primary Intracerebral Hemorrhage The FUNC Score". In: *Stroke* 39.8, pp. 2304–2309. ISSN: 0039-2499, 1524-4628.

DOI: 10.1161/STROKEAHA.107.512202. URL: <http://stroke.ahajournals.org/content/39/8/2304> (visited on 11/17/2015).

- Castellanos, M., R. Leira, J. Tejada, A. Gil-Peralta, A. Dàvalos, and J. Castillo (2005). "Predictors of good outcome in medium to large spontaneous supratentorial intracerebral haemorrhages". In: *Journal of Neurology, Neurosurgery & Psychiatry* 76.5, pp. 691–695. ISSN: , 1468-330X. DOI: 10.1136/jnnp.2004.044347. URL: <http://jnnp.bmjjournals.com/content/76/5/691> (visited on 11/17/2015).
- Kothari, Rashmi U., Thomas Brott, Joseph P. Broderick, William G. Barsan, Laura R. Sauerbeck, Mario Zuccarello, and Jane Khoury (1996). "The ABCs of Measuring Intracerebral Hemorrhage Volumes". In: *Stroke* 27.8, pp. 1304–1305. ISSN: 0039-2499, 1524-4628. DOI: 10.1161/01.STR.27.8.1304. URL: <http://stroke.ahajournals.org/content/27/8/1304> (visited on 12/09/2014).
- Divani, Afshin A., Shahram Majidi, Xianghua Luo, Fotis G. Soutsian, Jie Zhang, Aviva Abosch, and Ramachandra P. Tummala (2011). "The ABCs of accurate volumetric measurement of cerebral hematoma". In: *Stroke* 42.6, pp. 1569–1574. URL: <https://stroke.ahajournals.org/content/42/6/1569.full> (visited on 06/17/2014).
- Webb, Alastair J. S., Natalie L. Ullman, Tim C. Morgan, John Muschelli, Joshua Kornbluth, Issam A. Awad, Stephen Mayo, Michael Rosenblum, Wendy Ziai, Mario Zuccarello, Francois Aldrich, Sayona John, Sagi Harnof, George Lopez, William C. Broaddus, Christine Wijman, Paul Vespa, Ross Bullock, Stephen J. Haines, Salvador Cruz-Flores, Stan Tuhrim, Michael D. Hill, Raj Narayan, Daniel F. Hanley, Ashis H. Tayal, Alberto Torres Diaz, Alan Hoffer, Asma Moheet, Sachin Agarwal, Mason Markowski, Guy Rosenthal, Inam Kureshi, Panayiotis Varelas, Pedro Telles Cougo Pinto, Joan Marti-Fabregas, Paul Camarata, David B. Seder, William Freeman, Ann Helms, Christos Lazaridis, Kavian Shahi, David Sinclair, Martin Gizzi, David LeDoux, Michel Torbey, Kevin Cockroft, David Antezana, Julius Gene Latorre, Nevo Margalit, H. Mark Crabtree, Darren Lovick, Chitra Venkatasubramanian, Michael Weaver, Jack Jallo, Michael Reinert, Ken Butcher, Jody Leonardo, Agnieszka Ardel, Opeolu Adeoye, Laszlo Csiba, Katja Wartenberg, Julian Bosel, Fernando Testai, Harold Adams, Carsten Hobohm, Thomas Kerz, Lawrence Wechsler, Paul Gardner, David Decker, Diederik Bulters, Nicole Gonzales, Jean-Louis Caron, Christiana Hall, Safdar Ansari, Andreas Luft, Fuat Arikhan, R. Scott Graham, and

- Kristi Tucker (2015). "Accuracy of the ABC/2 Score for Intracerebral Hemorrhage Systematic Review and Analysis of MISTIE, CLEAR-IVH, and CLEAR III". In: *Stroke* 46.9, pp. 2470–2476. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/STROKEAHA.114.007343](https://doi.org/10.1161/STROKEAHA.114.007343). URL: <http://stroke.ahajournals.org/content/46/9/2470> (visited on 11/16/2015).
- Pedraza, Salvador, Josep Puig, Gerard Blasco, Josep Daunis-i Estadella, Imma Boada, Anton Bardera, Mar Castellanos, and Joaquín Serena (2012). "Reliability of the ABC/2 Method in Determining Acute Infarct Volume". In: *Journal of Neuroimaging* 22.2, pp. 155–159. ISSN: 1552-6569. DOI: [10.1111/j.1552-6569.2011.00588.x](https://doi.org/10.1111/j.1552-6569.2011.00588.x). URL: <http://onlinelibrary.wiley.com/doi/10.1111/j.1552-6569.2011.00588.x/abstract> (visited on 01/21/2015).
- Hussein, Haitham M., Nauman A. Tariq, Yuko Y. Palesch, and Adnan I. Qureshi (2013). "Reliability of Hematoma Volume Measurement at Local Sites in a Multicenter Acute Intracerebral Hemorrhage Clinical Trial". In: *Stroke* 44.1, pp. 237–239. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/STROKEAHA.112.667220](https://doi.org/10.1161/STROKEAHA.112.667220). URL: <http://stroke.ahajournals.org/content/44/1/237> (visited on 01/21/2015).
- Wang, Shuo, Min Lou, Tian Liu, Deqi Cui, Xiaomei Chen, and Yi Wang (2013). "Hematoma Volume Measurement in Gradient Echo MRI Using Quantitative Susceptibility Mapping". In: *Stroke* 44.8, pp. 2315–2317. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/STROKEAHA.113.001638](https://doi.org/10.1161/STROKEAHA.113.001638). URL: <http://stroke.ahajournals.org/content/44/8/2315> (visited on 04/23/2015).
- Carhuapoma, Ricardo J, Daniel F Hanley, Mousumi Banerjee, and Norman J Beauchamp (2003). "Brain Edema After Human Cerebral Hemorrhage A Magnetic Resonance Imaging Volumetric Analysis". In: *Journal of neurosurgical anesthesiology* 15.3, pp. 230–233.
- Gillebert, Céline R., Glyn W. Humphreys, and Dante Mantini (2014). "Automated delineation of stroke lesions using brain CT images". In: *NeuroImage: Clinical* 4, pp. 540–548. ISSN: 2213-1582. DOI: [10.1016/j.nicl.2014.03.009](https://doi.org/10.1016/j.nicl.2014.03.009). URL: <http://www.sciencedirect.com/science/article/pii/S2213158214000394> (visited on 07/17/2014).
- Prakash, K. N. Bhanu, Shi Zhou, Tim C. Morgan, Daniel F. Hanley, and Wieslaw L. Nowinski (2012). "Segmentation and quantification of intraventricular/cerebral hemorrhage in CT scans by modified distance regularized level set evolution technique". In: *International Journal of Computer Assisted Radiology and Surgery* 7.5, pp. 785–798. ISSN: 1861-6410, 1861-6429. DOI: [10.1007/s11548-012-0670-0](https://doi.org/10.1007/s11548-012-0670-0). URL: <http://link.springer.com/article/10.1007/s11548-012-0670-0> (visited on 10/03/2014).

- Loncaric, Sven, Dubravko Cosic, and Atam P. Dhawan (1996). "Hierarchical segmentation of CT head images". In: *Proc IEEE EMBS*. doi 10, p. 1109. URL: http://www.researchgate.net/publication/2690468_Hierarchical_Segmentation_of_CT_Head_Images/file/d912f50edddd1ca672.pdf (visited on 06/17/2014).
- Loncaric, Sven, Atam P. Dhawan, Dubravko Cosic, Domagoj Kovacević, Joseph Broderick, and Thomas Brott (1999). "Quantitative intracerebral brain hemorrhage analysis". In: *Medical Imaging'99*. International Society for Optics and Photonics, pp. 886–894. URL: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=980754> (visited on 06/17/2014).
- Pérez, Noel, José A. Valdés, Miguel A. Guevara, Luis A. Rodríguez, and J. M. Molina (2007). "Set of methods for spontaneous ICH segmentation and tracking from CT head images". In: *Progress in Pattern Recognition, Image Analysis and Applications*. Springer, pp. 212–220. URL: http://link.springer.com/chapter/10.1007/978-3-540-76725-1_23 (visited on 06/17/2014).
- Mould, W. Andrew, J. Ricardo Carhuapoma, John Muschelli, Karen Lane, Timothy C. Morgan, Nichol A. McBee, Amanda J. Bistran-Hall, Natalie L. Ullman, Paul Vespa, Neil A. Martin, Issam Awad, Mario Zuccarello, and Daniel F. Hanley (2013). "Minimally Invasive Surgery Plus Recombinant Tissue-type Plasminogen Activator for Intracerebral Hemorrhage Evacuation Decreases Perihematomal Edema". In: *Stroke* 44.3, pp. 627–634. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/STROKEAHA.111.000411](https://doi.org/10.1161/STROKEAHA.111.000411). URL: <http://stroke.ahajournals.org/content/44/3/627> (visited on 03/07/2014).
- Bergström, M, K Ericson, B Levander, P Svendsen, and S Larsson (1977). "Variation with time of the attenuation values of intracranial hematomas". In: *Journal of computer assisted tomography* 1.1, pp. 57–63. ISSN: 0363-8715.
- Smith, Eric E., Jonathan Rosand, and Steven M. Greenberg (2006). "Imaging of Hemorrhagic Stroke". In: *Magnetic Resonance Imaging Clinics of North America* 14.2, pp. 127–140. ISSN: 1064-9689. DOI: [10.1016/j.mric.2006.06.002](https://doi.org/10.1016/j.mric.2006.06.002). URL: <http://www.sciencedirect.com/science/article/pii/S1064968906000298> (visited on 03/06/2014).
- Rorden, Chris and Matthew Brett (2000). "Stereotaxic Display of Brain Lesions". In: *Behavioural Neurology* 12.4, pp. 191–200. ISSN: 0953-4180. DOI: [10.1155/2000/421719](https://doi.org/10.1155/2000/421719). URL: <http://www.hindawi.com/journals/bn/2000/421719/abs/> (visited on 03/06/2014).

- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.R-project.org/>.
- Muschelli, John, Elizabeth Sweeney, Martin Lindquist, and Ciprian Crainiceanu (2015a). “fslr: Connecting the FSL Software with R”. In: *R Journal* 7.1, pp. 163–175.
- Jenkinson, Mark, Christian F. Beckmann, Timothy E. J. Behrens, Mark W. Woolrich, and Stephen M. Smith (2012). “FSL”. In: *NeuroImage* 62.2, pp. 782–790. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2011.09.015](https://doi.org/10.1016/j.neuroimage.2011.09.015). URL: <http://www.sciencedirect.com/science/article/pii/S1053811911010603> (visited on 01/15/2014).
- Avants, Brian B., Nicholas J. Tustison, Gang Song, Philip A. Cook, Arno Klein, and James C. Gee (2011a). “A reproducible evaluation of ANTs similarity metric performance in brain image registration”. In: *NeuroImage* 54.3, pp. 2033–2044. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2010.09.025](https://doi.org/10.1016/j.neuroimage.2010.09.025). URL: <http://www.sciencedirect.com/science/article/pii/S1053811910012061> (visited on 11/16/2015).
- Smith, Stephen M. (2002). “Fast robust automated brain extraction”. In: *Human Brain Mapping* 17.3, pp. 143–155. ISSN: 1097-0193. DOI: [10.1002/hbm.10062](https://doi.org/10.1002/hbm.10062). URL: <http://onlinelibrary.wiley.com/doi/10.1002/hbm.10062/abstract> (visited on 01/15/2014).
- Muschelli, John, Natalie L. Ullman, W. Andrew Mould, Paul Vespa, Daniel F. Hanley, and Ciprian M. Crainiceanu (2015c). “Validated automatic brain extraction of head CT images”. In: *NeuroImage*. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2015.03.074](https://doi.org/10.1016/j.neuroimage.2015.03.074). URL: <http://www.sciencedirect.com/science/article/pii/S1053811915002700> (visited on 04/20/2015).
- Rorden, Christopher, Leonardo Bonilha, Julius Fridriksson, Benjamin Bender, and Hans-Otto Karnath (2012). “Age-specific CT and MRI templates for spatial normalization”. In: *NeuroImage* 61.4, pp. 957–965. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2012.03.020](https://doi.org/10.1016/j.neuroimage.2012.03.020). URL: <http://www.sciencedirect.com/science/article/pii/S1053811912002935> (visited on 01/15/2014).
- Frequently Asked Questions for FLIRT*. <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT/FAQ>.
- Avants, Brian B, Nicholas J Tustison, Jue Wu, Philip A Cook, and James C Gee (2011b). “An open source multivariate framework for n-tissue segmentation with evaluation on public data”. In: *Neuroinformatics* 9.4, pp. 381–400.

- Avants, B. B., C. L. Epstein, M. Grossman, and J. C. Gee (2008). "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain". In: *Medical Image Analysis*. Special Issue on The Third International Workshop on Biomedical Image Registration - WBIR 2006 12.1, pp. 26–41. ISSN: 1361-8415. DOI: [10.1016/j.media.2007.06.004](https://doi.org/10.1016/j.media.2007.06.004). URL: <http://www.sciencedirect.com/science/article/pii/S1361841507000606> (visited on 10/28/2014).
- Tibshirani, Robert (1996). "Regression Shrinkage and Selection via the Lasso". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 58.1, pp. 267–288. ISSN: 0035-9246. URL: <http://www.jstor.org/stable/2346178> (visited on 11/18/2015).
- Friedman, Jerome, Trevor Hastie, and Rob Tibshirani (2010). "Regularization Paths for Generalized Linear Models via Coordinate Descent". In: *Journal of statistical software* 33.1, pp. 1–22. ISSN: 1548-7660. URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2929880/> (visited on 11/18/2015).
- Hastie, Trevor and Robert Tibshirani (1986). "Generalized additive models". In: *Statistical science*, pp. 297–310. URL: <http://www.jstor.org/stable/2245459> (visited on 11/17/2015).
- Hastie, Trevor J. and Robert J. Tibshirani (1990). *Generalized additive models*. Vol. 43. CRC Press. URL: https://books.google.com/books?hl=en&lr=&id=qa29r1Ze1coC&oi=fnd&pg=PR13&dq=Generalized+additive+models&ots=j3-UmlwXmJ&sig=1qxwMbQHaHxAJf-Cgqvy_rGodR0 (visited on 11/17/2015).
- Wood, Simon N. (2011). "Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73.1, pp. 3–36. URL: <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9868.2010.00749.x/pdf> (visited on 11/17/2015).
- Wood, Simon N., Yannig Goude, and Simon Shaw (2015). "Generalized additive models for large data sets". In: *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 64.1, pp. 139–155. ISSN: 1467-9876. DOI: [10.1111/rssc.12068](https://doi.org/10.1111/rssc.12068). URL: <http://onlinelibrary.wiley.com/doi/10.1111/rssc.12068/abstract> (visited on 11/18/2015).
- Breiman, Leo (2001). "Random forests". In: *Machine learning* 45.1, pp. 5–32.
- Liaw, Andy and Matthew Wiener (2002). "Classification and Regression by randomForest". In: *R News* 2.3, pp. 18–22. URL: <http://CRAN.R-project.org/doc/Rnews/>.

- Dice, Lee R. (1945). "Measures of the amount of ecologic association between species". In: *Ecology* 26.3, pp. 297–302. URL: <http://www.jstor.org/stable/1932409> (visited on 09/09/2014).
- Perona, Pietro, Takahiro Shiota, and Jitendra Malik (1994). "Anisotropic diffusion". In: *Geometry-driven diffusion in computer vision*. Springer, pp. 73–92.
- Chang, Winston, Joe Cheng, JJ Allaire, Yihui Xie, and Jonathan McPherson (2015). *shiny: Web Application Framework for R*. URL: <http://CRAN.R-project.org/package=shiny>.
- Muschelli, John, Natalie L Ullman, Elizabeth M Sweeney, Ani Eloyan, Neil Martin, Paul Vespa, Daniel F Hanley, and Ciprian M Crainiceanu (2015b). "Quantitative Intracerebral Hemorrhage Localization". In: *Stroke* 46.11, pp. 3270–3273.
- Jordan, Lori C, Jonathan T Kleinman, and Argye E Hillis (2009). "Intracerebral hemorrhage volume predicts poor neurologic outcome in children". In: *Stroke* 40.5, pp. 1666–1671.

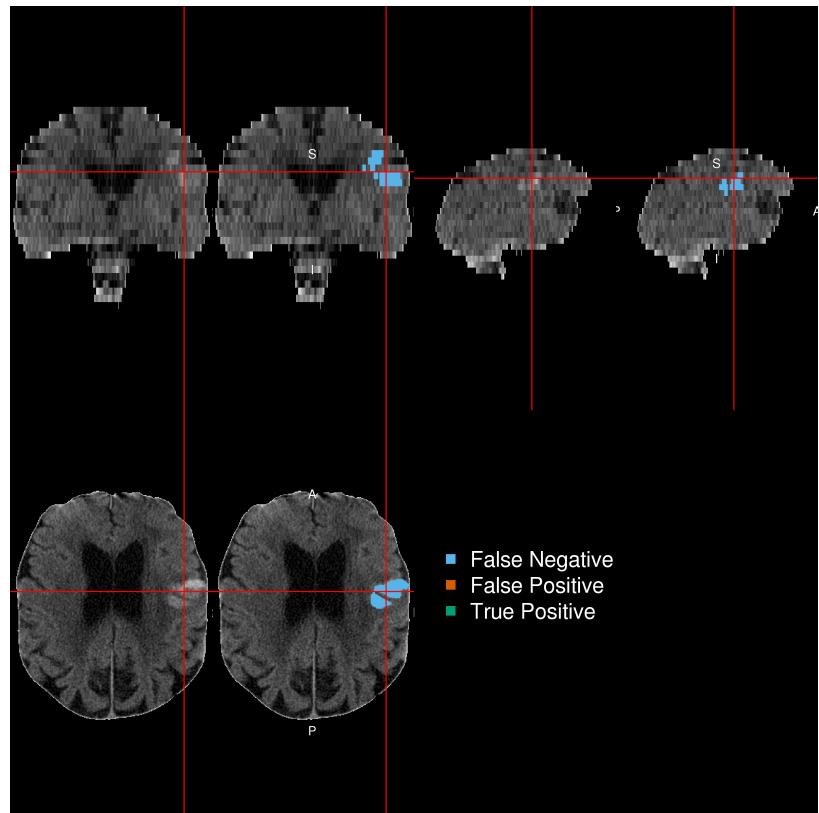


Figure 4.6: Patient with Lowest Dice Similarity Index. We present the patient with the lowest Dice Similarity Index (DSI), a measure of spatial overlap, from the chosen predictor model fit with a random forest. The lowest DSI was 0. The green indicates a correct classification of ICH from the model, blue indicates a false negative, where the manual segmentation denoted the area to be ICH but the predicted one did not, and red indicates a false positive, where the predicted segmentation denoted the area to be ICH but the manual one did not.

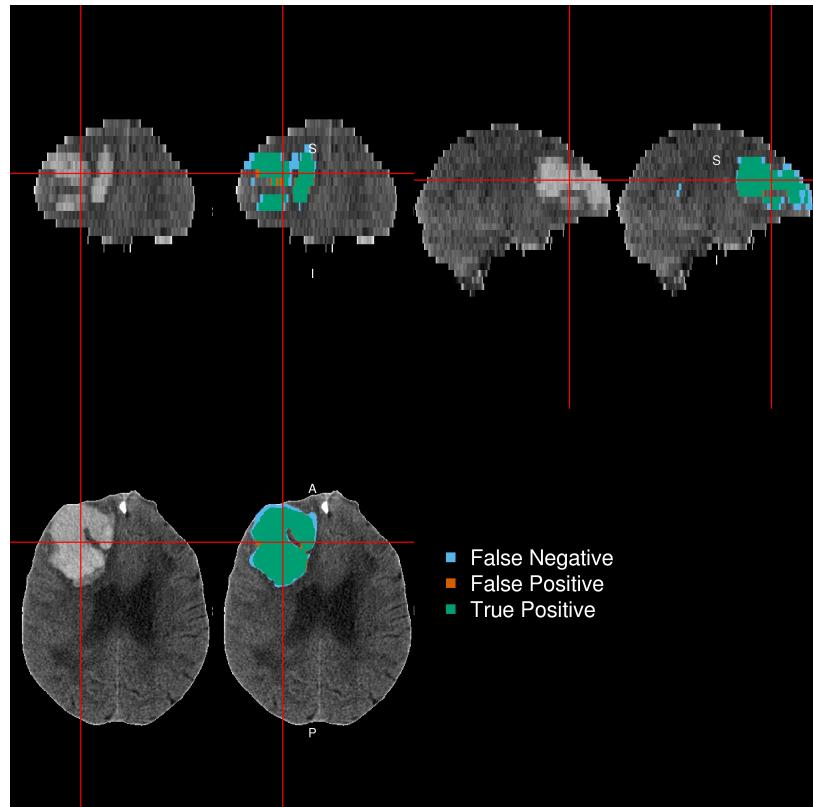


Figure 4.7: Patient with 25th Quantile Dice Similarity Index. We present the patient with the 25th quantile Dice Similarity Index (DSI), a measure of spatial overlap, from the chosen predictor model fit with a random forest. The 25th quantile DSI was 0.872. The green indicates a correct classification of ICH from the model, blue indicates a false negative, where the manual segmentation denoted the area to be ICH but the predicted one did not, and red indicates a false positive, where the predicted segmentation denoted the area to be ICH but the manual one did not.

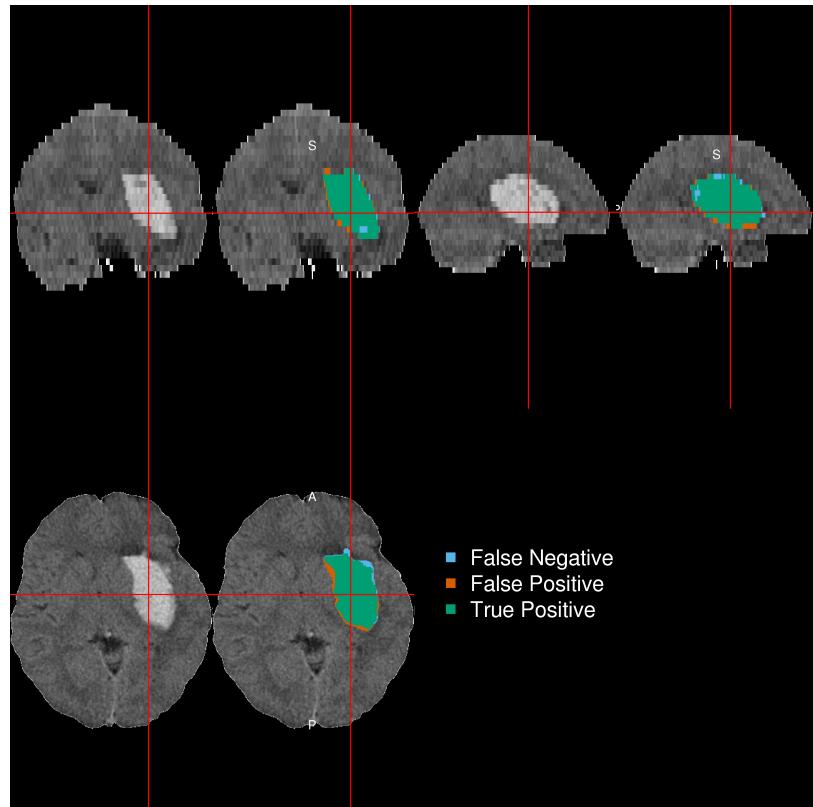


Figure 4.8: Patient with 75th Quantile Dice Similarity Index. We present the patient with the 75th quantile Dice Similarity Index (DSI), a measure of spatial overlap, from the chosen predictor model fit with a random forest. The 75th quantile DSI was 0.928. The green indicates a correct classification of ICH from the model, blue indicates a false negative, where the manual segmentation denoted the area to be ICH but the predicted one did not, and red indicates a false positive, where the predicted segmentation denoted the area to be ICH but the manual one did not.

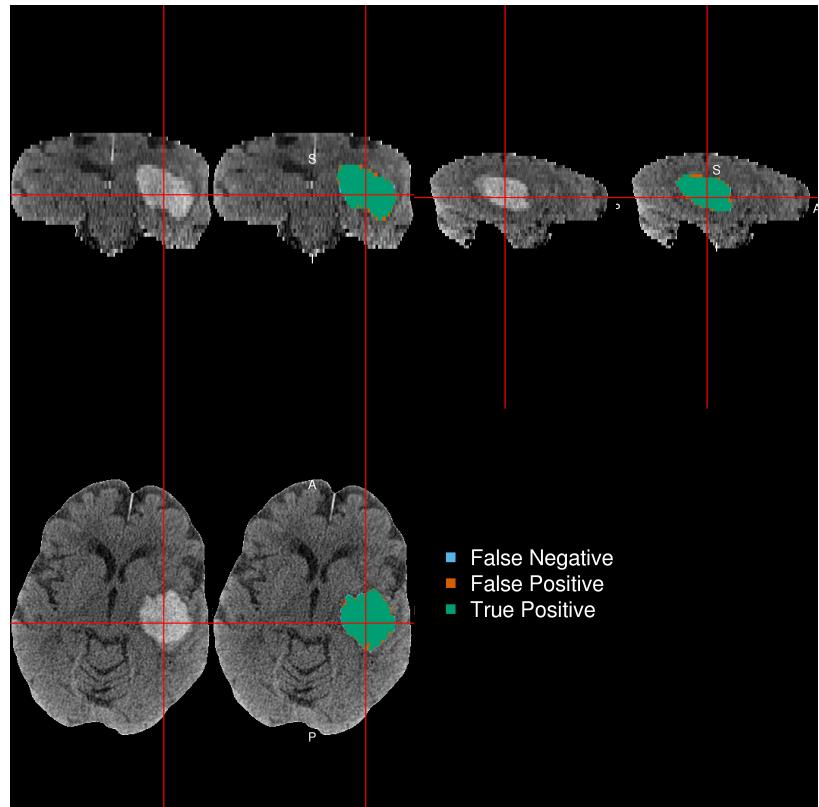


Figure 4.9: Patient with Highest Dice Similarity Index. We present the patient with the highest Dice Similarity Index (DSI), a measure of spatial overlap, from the chosen predictor model fit with a random forest. The highest DSI was 0.956. The green indicates a correct classification of ICH from the model, blue indicates a false negative, where the manual segmentation denoted the area to be ICH but the predicted one did not, and red indicates a false positive, where the predicted segmentation denoted the area to be ICH but the manual one did not.

4.5 Supplemental Material

4.5.1 Examples of Dice Similarity Index in Test Scans

4.5.2 Model Specification

Let $Y_i(v)$ represent the binary hemorrhage mask indicator for voxel v , from patient i , and $x_{i,v}(k)$ represent the predictor image for image k , $k = 1, \dots, 21$.

$$\text{logit}(P(Y_i(v) = 1)) = \beta_0 + \sum_{k=1}^{21} x_{i,k}(v)\beta_k$$

The coefficients for the logistic model are (in log odds or log odds ratios):

Predictor	Beta
Intercept	1.008
Neighborhood mean	0.051
Neighborhood sd	0.000
Neighborhood skew	0.065
Neighborhood kurtosis	-0.352
Image intensity (HU)	-0.172
Threshold ($>= 40$ and $<= 80$)	-0.151
Within-plane Coronal	-0.632
Within-plane Sagittal	-0.249
Within-plane Axial	1.037
Winsorized standardized (20\% trim)	0.547
Percentage thresholded neighbors	2.061
Atropos probability image	0.150
Percent of zero neighbors	-9.180
Indicator of any zero neighbors	0.071
Distance to Image centroid	-0.087
Gaussian smooth (5mm ³)	-0.051
Gaussian smooth (10mm ³)	0.550
Gaussian smooth (20mm ³)	-0.390
Standardized-to-template Intensity	1.460
Contralateral difference	0.033

The specification for the functional form of the model fit with the LASSO

penalty, is the same, but optimizes the following criteria:

$$\min_{\beta} - \left(\frac{1}{\sum_i V_i} \sum_i Y_i(v) \times X_i(v)\beta - \log \left(1 + e^{X_i(v)\beta} \right) \right) + \lambda \sum_k |\beta_k|$$

References

- Sahni, Ramandeep and Jesse Weinberger (2007). "Management of intracerebral hemorrhage". In: *Vascular Health and Risk Management* 3.5, pp. 701–709. ISSN: 1176-6344. URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2291314/> (visited on 04/04/2014).
- Broderick, J. P., T. G. Brott, J. E. Duldner, T. Tomsick, and G. Huster (1993). "Volume of intracerebral hemorrhage. A powerful and easy-to-use predictor of 30-day mortality." In: *Stroke* 24.7, pp. 987–993. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/01.STR.24.7.987](https://doi.org/10.1161/01.STR.24.7.987). URL: <http://stroke.ahajournals.org/content/24/7/987> (visited on 04/24/2014).
- Hemphill, J. Claude, David C. Bonovich, Lavrentios Besmertis, Geoffrey T. Manley, and S. Claiborne Johnston (2001). "The ICH Score A Simple, Reliable Grading Scale for Intracerebral Hemorrhage". In: *Stroke* 32.4, pp. 891–897. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/01.STR.32.4.891](https://doi.org/10.1161/01.STR.32.4.891). URL: <http://stroke.ahajournals.org/content/32/4/891> (visited on 03/25/2014).
- Tuhrim, Stanley, Deborah R. Horowitz, Michael Sacher, and James H. Godbold (1999). "Volume of ventricular blood is an important determinant of outcome in supratentorial intracerebral hemorrhage". In: *Critical care medicine* 27.3, pp. 617–621. URL: http://journals.lww.com/ccmjournal/Abstract/1999/03000/Volume_of_ventricular_blood_is_anImportant.45.aspx (visited on 12/09/2014).
- Anderson, Craig S., Yining Huang, Ji Guang Wang, Hisatomi Arima, Bruce Neal, Bin Peng, Emma Heeley, Christian Skulina, Mark W. Parsons, Jong Sung Kim, and others (2008). "Intensive blood pressure reduction in acute cerebral haemorrhage trial (INTERACT): a randomised pilot trial". In: *The Lancet Neurology* 7.5, pp. 391–399. URL: <http://www.sciencedirect.com/science/article/pii/S1474442208700693> (visited on 07/08/2014).
- Anderson, Craig S., Yining Huang, Hisatomi Arima, Emma Heeley, Christian Skulina, Mark W. Parsons, Bin Peng, Qiang Li, Steve Su, Qing Ling Tao,

- Yue Chun Li, Jian Dong Jiang, Li Wen Tai, Jin Li Zhang, En Xu, Yan Cheng, Lewis B. Morgenstern, John Chalmers, and Ji Guang Wang (2010). "Effects of Early Intensive Blood Pressure-Lowering Treatment on the Growth of Hematoma and Perihematomal Edema in Acute Intracerebral Hemorrhage The Intensive Blood Pressure Reduction in Acute Cerebral Haemorrhage Trial (INTERACT)". In: *Stroke* 41.2, pp. 307–312. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/STROK.0.109.561795](https://doi.org/10.1161/STROK.0.109.561795). URL: <http://stroke.ahajournals.org/content/41/2/307> (visited on 12/09/2014).
- Qureshi, Adnan I., Yuko Y. Palesch, Renee Martin, Jill Novitzke, Salvador Cruz-Flores, Asad Ehtisham, Mustapha A. Ezzeddine, Joshua N. Goldstein, Jawad F. Kirmani, Haitham M. Hussein, M. Fareed K. Suri, Nauman Tariq, Yuan Liu, and Atach Investigators (2011). "Association of Serum Glucose Concentrations During Acute Hospitalization with Hematoma Expansion, Perihematomal Edema, and Three Month Outcome Among Patients with Intracerebral Hemorrhage". In: *Neurocritical Care* 15.3, pp. 428–435. ISSN: 1541-6933, 1556-0961. DOI: [10.1007/s12028-011-9541-8](https://doi.org/10.1007/s12028-011-9541-8). URL: <http://link.springer.com/article/10.1007/s12028-011-9541-8> (visited on 12/09/2014).
- Mayer, Stephan A., Nikolai C. Brun, Kamilla Begtrup, Joseph Broderick, Stephen Davis, Michael N. Diringer, Brett E. Skolnick, and Thorsten Steiner (2005). "Recombinant Activated Factor VII for Acute Intracerebral Hemorrhage". In: *New England Journal of Medicine* 352.8, pp. 777–785. ISSN: 0028-4793. DOI: [10.1056/NEJMoa042991](https://doi.org/10.1056/NEJMoa042991). URL: <http://www.nejm.org/doi/full/10.1056/NEJMoa042991> (visited on 12/09/2014).
- Morgan, T., M. Zuccarello, R. Narayan, P. Keyl, K. Lane, and D. Hanley (2008a). "Preliminary findings of the minimally-invasive surgery plus rtPA for intracerebral hemorrhage evacuation (MISTIE) clinical trial". In: *Cerebral Hemorrhage*. Springer, pp. 147–151. URL: http://link.springer.com/chapter/10.1007/978-3-211-09469-3_30 (visited on 09/09/2014).
- Morgan, T., I. Awad, P. Keyl, K. Lane, and D. Hanley (2008b). "Preliminary report of the clot lysis evaluating accelerated resolution of intraventricular hemorrhage (CLEAR-IVH) clinical trial". In: *Cerebral Hemorrhage*. Springer, pp. 217–220.
- Rost, Natalia S., Eric E. Smith, Yuchiao Chang, Ryan W. Snider, Rishi Chanderraj, Kristin Schwab, Emily FitzMaurice, Lauren Wendell, Joshua N. Goldstein, Steven M. Greenberg, and Jonathan Rosand (2008). "Prediction of Functional Outcome in Patients With Primary Intracerebral Hemorrhage The FUNC Score". In: *Stroke* 39.8, pp. 2304–2309. ISSN: 0039-2499, 1524-4628.

DOI: 10.1161/STROKEAHA.107.512202. URL: <http://stroke.ahajournals.org/content/39/8/2304> (visited on 11/17/2015).

- Castellanos, M., R. Leira, J. Tejada, A. Gil-Peralta, A. Dàvalos, and J. Castillo (2005). "Predictors of good outcome in medium to large spontaneous supratentorial intracerebral haemorrhages". In: *Journal of Neurology, Neurosurgery & Psychiatry* 76.5, pp. 691–695. ISSN: , 1468-330X. DOI: 10.1136/jnnp.2004.044347. URL: <http://jnnp.bmjjournals.com/content/76/5/691> (visited on 11/17/2015).
- Kothari, Rashmi U., Thomas Brott, Joseph P. Broderick, William G. Barsan, Laura R. Sauerbeck, Mario Zuccarello, and Jane Khoury (1996). "The ABCs of Measuring Intracerebral Hemorrhage Volumes". In: *Stroke* 27.8, pp. 1304–1305. ISSN: 0039-2499, 1524-4628. DOI: 10.1161/01.STR.27.8.1304. URL: <http://stroke.ahajournals.org/content/27/8/1304> (visited on 12/09/2014).
- Divani, Afshin A., Shahram Majidi, Xianghua Luo, Fotis G. Soutsian, Jie Zhang, Aviva Abosch, and Ramachandra P. Tummala (2011). "The ABCs of accurate volumetric measurement of cerebral hematoma". In: *Stroke* 42.6, pp. 1569–1574. URL: <https://stroke.ahajournals.org/content/42/6/1569.full> (visited on 06/17/2014).
- Webb, Alastair J. S., Natalie L. Ullman, Tim C. Morgan, John Muschelli, Joshua Kornbluth, Issam A. Awad, Stephen Mayo, Michael Rosenblum, Wendy Ziai, Mario Zuccarello, Francois Aldrich, Sayona John, Sagi Harnof, George Lopez, William C. Broaddus, Christine Wijman, Paul Vespa, Ross Bullock, Stephen J. Haines, Salvador Cruz-Flores, Stan Tuhrim, Michael D. Hill, Raj Narayan, Daniel F. Hanley, Ashis H. Tayal, Alberto Torres Diaz, Alan Hoffer, Asma Moheet, Sachin Agarwal, Mason Markowski, Guy Rosenthal, Inam Kureshi, Panayiotis Varelas, Pedro Telles Cougo Pinto, Joan Marti-Fabregas, Paul Camarata, David B. Seder, William Freeman, Ann Helms, Christos Lazaridis, Kavian Shahi, David Sinclair, Martin Gizzi, David LeDoux, Michel Torbey, Kevin Cockroft, David Antezana, Julius Gene Latorre, Nevo Margalit, H. Mark Crabtree, Darren Lovick, Chitra Venkatasubramanian, Michael Weaver, Jack Jallo, Michael Reinert, Ken Butcher, Jody Leonardo, Agnieszka Ardel, Opeolu Adeoye, Laszlo Csiba, Katja Wartenberg, Julian Bosel, Fernando Testai, Harold Adams, Carsten Hobohm, Thomas Kerz, Lawrence Wechsler, Paul Gardner, David Decker, Diederik Bulters, Nicole Gonzales, Jean-Louis Caron, Christiana Hall, Safdar Ansari, Andreas Luft, Fuat Arikhan, R. Scott Graham, and

- Kristi Tucker (2015). "Accuracy of the ABC/2 Score for Intracerebral Hemorrhage Systematic Review and Analysis of MISTIE, CLEAR-IVH, and CLEAR III". In: *Stroke* 46.9, pp. 2470–2476. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/STROKEAHA.114.007343](https://doi.org/10.1161/STROKEAHA.114.007343). URL: <http://stroke.ahajournals.org/content/46/9/2470> (visited on 11/16/2015).
- Pedraza, Salvador, Josep Puig, Gerard Blasco, Josep Daunis-i Estadella, Imma Boada, Anton Bardera, Mar Castellanos, and Joaquín Serena (2012). "Reliability of the ABC/2 Method in Determining Acute Infarct Volume". In: *Journal of Neuroimaging* 22.2, pp. 155–159. ISSN: 1552-6569. DOI: [10.1111/j.1552-6569.2011.00588.x](https://doi.org/10.1111/j.1552-6569.2011.00588.x). URL: <http://onlinelibrary.wiley.com/doi/10.1111/j.1552-6569.2011.00588.x/abstract> (visited on 01/21/2015).
- Hussein, Haitham M., Nauman A. Tariq, Yuko Y. Palesch, and Adnan I. Qureshi (2013). "Reliability of Hematoma Volume Measurement at Local Sites in a Multicenter Acute Intracerebral Hemorrhage Clinical Trial". In: *Stroke* 44.1, pp. 237–239. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/STROKEAHA.112.667220](https://doi.org/10.1161/STROKEAHA.112.667220). URL: <http://stroke.ahajournals.org/content/44/1/237> (visited on 01/21/2015).
- Wang, Shuo, Min Lou, Tian Liu, Deqi Cui, Xiaomei Chen, and Yi Wang (2013). "Hematoma Volume Measurement in Gradient Echo MRI Using Quantitative Susceptibility Mapping". In: *Stroke* 44.8, pp. 2315–2317. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/STROKEAHA.113.001638](https://doi.org/10.1161/STROKEAHA.113.001638). URL: <http://stroke.ahajournals.org/content/44/8/2315> (visited on 04/23/2015).
- Carhuapoma, Ricardo J, Daniel F Hanley, Mousumi Banerjee, and Norman J Beauchamp (2003). "Brain Edema After Human Cerebral Hemorrhage A Magnetic Resonance Imaging Volumetric Analysis". In: *Journal of neurosurgical anesthesiology* 15.3, pp. 230–233.
- Gillebert, Céline R., Glyn W. Humphreys, and Dante Mantini (2014). "Automated delineation of stroke lesions using brain CT images". In: *NeuroImage: Clinical* 4, pp. 540–548. ISSN: 2213-1582. DOI: [10.1016/j.nicl.2014.03.009](https://doi.org/10.1016/j.nicl.2014.03.009). URL: <http://www.sciencedirect.com/science/article/pii/S2213158214000394> (visited on 07/17/2014).
- Prakash, K. N. Bhanu, Shi Zhou, Tim C. Morgan, Daniel F. Hanley, and Wieslaw L. Nowinski (2012). "Segmentation and quantification of intraventricular/cerebral hemorrhage in CT scans by modified distance regularized level set evolution technique". In: *International Journal of Computer Assisted Radiology and Surgery* 7.5, pp. 785–798. ISSN: 1861-6410, 1861-6429. DOI: [10.1007/s11548-012-0670-0](https://doi.org/10.1007/s11548-012-0670-0). URL: <http://link.springer.com/article/10.1007/s11548-012-0670-0> (visited on 10/03/2014).

- Loncaric, Sven, Dubravko Cosic, and Atam P. Dhawan (1996). "Hierarchical segmentation of CT head images". In: *Proc IEEE EMBS*. doi 10, p. 1109. URL: http://www.researchgate.net/publication/2690468_Hierarchical_Segmentation_of_CT_Head_Images/file/d912f50edddd1ca672.pdf (visited on 06/17/2014).
- Loncaric, Sven, Atam P. Dhawan, Dubravko Cosic, Domagoj Kovacević, Joseph Broderick, and Thomas Brott (1999). "Quantitative intracerebral brain hemorrhage analysis". In: *Medical Imaging'99*. International Society for Optics and Photonics, pp. 886–894. URL: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=980754> (visited on 06/17/2014).
- Pérez, Noel, José A. Valdés, Miguel A. Guevara, Luis A. Rodríguez, and J. M. Molina (2007). "Set of methods for spontaneous ICH segmentation and tracking from CT head images". In: *Progress in Pattern Recognition, Image Analysis and Applications*. Springer, pp. 212–220. URL: http://link.springer.com/chapter/10.1007/978-3-540-76725-1_23 (visited on 06/17/2014).
- Mould, W. Andrew, J. Ricardo Carhuapoma, John Muschelli, Karen Lane, Timothy C. Morgan, Nichol A. McBee, Amanda J. Bistran-Hall, Natalie L. Ullman, Paul Vespa, Neil A. Martin, Issam Awad, Mario Zuccarello, and Daniel F. Hanley (2013). "Minimally Invasive Surgery Plus Recombinant Tissue-type Plasminogen Activator for Intracerebral Hemorrhage Evacuation Decreases Perihematomal Edema". In: *Stroke* 44.3, pp. 627–634. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/STROKEAHA.111.000411](https://doi.org/10.1161/STROKEAHA.111.000411). URL: <http://stroke.ahajournals.org/content/44/3/627> (visited on 03/07/2014).
- Bergström, M, K Ericson, B Levander, P Svendsen, and S Larsson (1977). "Variation with time of the attenuation values of intracranial hematomas". In: *Journal of computer assisted tomography* 1.1, pp. 57–63. ISSN: 0363-8715.
- Smith, Eric E., Jonathan Rosand, and Steven M. Greenberg (2006). "Imaging of Hemorrhagic Stroke". In: *Magnetic Resonance Imaging Clinics of North America* 14.2, pp. 127–140. ISSN: 1064-9689. DOI: [10.1016/j.mric.2006.06.002](https://doi.org/10.1016/j.mric.2006.06.002). URL: <http://www.sciencedirect.com/science/article/pii/S1064968906000298> (visited on 03/06/2014).
- Rorden, Chris and Matthew Brett (2000). "Stereotaxic Display of Brain Lesions". In: *Behavioural Neurology* 12.4, pp. 191–200. ISSN: 0953-4180. DOI: [10.1155/2000/421719](https://doi.org/10.1155/2000/421719). URL: <http://www.hindawi.com/journals/bn/2000/421719/abs/> (visited on 03/06/2014).

- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.R-project.org/>.
- Muschelli, John, Elizabeth Sweeney, Martin Lindquist, and Ciprian Crainiceanu (2015a). "fslr: Connecting the FSL Software with R". In: *R Journal* 7.1, pp. 163–175.
- Jenkinson, Mark, Christian F. Beckmann, Timothy E. J. Behrens, Mark W. Woolrich, and Stephen M. Smith (2012). "FSL". In: *NeuroImage* 62.2, pp. 782–790. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2011.09.015](https://doi.org/10.1016/j.neuroimage.2011.09.015). URL: <http://www.sciencedirect.com/science/article/pii/S1053811911010603> (visited on 01/15/2014).
- Avants, Brian B., Nicholas J. Tustison, Gang Song, Philip A. Cook, Arno Klein, and James C. Gee (2011a). "A reproducible evaluation of ANTs similarity metric performance in brain image registration". In: *NeuroImage* 54.3, pp. 2033–2044. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2010.09.025](https://doi.org/10.1016/j.neuroimage.2010.09.025). URL: <http://www.sciencedirect.com/science/article/pii/S1053811910012061> (visited on 11/16/2015).
- Smith, Stephen M. (2002). "Fast robust automated brain extraction". In: *Human Brain Mapping* 17.3, pp. 143–155. ISSN: 1097-0193. DOI: [10.1002/hbm.10062](https://doi.org/10.1002/hbm.10062). URL: <http://onlinelibrary.wiley.com/doi/10.1002/hbm.10062/abstract> (visited on 01/15/2014).
- Muschelli, John, Natalie L. Ullman, W. Andrew Mould, Paul Vespa, Daniel F. Hanley, and Ciprian M. Crainiceanu (2015c). "Validated automatic brain extraction of head CT images". In: *NeuroImage*. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2015.03.074](https://doi.org/10.1016/j.neuroimage.2015.03.074). URL: <http://www.sciencedirect.com/science/article/pii/S1053811915002700> (visited on 04/20/2015).
- Rorden, Christopher, Leonardo Bonilha, Julius Fridriksson, Benjamin Bender, and Hans-Otto Karnath (2012). "Age-specific CT and MRI templates for spatial normalization". In: *NeuroImage* 61.4, pp. 957–965. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2012.03.020](https://doi.org/10.1016/j.neuroimage.2012.03.020). URL: <http://www.sciencedirect.com/science/article/pii/S1053811912002935> (visited on 01/15/2014).
- Frequently Asked Questions for FLIRT*. <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT/FAQ>.
- Avants, Brian B, Nicholas J Tustison, Jue Wu, Philip A Cook, and James C Gee (2011b). "An open source multivariate framework for n-tissue segmentation with evaluation on public data". In: *Neuroinformatics* 9.4, pp. 381–400.

- Avants, B. B., C. L. Epstein, M. Grossman, and J. C. Gee (2008). "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain". In: *Medical Image Analysis*. Special Issue on The Third International Workshop on Biomedical Image Registration - WBIR 2006 12.1, pp. 26–41. ISSN: 1361-8415. DOI: [10.1016/j.media.2007.06.004](https://doi.org/10.1016/j.media.2007.06.004). URL: <http://www.sciencedirect.com/science/article/pii/S1361841507000606> (visited on 10/28/2014).
- Tibshirani, Robert (1996). "Regression Shrinkage and Selection via the Lasso". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 58.1, pp. 267–288. ISSN: 0035-9246. URL: <http://www.jstor.org/stable/2346178> (visited on 11/18/2015).
- Friedman, Jerome, Trevor Hastie, and Rob Tibshirani (2010). "Regularization Paths for Generalized Linear Models via Coordinate Descent". In: *Journal of statistical software* 33.1, pp. 1–22. ISSN: 1548-7660. URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2929880/> (visited on 11/18/2015).
- Hastie, Trevor and Robert Tibshirani (1986). "Generalized additive models". In: *Statistical science*, pp. 297–310. URL: <http://www.jstor.org/stable/2245459> (visited on 11/17/2015).
- Hastie, Trevor J. and Robert J. Tibshirani (1990). *Generalized additive models*. Vol. 43. CRC Press. URL: https://books.google.com/books?hl=en&lr=&id=qa29r1Ze1coC&oi=fnd&pg=PR13&dq=Generalized+additive+models&ots=j3-UmlwXmJ&sig=1qxwMbQHaHxAJf-Cgqvy_rGodR0 (visited on 11/17/2015).
- Wood, Simon N. (2011). "Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73.1, pp. 3–36. URL: <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9868.2010.00749.x/pdf> (visited on 11/17/2015).
- Wood, Simon N., Yannig Goude, and Simon Shaw (2015). "Generalized additive models for large data sets". In: *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 64.1, pp. 139–155. ISSN: 1467-9876. DOI: [10.1111/rssc.12068](https://doi.org/10.1111/rssc.12068). URL: <http://onlinelibrary.wiley.com/doi/10.1111/rssc.12068/abstract> (visited on 11/18/2015).
- Breiman, Leo (2001). "Random forests". In: *Machine learning* 45.1, pp. 5–32.
- Liaw, Andy and Matthew Wiener (2002). "Classification and Regression by randomForest". In: *R News* 2.3, pp. 18–22. URL: <http://CRAN.R-project.org/doc/Rnews/>.

- Dice, Lee R. (1945). "Measures of the amount of ecologic association between species". In: *Ecology* 26.3, pp. 297–302. URL: <http://www.jstor.org/stable/1932409> (visited on 09/09/2014).
- Perona, Pietro, Takahiro Shiota, and Jitendra Malik (1994). "Anisotropic diffusion". In: *Geometry-driven diffusion in computer vision*. Springer, pp. 73–92.
- Chang, Winston, Joe Cheng, JJ Allaire, Yihui Xie, and Jonathan McPherson (2015). *shiny: Web Application Framework for R*. URL: <http://CRAN.R-project.org/package=shiny>.
- Muschelli, John, Natalie L Ullman, Elizabeth M Sweeney, Ani Eloyan, Neil Martin, Paul Vespa, Daniel F Hanley, and Ciprian M Crainiceanu (2015b). "Quantitative Intracerebral Hemorrhage Localization". In: *Stroke* 46.11, pp. 3270–3273.
- Jordan, Lori C, Jonathan T Kleinman, and Argye E Hillis (2009). "Intracerebral hemorrhage volume predicts poor neurologic outcome in children". In: *Stroke* 40.5, pp. 1666–1671.

Chapter 5

Quantitative Localization of Supratentorial Intracranial Hemorrhage

5.1 Introduction

Intracranial hemorrhage (ICH) results from a blood vessel rupturing into brain tissues and possibly the ventricles. The classification and quantitative descriptions of hemorrhage location using X-ray computed tomography (CT) is complicated: it may extend into multiple brain areas, distend tissues, and break through the ventricular wall. Therefore, routine practice identifies one primary affected anatomic region (e.g. caudate) (Mendelow et al., 2013; Anderson et al., 2008; Acute Cerebral Hemorrhage (ATACH) investigators, 2010) or describes the hemorrhage in relation to a landmark (Ziai et al., 2014).

Detailed localization information can be obtained, however, by registering scans to a common template where labeled brain atlases are available. After registration, each patient scan is located in the same stereotaxic template space so information may be combined spatially across scans. We registered CT

images to a previously-published CT template (Rorden et al., 2012), located in MNI (Montreal Neurological Institute) space, using the provided software.

Our goal is to quantitatively and objectively characterize ICH location, compare this characterization to human labeling, and predict stroke severity scores using ICH location information.

5.2 Methods

To achieve this goal, we 1) created a 3-dimensional (3D) density map of supratentorial hemorrhages, 2) provided detailed quantification of hemorrhage engagement of individual neuroanatomic regions, 3) performed voxel-wise tests of ICH location on 2 severity scores: Glasgow Coma Scale (GCS) and National Institutes of Health Stroke Scale (NIHSS) and 4) summarized significantly predictive voxel regions into patient-level covariates.

5.2.1 Subjects and Demographics

The population studied consists of 111 patients from the MISTIE (Minimally Invasive Surgery plus recombinant-tissue plasminogen activator for Intracerebral Evacuation) (Morgan et al., 2008) trial recruited from 26 centers with lobar and deep ICHs $\geq 20\text{mL}$ in volume. This sample was 31.5% female with mean (SD) age of 60.8 (11.2) years.

5.2.2 Imaging Data

Standard diagnostic CT images were acquired under a standard protocol but with differences across sites. Scans were acquired using GE ($N=46$),

Siemens ($N=37$), Philips ($N=20$), and Toshiba ($N=8$) scanners, had gantry tilt ($N=87$), and the slice thickness of the image varied within some scans ($N=14$). Therefore, the scans analyzed had different voxel (volume element) dimensions and image resolution prior to registration to the template.

5.2.3 Hemorrhage Segmentation and Location Identification

ICH was manually segmented using the OsiriX (v.4.1, Pixmeo; Geneva, Switzerland) by expert readers. Readers employed a semi-automated approach: a Hounsfield unit range of 40 to 80 selected potential regions of ICH (Bergström et al., 1977; Smith, Rosand, and Greenberg, 2006), then these regions were manually adjusted. Readers identified the anatomic location most engaged by the ICH: putamen ($N=68$), lobar ($N=33$), globus pallidus ($N=6$), and thalamus ($N=4$). The initial mean (SD) ICH volume of this sample was 37.4 (20.1) mL.

5.2.4 Image Registration

The brain image was spatially registered to the CT template using the Clinical toolbox (Rorden et al., 2012) and the statistical parametric mapping software (SPM8, Wellcome Trust Centre for Neuroimaging, London, UK) in MATLAB (Mathworks, Natick, Massachusetts, USA). The binary hemorrhage mask was transformed into the template space. No scans were excluded due to inadequate registration, determined by visual inspection.

5.2.5 ICH Localization and Engagement

We estimated a 3D histogram of ICH localization: for every voxel in template space, we calculated the proportion of patients who have an ICH at that particular voxel.

We calculated spatial ICH engagement by neuroanatomic region using the “Eve” atlas (Oishi et al., 2008), which segments gray matter (GM) and white matter (WM) regions. Ventricular regions were not explicitly segmented in the Eve atlas. Any region not classified was reported as cerebrospinal fluid (CSF).

From this atlas, we estimated ICH engagement at the population level: 1) the percent of the ICH engaged by region (e.g. putamen engages 20% of the ICH) and 2) the percent of each region engaged by the ICH (e.g. ICH engages 78% of the putamen). These summaries of ICH engagement provide a finer description than one location (e.g. putamen).

In the study population, 1,045,174 voxels (denoted by V) had at least one patient with ICH. We limited our analysis to voxels in the template space where at least 10 patients exhibit ICH (166,202 voxels) to optimize the models for a substantial proportion of the study population.

We tested the association between hemorrhage location and stroke severity by running voxel-wise t-tests, testing the null hypothesis $H_0(v) : \mu_1(v) = \mu_0(v)$, where $\mu_1(v)$ corresponds to the mean score (NIHSS or GCS) in patients where $ICH=1$ at voxel v , similarly for $\mu_0(v)$.

We created a patient-level covariate that summarizes ICH location information by selecting regions using a sequence of nested p-value thresholds. We

call these regions “highest predictive regions” (HPR) because they contain the voxels that are most predictive of the severity scores. We obtained 6 different HPR, 3 based on the smallest 1000, 2000, or 3000 lowest p-values and three based on p-values thresholds of .05, .01, and .001. For each HPR, we calculated the HPR “coverage” for scan i :

$$\text{HPR Coverage}_i = \frac{\# \text{ Voxels classified ICH in HPR for scan } i}{\# \text{ Voxels in HPR}} \times 100\%$$

and used this as a predictor of the severity score (Y_i), adjusting for age, sex, and baseline ICH volume (ICHVol):

$$Y_i = \beta_0 + \beta_1 \text{Coverage}_i + \gamma_1 \text{Age}_i + \gamma_2 \text{Sex}_i + \gamma_3 \text{ICHVol}_i + \epsilon_i \quad (5.1)$$

We compared model (5.1) using adjusted R^2 to one using a categorical indicator of the expert-specified ICH location, with categories: thalamus, globus pallidus, putamen, and lobar. A permutation test was used: the severity score was randomly permuted, the HPR was estimated, and the adjusted R^2 was calculated. The p-value was obtained by comparing the distribution of permuted adjusted R^2 with the estimated adjusted R^2 from the HPR model.

5.3 Results

5.3.1 Prevalence of ICH Engagement in the Brain

Figure 5.1 represents the 3D histogram of hemorrhage prevalence, where colors represent the percentage of patients with ICH engagement at that voxel. ICH is distributed medially in the brain in this cohort, with a lower concentration at the cortical surface and higher on the left side of the brain.

The majority of voxels have a low prevalence of ICH engagement; the median number of patients with ICH at a voxel is 3 (3%), though some voxels ($V=5685$) have a high prevalence of $>40\%$ of the sample population.

5.3.1.1 ICH Localization and Engagement

Table 5.1 represents the 10 most-engaged regions for the population 3D histogram. The engagement represents the percent engagement of a specific area compared to all areas engaged. The 3D histogram of ICH is engaged in areas of the insula, putamen, and primarily the CSF (i.e. ventricles and subarachnoid spaces).

We also calculated the engagement of the thalamus, putamen, and globus pallidus by the population 3D histogram. The population engagement represents the mean proportion of ICH prevalence in the population for that brain region. On average, 23% of the putamen, 20% of the globus pallidus, and 8% of the thalamus are engaged with ICH from patients in this study.

Registering images with large deformations caused by the hemorrhage, even those using non-linear registration, will likely have mis-registration artifacts. Lateral shift caused by the hemorrhage, resulting in compression of ventricles, may lead to mapping some areas of the hemorrhage or tissue to the ventricles in template space. Combining multiple registration approaches may achieve better results.

5.3.1.2 Highest Predictive Region Analysis

For illustration, panel (a) in Figure 5.2 displays the region of p-values smaller than .01 for the voxel-wise t-tests of NIHSS and panel (b) displays voxels with the smallest 1000 p-values testing GCS score. These regions were the HPR with the highest adjusted R^2 in (5.1). We show three orthographic slices for each HPR; the entire regions are available in MNI coordinates.

Figure 5.2c displays the NIHSS score as a function of the coverage of the HPR from Figure 5.2a. Similarly, Figure 5.2d displays the GCS score as a function of the coverage of the HPR in Figure 5.2b. The blue line represents a non-parametric LOESS (local regression) (Cleveland, Grosse, and Shyu, 1992) fit and the red line represents an unadjusted linear model fit. As expected, the larger the HPR coverage the higher (more severe stroke) the NIHSS score and the lower (deeper unconsciousness) the GCS score.

Adjusted R^2 model estimates indicated all HPR coverage models strongly outperform reader-classified location models: the adjusted R^2 almost doubled from 0.129 for the reader-classified location model to 0.254 for NIHSS; the adjusted R^2 more than tripled from 0.069 for the reader-classified location model to 0.214 for the HPR coverage model for GCS.

The permutation test p-values were $< .001$ for HPR of NIHSS and $< .01$ for HPR of GCS, indicating that the selected HPRs were more predictive than HPRs obtained using the same selection procedure when there is no association between location and stroke severity scores.

5.4 Discussion

We have characterized the localization of ICH in a population from prospective clinical trial images using a 3D histogram. We found the well-described medial location of most supratentorial ICHs. We can create 3D histograms based on subgroups or different study populations and test differences between groups at a voxel level using proportion tests, allowing a fine-scale comparison ICH location across groups.

We also demonstrate how labeled atlases can automatically describe ICH engagement by neuroanatomic regions at a patient or population level. These measures are more interpretable for clinical relevance and may translate to better determination of disability.

5.4.1 Summary

The summary of Eve-atlas ICH engagement presented provided a finer description of location than previously possible by human readers. This type of analysis provides a framework for derivation and testing measures of ICH engagement. We hope this method will engage others with larger data sets and methodological skills to enhance the use of quantitative localization.

5.5 Sources of Funding

The project described was supported by the grant RO1EB012547 (NIH), T32AG000247 (NIA), R01NS046309, RO1NS060910, RO1NS085211, R01NS046309, U01NS080824 and U01NS062851 (NINDS), and RO1MH095836 (NIMH).

5.6 Disclosures

Johns Hopkins University holds a use patent for intraventricular tissue plasminogen activator.

5.7 Tables

Area	Population Prevalence
CSF (ventricular & subarachnoid spaces)	7.9
Insular	7.6
Superior temporal gyrus	5.5
Putamen	4.8
External capsule	3.9
Superior corona radiata	3.7
Precentral gyrus	3.3
Precentral WM	3.1
Superior temporal WM	3.1
Posterior limb of internal capsule	3.0

Table 5.1: Distribution of the Top 10 Areas of Engagement of Highest Predictive Regions.

5.8 Figures

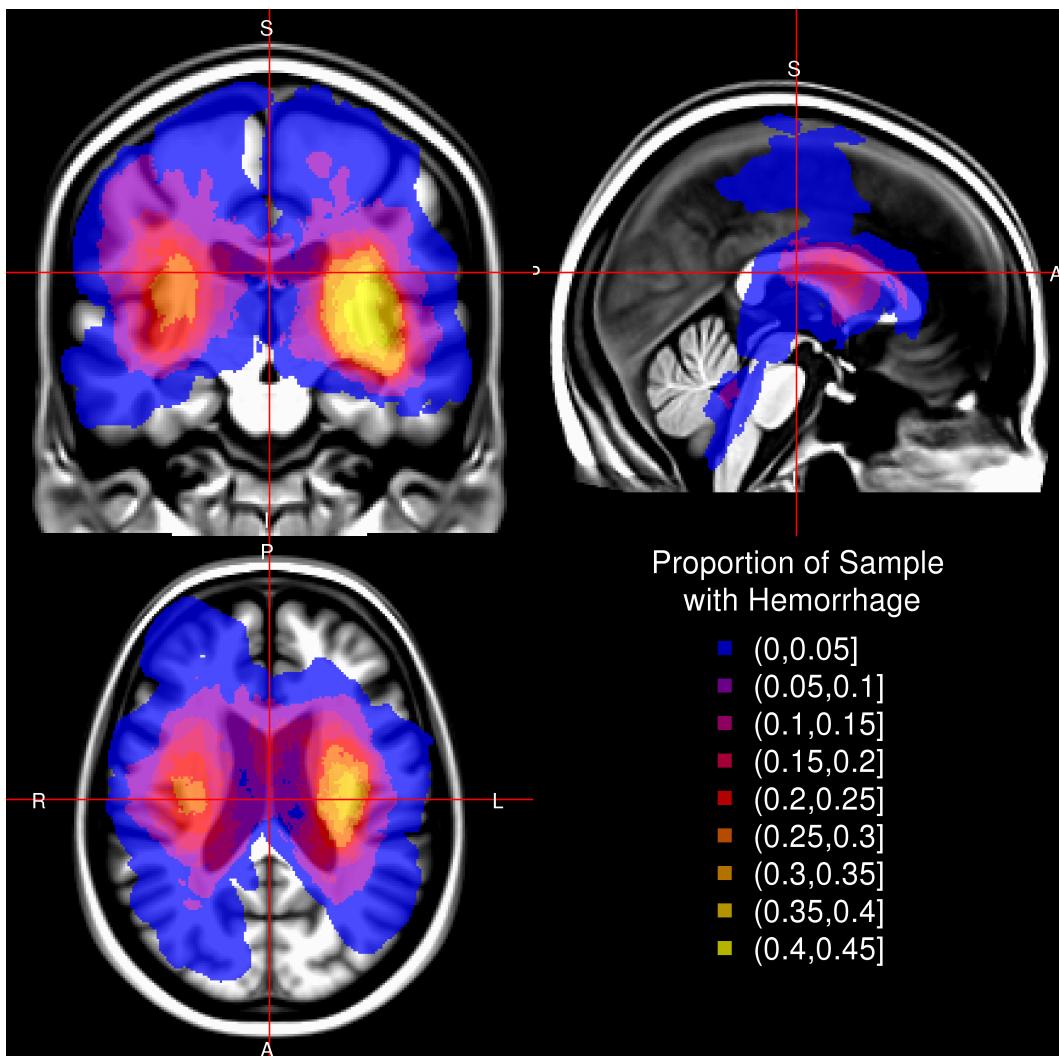


Figure 5.1: ICH Engagement Prevalence. The proportion of patients with ICH engaging a given voxel are represented in a 3D histogram (right side of image is left side of brain) overlaid on an MRI T1 template. There is a higher prevalence of ICH on the left side of the brain, localized in the middle of the brain, with few extensions in the anterior and posterior areas. The interactive version of this figure is located at http://muscchellij2.github.io/CT_Pipeline/index.html.

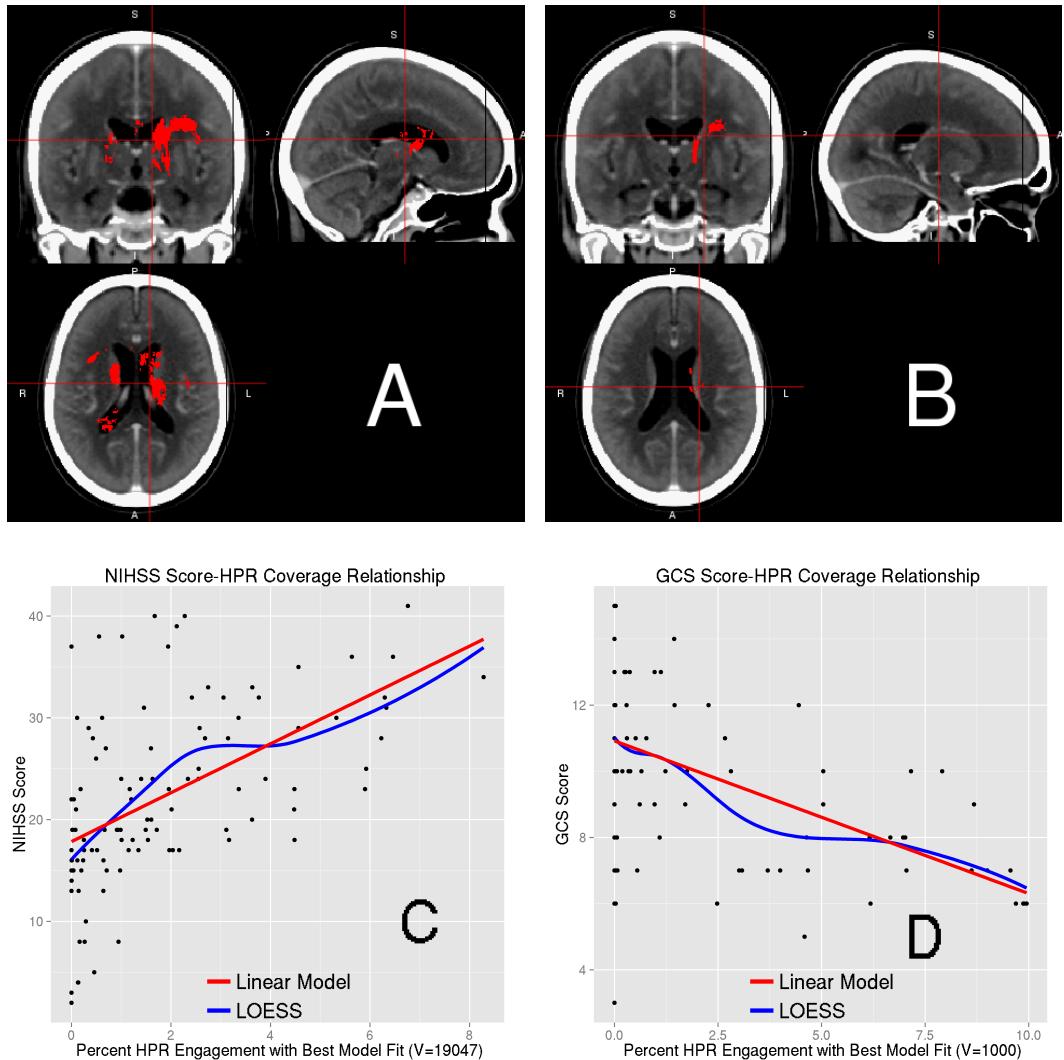


Figure 5.2: Highest Predictive Region (HPR) Analysis. Panels (a) and (b) correspond to the HPR for the top-performing model for NIHSS and GCS scores. The HPR in (a) represents a p-value threshold of .0100 (19047 voxels) for the voxel-wise p-value of ICH on NIHSS. The HPR in (b) represents 1000 with the lowest p-values for the voxel-wise ICH on GCS score regressions. Panels (c) and (d) plot the relationship of the HPR coverage and severity score. The red line represents a linear fit, the blue line—a LOESS fit. The larger the HPR coverage the higher (more severe) the NIHSS and the lower (deeper unconsciousness) the GCS.

References

- Mendelow, A David, Barbara A Gregson, Elise N Rowan, Gordon D Murray, Anil Ghokar, and Patrick M Mitchell (2013). "Early surgery versus initial conservative treatment in patients with spontaneous supratentorial lobar intracerebral haematomas (STICH II): a randomised trial". In: *The Lancet* 382.9890, pp. 397–408. ISSN: 0140-6736. DOI: [10.1016/S0140-6736\(13\)60986-1](https://doi.org/10.1016/S0140-6736(13)60986-1). URL: <http://www.sciencedirect.com/science/article/pii/S0140673613609861> (visited on 02/26/2014).
- Anderson, Craig S., Yining Huang, Ji Guang Wang, Hisatomi Arima, Bruce Neal, Bin Peng, Emma Heeley, Christian Skulina, Mark W. Parsons, Jong Sung Kim, and others (2008). "Intensive blood pressure reduction in acute cerebral haemorrhage trial (INTERACT): a randomised pilot trial". In: *The Lancet Neurology* 7.5, pp. 391–399. URL: <http://www.sciencedirect.com/science/article/pii/S1474442208700693> (visited on 07/08/2014).
- Acute Cerebral Hemorrhage (ATACH) investigators, Antihypertensive Treatment of (2010). "Antihypertensive treatment of acute cerebral hemorrhage *. [Miscellaneous Article]". In: *Critical Care Medicine February 2010* 38.2, pp. 637–648. ISSN: 0090-3493. DOI: [10.1097/CCM.0b013e3181b9e1a5](https://doi.org/10.1097/CCM.0b013e3181b9e1a5). URL: <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=ovftk&AN=00003246-201002000-00038> (visited on 07/08/2014).
- Ziai, Wendy C., Stanley Tuhrim, Karen Lane, Nichol McBee, Kennedy Lees, Jesse Dawson, Kenneth Butcher, Paul Vespa, David W. Wright, Penelope M. Keyl, A. David Mendelow, Carlos Kase, Christine Wijman, Marc LaPointe, Sayona John, Richard Thompson, Carol Thompson, Steven Mayo, Pat Reilly, Scott Janis, Issam Awad, Daniel F. Hanley, and CLEAR III Investigators (2014). "A multicenter, randomized, double-blinded, placebo-controlled phase III study of Clot Lysis Evaluation of Accelerated Resolution of Intraventricular Hemorrhage (CLEAR III)". In: *International Journal of Stroke* 9.4, pp. 536–542. ISSN: 1747-4949. DOI: [10.1111/ijjs.12097](https://doi.org/10.1111/ijjs.12097). URL:

<http://onlinelibrary.wiley.com/doi/10.1111/ijjs.12097/abstract> (visited on 11/18/2014).

- Rorden, Christopher, Leonardo Bonilha, Julius Fridriksson, Benjamin Bender, and Hans-Otto Karnath (2012). "Age-specific CT and MRI templates for spatial normalization". In: *NeuroImage* 61.4, pp. 957–965. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2012.03.020](https://doi.org/10.1016/j.neuroimage.2012.03.020). URL: <http://www.sciencedirect.com/science/article/pii/S1053811912002935> (visited on 01/15/2014).
- Morgan, T., M. Zuccarello, R. Narayan, P. Keyl, K. Lane, and D. Hanley (2008). "Preliminary findings of the minimally-invasive surgery plus rtPA for intracerebral hemorrhage evacuation (MISTIE) clinical trial". In: *Cerebral Hemorrhage*. Springer, pp. 147–151. URL: http://link.springer.com/chapter/10.1007/978-3-211-09469-3_30 (visited on 09/09/2014).
- Bergström, M, K Ericson, B Levander, P Svendsen, and S Larsson (1977). "Variation with time of the attenuation values of intracranial hematomas". In: *Journal of computer assisted tomography* 1.1, pp. 57–63. ISSN: 0363-8715.
- Smith, Eric E., Jonathan Rosand, and Steven M. Greenberg (2006). "Imaging of Hemorrhagic Stroke". In: *Magnetic Resonance Imaging Clinics of North America* 14.2, pp. 127–140. ISSN: 1064-9689. DOI: [10.1016/j.mric.2006.06.002](https://doi.org/10.1016/j.mric.2006.06.002). URL: <http://www.sciencedirect.com/science/article/pii/S1064968906000298> (visited on 03/06/2014).
- Oishi, Kenichi, Karl Zilles, Katrin Amunts, Andreia Faria, Hangyi Jiang, Xin Li, Kazi Akhter, Kegang Hua, Roger Woods, Arthur W. Toga, G. Bruce Pike, Pedro Rosa-Neto, Alan Evans, Jiangyang Zhang, Hao Huang, Michael I. Miller, Peter C. M. van Zijl, John Mazziotta, and Susumu Mori (2008). "Human brain white matter atlas: Identification and assignment of common anatomical structures in superficial white matter". In: *NeuroImage* 43.3, pp. 447–457. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2008.07.009](https://doi.org/10.1016/j.neuroimage.2008.07.009). URL: <http://www.sciencedirect.com/science/article/pii/S1053811908008240> (visited on 03/19/2014).
- Cleveland, William S., Eric Grosse, and William M. Shyu (1992). "Local regression models". In: *Statistical models in S*, pp. 309–376. URL: ftp://plan9mail.cs.bell-labs.com/cm/ms/departments/sia/doc/smoothing_econhb.ps (visited on 09/11/2014).

Chapter 6

Validated Automatic Brain Extraction of Head CT Images

6.1 Introduction

X-ray computed tomography (CT) scanning of the brain is widely available and is a commonly used diagnostic tool in clinical settings (Sahni and Weinberger, 2007; Chalela et al., 2007; Schellinger et al., 1999). 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 and 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 contain non-brain human tissues and 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. (2014) for an overview of methods). While an extensive literature accompanied by software exist for brain MRI scans, the same is not true for brain CT scans. Smith (2002) introduced and validated the Brain Extraction Tool (BET), a function of the FSL (Jenkinson et al., 2012) neuroimaging software (v5.0.4), to automatically extract the brain from MRI scans. Here we propose to adapt BET and validate its brain extraction performance for CT scans.

BET adaptations for this purpose have been presented before in Solomon et al. (2007). Although the method is similar to that outlined below, using thresholding and then applying BET, the authors did not publish the specific details of the method nor any code to evaluate it. To replicate the method from Solomon et al. (2007), Rorden et al. (2012) thresholded voxels to be under 100 Hounsfield units, manually adjusted the image intensity to enhance the soft tissue in the brain, and then BET was applied with a fractional intensity of 0.35. Therefore, brain extraction methods from Rorden et al. (2012) and Solomon et al. (2007) strongly parallels the proposed method described below, but neither studies presented a formal validation against a set of manually segmented brain images.

Mandell et al. (2014a) has recently proposed a brain extraction method for CT scans and has done a validation against manual segmentation. This method was also performed on a set of brains with disease (Mandell et al., 2014b; Mandell et al., 2014c). This method is not fully automated, however. Mandell et al. (2014a) has formally validated a brain extraction method against

manually segmented images, but the method requires user interaction.

Thus, the goals of our study are to propose an automated method that has been formally validated against a set of manually segmented images and estimate brain extraction performance of this method in a large number of CT scans.

6.2 Methods

6.2.1 Participants and CT data

We used CT images from patients enrolled in the MISTIE (Minimally Invasive Surgery plus recombinant-tissue plasminogen activator for Intracerebral Evacuation) and ICES (Intraoperative CT-Guided Endoscopic Surgery) stroke trials (Morgan et al., 2008). Inclusion criteria into the study included: 18 to 80 years of age, spontaneous supratentorial intracerebral hemorrhage above 20 milliliters (mL) in size (for full criteria, see Mould et al. (2013)). The population analyzed here had a mean (SD) age was 60.6 (11.6) years, was 66.9% male, and was 55.6% Caucasian, 30.1% African American, 9.8% Hispanic, and 4.5% Asian or Pacific islander. CT data were collected as part of the Johns Hopkins Medicine IRB-approved MISTIE research studies with written consent from participants.

6.2.2 Imaging Data

6.2.2.1 Validation of Automated Head Segmentation

For the validation of automated segmentation against gold standard manual segmentation, we analyzed 36 scans, corresponding to 36 unique patients. The study protocol was executed with minor, but important, differences across the 19 sites. Scans were acquired using Siemens ($N = 14$), GE ($N = 11$), Philips ($N = 10$), and Toshiba ($N = 1$) scanners. Gantry tilt was observed in 21 scans. Slice thickness of the image varied within the scan for 7 scans. For example, a scan may have 10 millimeter (mm) slices at the top and bottom of the brain and 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.

6.2.3 Manual and Automated Brain Extraction

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 (reader 1: NU). CT brain images and the binary brain tissue mask obtained using manual segmentation were exported from OsiriX to DICOM format.

6.2.4 Image Processing

The image processing pipeline is provided in Figure 6.1. Images with gantry tilt were corrected using a customized MATLAB (The Mathworks, Natick,

Massachusetts, USA) user-written script (<http://bit.ly/1ltIM8c>). Although gantry tilt correction is not inherently necessary for brain extraction, it is required for rigid co-registration of scans within a patient, which is a common processing step in longitudinal analysis of images post brain extraction.

Images were converted to the Neuroimaging Informatics Technology Initiative (NIfTI) data format using dcm2nii (2009 version, provided with MRIcro (Rorden and Brett, 2000)). Images were constrained to values between –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 converted slice and the NIfTI format assumes homogeneous thickness throughout the image. This loss of information, if not properly accounted for, affects volume estimation, which relies on accurate pixel dimensions in millimeters. Variable slice thickness should have no affect on the other estimates of performance described below as they are calculated at a voxel level and do not rely on pixel resolution. Although the NIfTI images store the data with only one pixel dimension for the height of the voxel, we use the ImagePositionPatient DICOM field to determine the accurate height of each voxel to calculate an accurate volume.

Each image was thresholded using the brain tissue range (0 – 100 HU); voxels outside this range were set to 0 HU. In one variant of the pipeline, data were smoothed using a 3-dimensional (3D) Gaussian kernel ($\sigma = 1\text{mm}^3$) and re-thresholded to 0 – 100 HU; in the other, data were not smoothed. BET was then applied, varying the fractional intensity (FI) parameter to determine its influence on performance: we used values of 0.35 (as used in Rorden et al.

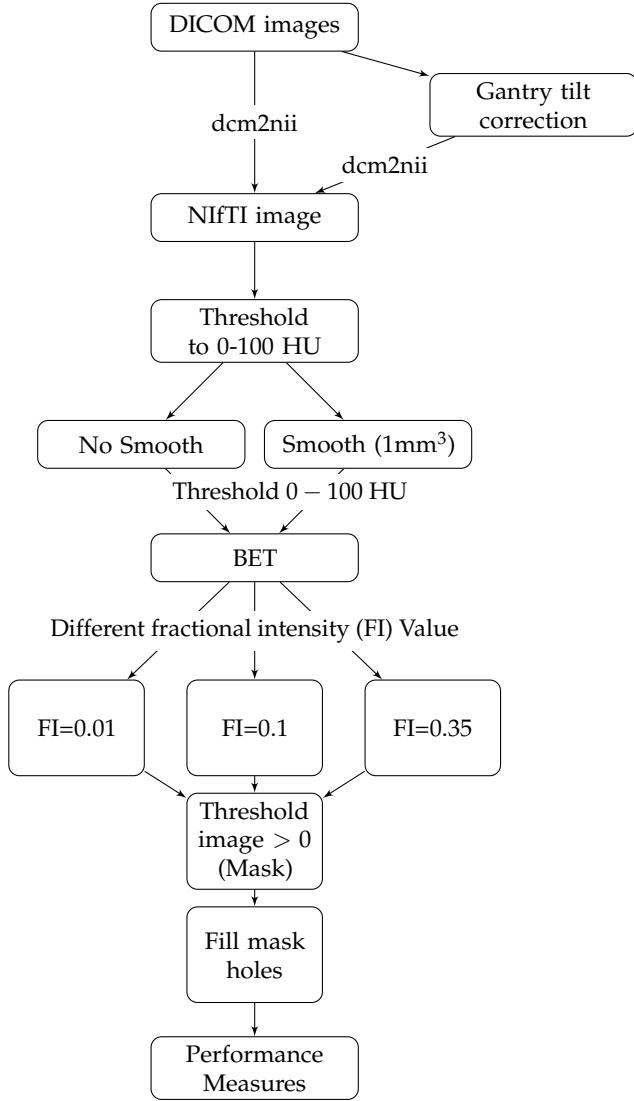


Figure 6.1: Brain Extraction Processing Pipeline. Images in DICOM (Digital Imaging and Communications in Medicine) format were gantry tilt corrected if necessary and converted to NIfTI (Neuroimaging Informatics Technology Initiative) format using dcm2nii. After NIfTI conversion, the data is thresholded to tissue ranges of 0 – 100 Hounsfield units (HU). In one variant of the pipeline, the data was smoothed using a 3-dimensional Gaussian kernel ($\sigma = 1\text{mm}^3$) and re-thresholded to 0 – 100 HU; in the other, the data was not smoothed. BET was applied to the image using 3 different fractional intensity (FI) values: 0.01, 0.1, and 0.35. The resultant image was masked to values greater than 0 HU and FSL was used to fill in any holes. These filled masks were used in comparison to the manually segmented image.

(2012)), 0.1, and 0.01.

The FI parameter varies between 0 and 1 and determines the location of the edge of the segmented brain image; smaller values correspond to larger brain masks. Smith (2002) describes that the FI parameter determines a local threshold t_l by the following equation:

$$t_l = (I_{\max} - t_2) \times \text{FI} + t_2$$

where I_{\max} is a local maximum intensity along a line from an outer surface vertex pointing inward to the image center and t_2 is the 2nd percentile of the image distribution. As a result of thresholding in our pipeline, t_2 equals 0 HU and I_{\max} must lie between 0 and 100 HU. Therefore, after thresholding,

$$t_l = I_{\max} \times \text{FI}.$$

With an I_{\max} of 100 HU, using an FI lower than 0.01 results in t_l less than 1 HU, but greater than 0 HU. As CT data is stored as integers, no intensities lie between 0 and 1, so a t_l between 0 or 1 HU should provide similar local thresholds. Therefore, we chose 0.01 as a lower limit for testing FI.

After BET was applied, we created a brain mask taking values > 0 HU and filled the holes in the mask (using `fslmaths -fillh`).

6.2.5 Measuring and Testing Brain Extraction Performance

We compared the masks obtained using the various choices of parameters to the manually segmented images. Four common measurements of performance were calculated for each image: sensitivity, specificity, accuracy, and the Dice

Similarity Index (DSI) (Dice, 1945). For each measure, higher values indicate better agreement with the manual segmentation. See Inline Supplementary Methods 1 for the calculation of each measure.

[Insert Supplementary Methods 1 here]

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

From each scan, we also calculated the intracranial volume (ICV), defined as all voxels inside the skull, by multiplying the number of voxels in the resulting mask by the dimensions of each voxel. We calculated the ICV ratio comparing manual and automated segmentation: $\frac{\text{ICV}_{\text{automated}}}{\text{ICV}_{\text{manual}}}$. A ratio of 1 indicates the same volume; greater than 1 indicates over-estimation of ICV; less than 1 indicates underestimation of ICV. As adjustment for ICV has been shown to reduce inter-subject variation in volumetric studies (Whitwell et al., 2001), we wish to estimate ICV accurately.

6.2.6 Consistency of Manual Brain Extraction

As manual segmentation can have intra-reader variability, another reader (reader 2: AM) manually segmented brain tissue on the 36 scans. We additionally estimated all four performance measurements, using the manual segmentation from reader 2 as the gold standard. We also estimated the ICV from the segmentation from reader 2. We calculated the ICV ratio $\left(\frac{\text{ICV}_{\text{reader 2}}}{\text{ICV}_{\text{reader 1}}}\right)$ and the correlation of ICV estimates across readers.

6.2.7 Failure Rate and Intraclass Correlation Estimate

Although comparison of automated methods to a manual gold standard is ideal, manual segmentation requires a significant amount of time. Therefore, for a large number of scans, this procedure is impractical. As multiple CT scans are obtained from patients in the MISTIE trial, we can estimate the reliability of our proposed brain extraction pipelines without manual segmentation by comparing intracranial volumes of the same patient on subsequent scans. Moreover, we can estimate failure rate of each pipeline.

For these tasks, we collected 1160 scans. Of these scans, we excluded 27 scans due to craniotomy and 38 due to the gantry tilt correction forcing areas of the brain outside the field of view. We executed the previous brain extraction pipelines on the remaining 1095 scans. Of these scans, we visually assessed the quality of brain extraction: any scan excluding a significant portion of the brain or having holes due to mask self-intersection were classified as a failure. These scans represent 129 patients from 26 sites, with a mean (SD) of 8.5 (2.8) scans per patient. Scans were acquired using Siemens ($N = 492$), GE ($N = 298$), Philips ($N = 207$), Toshiba ($N = 66$), Neurologica ($N = 30$), and Picker ($N = 2$) scanners. We estimated the failure rate for each processing pipeline and used a Fisher's exact test to test whether failure rates differed across scanners.

For each scan, we calculated the ICV. Using only the scans with successful brain extraction, we estimated the intraclass correlation (ICC) and its confidence interval (CI) of ICV using a one-way ANOVA, where a patients was treated as a group, for unbalanced repeated measures (Searle,

2012; Thomas and Hultquist, 1978; Donner, 1979; Lessells and Boag, 1987) using the ICC package (Wolak, Fairbairn, and Paulsen, 2012) in R (<http://cran.r-project.org/>).

6.3 Results

6.3.1 Manual and Automated Brain Extraction

The following estimates use the manual segmentation from reader 1 as the gold standard. Figure 6.2a illustrates the performance of each variation of the BET pipeline in Figure 6.1. The pipelines using smoothing (top panel) perform better than the unsmoothed pipelines (bottom panel) on all measures except specificity (all $p < 0.01$, uncorrected for multiplicity). BET also performed poorly on some scans without smoothing.

Figure 6.2b displays the performance for brain extraction for the pipelines using smoothed images. Because the performance for all metrics was high when using smoothed images, it was necessary to 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; thus, we will focus and compare results for these values of FI only for the case when BET was applied to smoothed images. Using an FI of 0.01 had a higher median sensitivity (0.9901) than an FI of 0.1 (0.9884, median difference: 0.0014, $p < 0.001$), accuracy (0.9971 vs. 0.9971; median difference: 0.0001, $p < 0.001$), and DSI (0.9895 vs. 0.9894; median difference: 0.0004, $p < 0.001$) and lower specificity (0.9981 vs. 0.9982; median difference: -0.0001, $p < 0.001$). Overall, regardless of p-values, these measures are all very high, indicating that multiple choices of parameters work well for brain extraction after CT

image processing. Moreover, a Bonferroni correction for multiple comparisons yields the same conclusions.

The mean (SD) ICV ratio was 1.002 (0.0079) using an FI of 0.01 and 1 (0.0081) using an FI of 0.1. Both mean ratios are close to 1 with a small variance, indicating the ICV estimates are similar for automated and manual segmentation.

The above results indicate that using smoothed data and an FI of 0.01 or 0.1 had high performance when compared to the manual segmentation of reader 1. The results were similar using the scan-wise union of the segmentation from reader 1 and reader 2. Using the manual segmentation from reader 2 or the scan-wise intersection of the segmentation from reader 1 and reader 2, the median values using 0.1 had higher marginally performance than using 0.01 for DSI, accuracy, and specificity, but lower performance for sensitivity. See Inline Supplementary Figure 1 for the distribution of performance metrics for each segmentation.

[Inline Supplementary Figure 1]

Regardless of which manual segmentation was used, estimates of performance for each scan using smoothed data and an FI of 0.01 or 0.1 remained above 0.95. Thus, these pipelines perform well, yet one FI may not perform universally better than the other.

6.3.2 Consistency of Manual Brain Extraction

When comparing manual segmentations, we used the manual segmentation from reader 1 as the gold standard and the segmentation from reader 2 as the

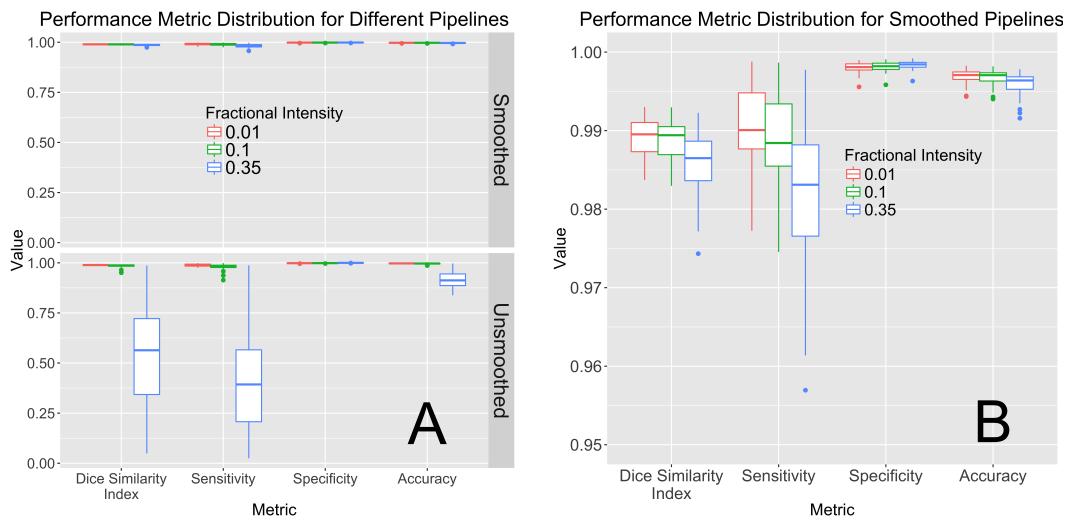


Figure 6.2: Performance Metric Distribution for Different Pipelines. Panel a displays the boxplots for performance measures when running the pipeline with a different fractional intensity (FI), using smoothed data (top) or unsmoothed data (bottom). Panel b presents the smoothed data only, rescaled to 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 has high performance on all measures.

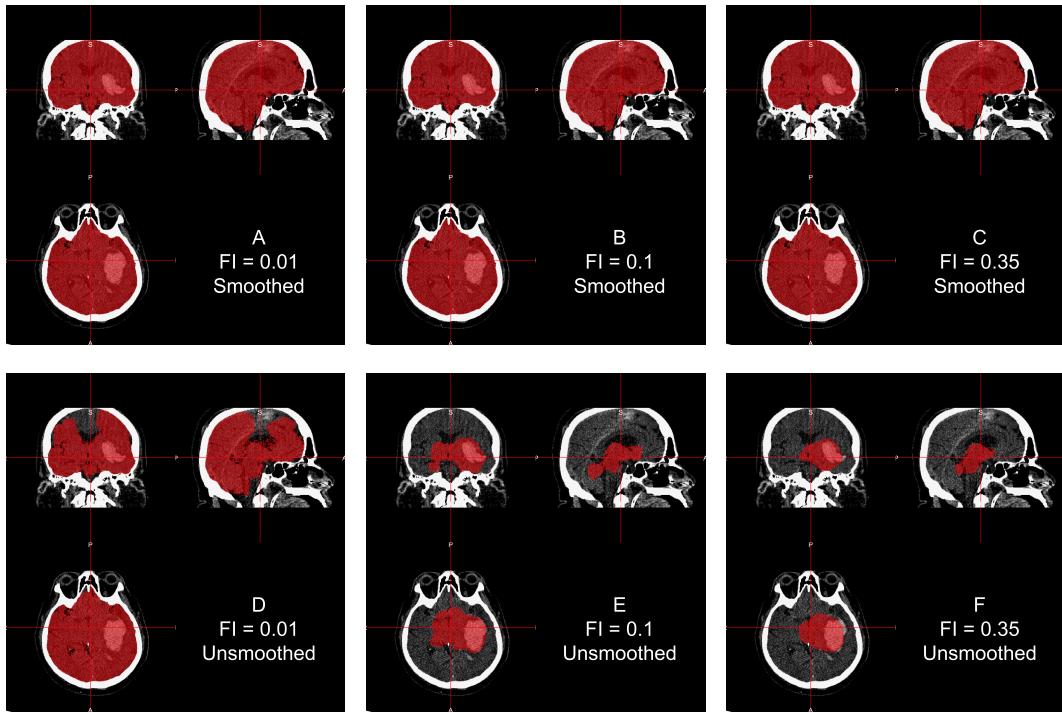


Figure 6.3: Example Case where Smoothing before BET is Required. For one subject, the CT image is displayed with the brain extracted mask in red after running all pipelines. Panels **a**, **b**, and **c** represent applying BET using FI of 0.01, 0.1, and 0.35, respectively, to smoothed data. Panels **d**, **e**, and **f** correspond to applying BET using FI 0.01, 0.1, and 0.35 on unsmoothed data. Smoothing images improves brain extraction with BET.

test segmentation similar to the automated segmentation above. The mean (SD) was 0.989 (0.0030) for DSI, 0.997 (0.0010) for accuracy, 0.982 (0.0060) for sensitivity, and 0.999 (0.0003) for specificity.

The estimated mean (SD) ratio of the ICV was 0.988 (0.0068). The correlation (95% confidence interval) of ICV was 0.998 (0.997, 0.999). See [Inline Supplementary Figure 2](#) for the comparison of ICV estimates from reader 1 and reader 2.

[[Inline Supplementary Figure 2](#)]

Overall, we observe high agreement of segmentation between raters and the estimates of performance in automated segmentation to be similar to multiple reader segmentation. Differences between manual segmentation occurred on the boundary between bone and non-bone areas towards the surface of the cortex and inferior regions of the brain, where one may or may not classify areas as spinal cord and not part of the brain stem. Thus, the difference observed in the performance of FI of 0.01 compared to 0.1 when using a reader 2 as the gold standard are likely due to these areas.

6.3.3 Failure Rate and Intraclass Correlation Estimate

Although Figure 6.2 indicates that using FI of 0.01 or 0.1 provides adequate brain extraction results for the cases analyzed, they perform relatively well regardless whether or not the data are smoothed. Figure 6.3 displays an example where using unsmoothed data performs poorly for these FIs, demonstrating why smoothing may be necessary for some scans. This is a high-resolution scan, with voxel size $0.49\text{mm} \times 0.49\text{mm} \times 1$, which may result in more noise in the image that may affect the performance of BET. Moreover, in Table 6.1, the estimated failure rates were lower using the smoothed data compared to the unsmoothed data. We observe the lowest failure rate in the pipelines using smoothed data and an FI of 0.01 or 0.1. Though this represents a large number of scans, failure rates may also be affected by patient-level characteristics, including the center where the patient was scanned.

As multiple scanners were used, we wanted to determine if the failure rate was different across scanners. In Table 6.2, we present the failure rate

Fractional Intensity	Failure Scans: N (%)	
	Unsmoothed	Smoothed
0.01	161 (14.7%)	57 (5.2%)
0.1	192 (17.5%)	80 (7.3%)
0.35	1068 (97.5%)	154 (14.1%)

Table 6.1: Failure Rates for each Processing Pipeline of Brain Extraction of the 1095 Scans Analyzed.

for each scanner, using smoothed data and an FI of 0.01. The failure rates for all scanner types other than Neurologica were below 6%. Although we see a failure rate above 16% in the Neurologica scanners, only 30 scans were acquired using this type of scanner. Moreover, a Fisher's exact test did not find a difference in the failure rates across scanner type ($p = 0.110$).

Scanner Type	Failure Rate: Fail/N (%)
Siemens	28/492 (5.7%)
GE	15/298 (5.0%)
Philips	7/207 (3.4%)
Toshiba	2/66 (3.0%)
Neurologica	5/30 (16.7%)
Picker	0/2 (0.0%)

Table 6.2: Failure Rates for Different Scanner Types using Smoothed Data and an FI of 0.01 Processing Pipeline of Brain Extraction of the 1095 Scans Analyzed.

The ICC estimate was high using the successfully brain extracted scans from the smoothed data with an FI of 0.01 (ICC: 0.929, 95% CI: 0.91, 0.945) and 0.1 (ICC: 0.928, 95% CI: 0.909, 0.944). In Figure 6.4, we illustrate the ICV estimates, using an FI of 0.01 and smoothed data, for successful brain extraction in scans 10 or fewer days post baseline scan (gray lines). The black lines represent ICV estimates over time for 10 randomly selected patients. The blue line is a local regression (LOESS) (Cleveland, Grosse, and Shyu, 1992)

line, which represents an estimate of the average ICV over time. This LOESS line is relatively flat, indicating that the ICV estimate averaged over patients is stable. We also observe that although within-patient variability exists for ICV estimates, the variability across patients is greater.

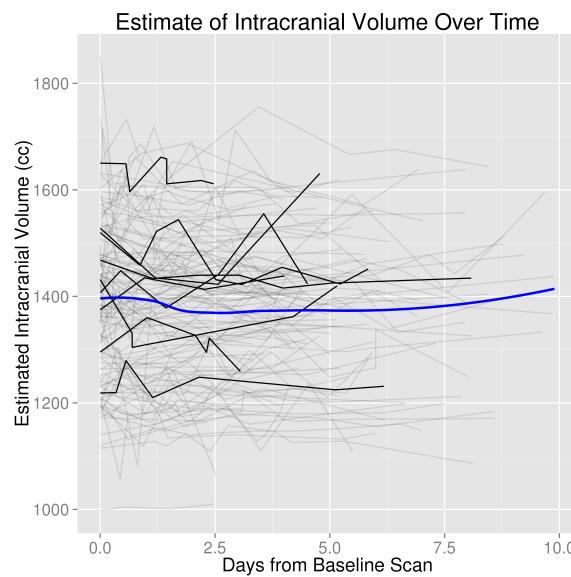


Figure 6.4: Intracranial Volume (ICV) Estimates for Scans Less than 10 Days Post-Baseline. These estimates are from the brain extraction pipeline using and FI of 0.01 and smoothed data. Each separate line represents an individual patient. The black lines represent ICV estimates over time for 10 randomly selected patients. The blue line is a local regression (LOESS) scatterplot smoother.

6.4 Discussion

Quantitative procedures based on data contained in head CT images require brain-only images for analysis. We have introduced the first validated, fully automated brain extraction pipeline for head CT images based on widely used, existing software. Validation was done using gold-standard manual segmentations of brain tissue and multiple measurements of intracranial

volume per patient. A novel finding is that smoothing the data, as opposed to other studies which have smoothed at other points in the process (Mandell et al., 2014a), using a conservative smoother (1mm^3 3D Gaussian kernel) and using an FI of 0.01 or 0.1 provides good brain extraction for the sample studied. These choices make a large difference in the performance of the algorithms and have not been previously reported in the literature.

Although the sample size was relatively small for the gold standard validation, the CT images used are from different people, different centers, and different scanners. We have also shown that failure rates are low (5%) using smoothed data and an FI of 0.01 in a large number of scans. We are using a population of patients with intracranial hemorrhage and the accuracy of BET may be dependent on factors such as hematoma size, which may change the distribution of Hounsfield units or compress brain structures. We observed good performance of BET in these patients using the parameters described, which may indicate even better performance for individuals with no observed pathology. BET also performs well in scans from follow-up scans with no hemorrhage included in the longitudinal ICV estimates; these scans should be more similar to scans from patients without hemorrhage.

We did not, however, rigorously test this pipeline against a set of different levels of noise (as was done using MRI in (Mandell et al., 2014a)), convolution filters, scanning artifacts, or scanning parameters. As noted in Mandell et al. (2014a), there is no standard ground truth for CT scans. However, as these scans are those presented for evaluation in many diagnostic cases, we believe these do represent clinically relevant scans that would be analyzed. As the

failure rates are low, we believe this method may be adapted for different settings of acquisition for CT.

A percentage of scans in this study did not have successful brain extraction, however. Additional steps may be done to decrease the failure rate of the proposed method: 1) using registration to a CT template (Rorden et al., 2012) to remove slices of the neck below the head can achieve better brain-center estimation used by BET; 2) performing BET using a higher smoothness constraint that may reduce potential holes in the brain mask caused by mask intersection; and 3) using a dilation followed by an erosion operation to fill any holes that are caused by thresholding but not filled by the hole filling operation above.

One additional concern is how general this method is to other populations as the data is from an older adult population, such as the pediatric population (Mandell et al., 2014b) analyzed. Gousias et al. (2008) analyzed 33 2-year-old children who had been born prematurely and found that after preprocessing of co-registration and neck removal, and dilation, BET was adequate for brain extraction in MR images, and then brain labeling. Moreover, Shi et al. (2012) demonstrated BET performed well in automated brain extraction for pediatric MR images of neonates < 2 months ($N = 90$), 1-2 year infants ($N = 141$) and 5-18 year old children ($N = 60$). Although CT scans may have differences compared to MRI, we believe this method should be robust at least to children, but would like to validate our method on additional populations.

Overall, good performance using CT acquired under different scanners and different scanning parameters indicate that the approach described here

will likely generalize in addition to the fact that CT scan data are expressed in standardized units (Hounsfield units). Moreover, the robust success of BET as a method is another indicator that our proposed method has a high likelihood of generalizability.

After creating an accurate brain mask, secondary image processing or estimation steps can be performed. These include intensity normalization, segmentation, and image registration. Moreover, ICV estimates can be used as potential factor for adjustment in analysis (Whitwell et al., 2001). Additionally, extraction of structures within the brain, such the cerebrospinal fluid (CSF), which can be estimated from the method described in Volkau, Puspitasari, and Nowinski (2010), may have fewer errors if performed on the segmented image as the process is not computed over voxels outside of the brain. We believe that successful brain extraction is fundamental for calculating quantitative measures on the brain and performing necessary secondary operations required for analysis.

The research presented here is fully reproducible and we provide ready-to-use software for CT brain extraction. The R function designed to perform brain extraction is located at http://bit.ly/CTBET_RCODE and example bash script for command-line FSL can be downloaded here http://bit.ly/CTBET_BASH. As our software is publicly available and is based on open-source, free programs (FSL and R), the proposed method is readily available to all users.

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 U01NS062851 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 $I_{im} = 0$. Let the total number of voxels be denoted by V . The number of true positive voxels is defined as:

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

Sensitivity is defined as

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

specificity is defined as

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

overall accuracy is defined as:

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

and the Dice Similarity Index (DSI) is defined as

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

Inline Supplementary Figure 1

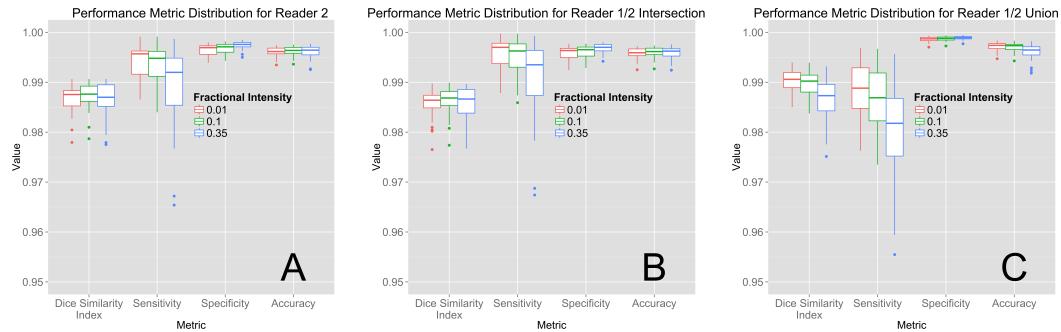


Figure 6.5: We display the boxplots for performance measures of the automated segmentation when using smoothed data with different fractional intensity (FI) with the gold standard being the manual segmentation from reader 2 (a), scan-wise intersection of the manual segmentation from reader 1 and reader 2 (b), or scan-wise union of the manual segmentation from reader 1 and reader 2 (c). Overall, using an FI of 0.01 and 0.1 perform high on all measures, regardless of manual segmentation used as the gold standard.

Inline Supplementary Figure 2

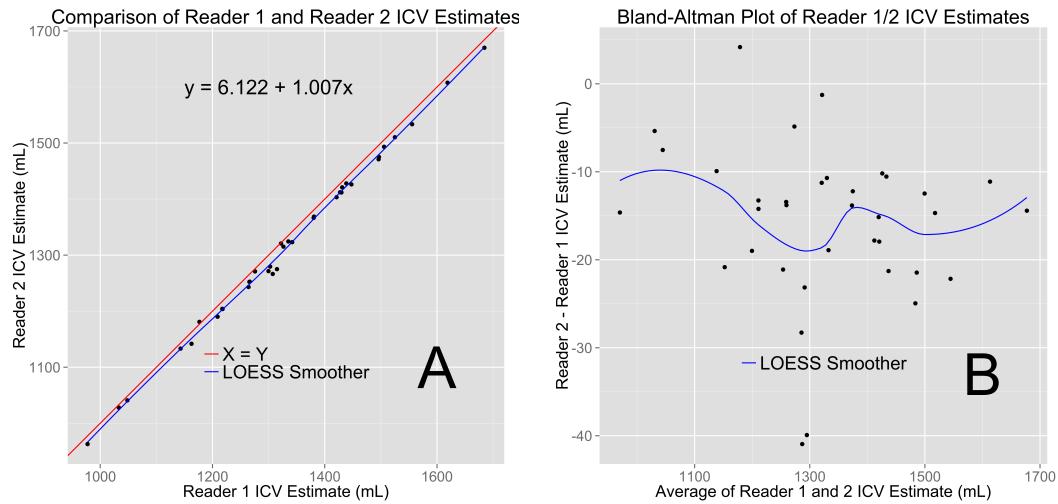


Figure 6.6: Panel **a** displays the intracranial volume (ICV) estimate from the manual segmentation of reader 1 versus reader 2. The blue line represents a LOESS scatterplot smoother of the data. The red line represents a linear fit. The slope is approximately 1 and the intercept is approximately 6 mL, indicating strong agreement of the estimates. The Bland-Altman plot in panel **b** denotes that there is no strong effect of the size of segmentation on the difference, but the ICV of reader 1 is higher on average than that of reader 2. These differences are small compared to the value of the ICV estimate, however.

References

- Sahni, Ramandeep and Jesse Weinberger (2007). "Management of intracerebral hemorrhage". In: *Vascular Health and Risk Management* 3.5, pp. 701–709. ISSN: 1176-6344.
- Chalela, Julio A, Chelsea S Kidwell, Lauren M Nentwich, Marie Luby, John A Butman, Andrew M Demchuk, Michael D Hill, Nicholas Patronas, Lawrence Latour, and Steven Warach (2007). "Magnetic resonance imaging and computed tomography in emergency assessment of patients with suspected acute stroke: a prospective comparison". In: *The Lancet* 369.9558, pp. 293–298.
- Schellingen, Peter D, Olav Jansen, Jochen B Fiebach, Werner Hacke, and Klaus Sartor (1999). "A standardized MRI stroke protocol comparison with CT in hyperacute intracerebral hemorrhage". In: *Stroke* 30.4, pp. 765–768.
- Wang, Yaping, Jingxin Nie, Pew-Thian Yap, Gang Li, Feng Shi, Xiujuan Geng, Lei Guo, Dinggang Shen, Alzheimer's Disease Neuroimaging Initiative, and others (2014). "Knowledge-guided robust mri brain extraction for diverse large-scale neuroimaging studies on humans and non-human primates". In: *PloS one* 9.1, e77810.
- Smith, Stephen M. (2002). "Fast robust automated brain extraction". en. In: *Human Brain Mapping* 17.3, pp. 143–155. ISSN: 1097-0193. DOI: [10.1002/hbm.10062](https://doi.org/10.1002/hbm.10062).
- Jenkinson, Mark, Christian F. Beckmann, Timothy E. J. Behrens, Mark W. Woolrich, and Stephen M. Smith (2012). "FSL". In: *NeuroImage* 62.2, pp. 782–790. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2011.09.015](https://doi.org/10.1016/j.neuroimage.2011.09.015).
- Solomon, Jeffrey, Vanessa Raymont, Allen Braun, John A. Butman, and Jordan Grafman (2007). "User-friendly software for the analysis of brain lesions (ABLE)". In: *Computer Methods and Programs in Biomedicine* 86.3, pp. 245–254. ISSN: 0169-2607. DOI: [10.1016/j.cmpb.2007.02.006](https://doi.org/10.1016/j.cmpb.2007.02.006).
- Rorden, Christopher, Leonardo Bonilha, Julius Fridriksson, Benjamin Bender, and Hans-Otto Karnath (2012). "Age-specific CT and MRI templates for

- “spatial normalization”. In: *NeuroImage* 61.4, pp. 957–965. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2012.03.020](https://doi.org/10.1016/j.neuroimage.2012.03.020).
- Mandell, Jason G, Jack W Langelaan, Andrew G Webb, and Steven J Schiff (2014a). “Volumetric brain analysis in neurosurgery: Part 1. Particle filter segmentation of brain and cerebrospinal fluid growth dynamics from MRI and CT images”. In: *Journal of Neurosurgery: Pediatrics*, pp. 1–12.
- Mandell, Jason G, Abhaya V Kulkarni, Benjamin C Warf, and Steven J Schiff (2014b). “Volumetric brain analysis in neurosurgery: Part 2. Brain and CSF volumes discriminate neurocognitive outcomes in hydrocephalus”. In: *Journal of Neurosurgery: Pediatrics*, pp. 1–8.
- Mandell, Jason G, Kenneth L Hill, Dan TD Nguyen, Kevin W Moser, Robert E Harbaugh, James McInerney, Brian Kaaya Nsubuga, John K Mugamba, Derek Johnson, Benjamin C Warf, et al. (2014c). “Volumetric brain analysis in neurosurgery: Part 3. Volumetric CT analysis as a predictor of seizure outcome following temporal lobectomy”. In: *Journal of Neurosurgery: Pediatrics*, pp. 1–11.
- Morgan, T., M. Zuccarello, R. Narayan, P. Keyl, K. Lane, and D. Hanley (2008). “Preliminary findings of the minimally-invasive surgery plus rtPA for intracerebral hemorrhage evacuation (MISTIE) clinical trial”. In: *Cerebral Hemorrhage*. Springer, pp. 147–151.
- Mould, W. Andrew, J. Ricardo Carhuapoma, John Muschelli, Karen Lane, Timothy C. Morgan, Nichol A. McBee, Amanda J. Bistran-Hall, Natalie L. Ullman, Paul Vespa, Neil A. Martin, Issam Awad, Mario Zuccarello, and Daniel F. Hanley (2013). “Minimally Invasive Surgery Plus Recombinant Tissue-type Plasminogen Activator for Intracerebral Hemorrhage Evacuation Decreases Perihematomal Edema”. en. In: *Stroke* 44.3, pp. 627–634. ISSN: 0039-2499, 1524-4628. DOI: [10.1161/STROKEAHA.111.000411](https://doi.org/10.1161/STROKEAHA.111.000411).
- Rorden, Chris and Matthew Brett (2000). “Stereotaxic Display of Brain Lesions”. en. In: *Behavioural Neurology* 12.4, pp. 191–200. ISSN: 0953-4180. DOI: [10.1155/2000/421719](https://doi.org/10.1155/2000/421719).
- Dice, Lee R. (1945). “Measures of the amount of ecologic association between species”. In: *Ecology* 26.3, pp. 297–302.
- Whitwell, Jennifer L, William R Crum, Hilary C Watt, and Nick C Fox (2001). “Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging”. In: *American Journal of Neuroradiology* 22.8, pp. 1483–1489.
- Searle, Shayle R. (2012). *Linear models*. John Wiley & Sons.

- Thomas, James D. and Robert A. Hultquist (1978). "Interval estimation for the unbalanced case of the one-way random effects model". In: *The Annals of Statistics*, pp. 582–587.
- Donner, Allan (1979). "The use of correlation and regression in the analysis of family resemblance". In: *American journal of epidemiology* 110.3, pp. 335–342.
- Lessells, C. M. and Peter T. Boag (1987). "Unrepeatable repeatabilities: a common mistake". In: *The Auk*, pp. 116–121.
- Wolak, Matthew E., Daphne J. Fairbairn, and Yale R. Paulsen (2012). "Guidelines for estimating repeatability". en. In: *Methods in Ecology and Evolution* 3.1, pp. 129–137. ISSN: 2041-210X. DOI: [10.1111/j.2041-210X.2011.00125.x](https://doi.org/10.1111/j.2041-210X.2011.00125.x).
- Cleveland, William S., Eric Grosse, and William M. Shyu (1992). "Local regression models". In: *Statistical models in S*, pp. 309–376.
- Gousias, Ioannis S, Daniel Rueckert, Rolf A Heckemann, Leigh E Dyet, James P Boardman, A David Edwards, and Alexander Hammers (2008). "Automatic segmentation of brain MRIs of 2-year-olds into 83 regions of interest". In: *Neuroimage* 40.2, pp. 672–684.
- Shi, Feng, Li Wang, Yakang Dai, John H Gilmore, Weili Lin, and Dinggang Shen (2012). "LABEL: pediatric brain extraction using learning-based meta-algorithm". In: *Neuroimage* 62.3, pp. 1975–1986.
- Volkau, Ihar, Fiftarina Puspitasari, and Wieslaw L Nowinski (2010). "Ventricle boundary in CT: partial volume effect and local thresholds". In: *Journal of Biomedical Imaging* 2010, p. 15.

Chapter 7

Discussion and Conclusion

7.1 Neuroimaging and R

Neuroimaging has become an increasingly popular data type for statisticians to analyze. Enhancing familiar tools to statisticians, such as R, with the ability to process and analyze neuroimaging data helps statisticians provide complete analyses of data, from raw data to the finished product. As I have shown in chapter 3, the packages I have created, largely **fslr**, aid in that process and provides the necessary functionality.

Though **fslr** is highly popular, encouraging others to port their software into R is crucial, as it is commonly overlooked by those in the neuroimaging community. Many in that community may not know the capabilities above modeling and visualization that R has, such as Shiny applications and reproducible reporting Xie, 2015b; Xie, 2015a; Xie, 2014. Moreover, those developing packages within R may require some standardization of packages, objects, and syntax, similar to what has been done with genomics in Bioconductor (Huber et al., 2015).

7.2 CT Analysis

Although neuroimaging analysis is dominated by MRI, an increasing number of articles are being published on CT analysis. As many tools for MRI have been developed, many analyses for CT may try to use these tools. Some may work as expected, but many likely need to be adapted to CT data. Also, these methods should be validated in a new cohort of patients. The Phase III of the MISTIE trial (MISTIE III) has enrolled over 250 patients, is of the same inclusion criteria as MISTIE II and is a great data set on which to validate.

7.2.1 ICH Segmentation

In chapter 4, we have shown that hemorrhage segmentation can be done accurately when compared to manual segmentation. We can assign the voxels that have hemorrhage to determine the location of the hemorrhage and estimate the volume. We have not considered, however, measures of uncertainty around these volume estimates. In the case of logistic regression, each voxel is a predicted probability, for which there are methods to determine a prediction interval around this estimate. We can use these at each voxel and get a voxel-wise prediction map of the probability of hemorrhage, but this does not take into account the fact that there are multiple voxels with a complex dependence structure. Incorporating these measures some uncertainty could then be aggregated to create a prediction interval for the volume of the hemorrhage. Also, although we have over 100 test scans, validating this in MISTIE III will be crucial as it will enroll 500 patients not in the data set used in this work.

The segmentation of ICH presented has only analyzed baseline scans. This was due to the fact that the patient scans may be very different after treatment assignment. Follow-up scans on these patients may have anomalies or artifacts not present in the baseline data, especially those who have undergone surgery. Thus, a validation of this method may need to be done on post-operative or longitudinal scans.

7.2.2 Predictive Regions of Stroke Severity

In chapter 5, we have shown an exploratory and descriptive analysis of hemorrhage location in the brain. We have quantitatively described where hemorrhages occur in the population and an automated way to describe hemorrhage engagement at a population or individual level using previously-made atlases. We have also shown how voxel-wise testing of hemorrhage location compared to NIHSS may yield areas that are predictive of stroke severity over and above standard covariates. Even more than in ICH segmentation, validation of these regions in MISTIE III will be crucial. Using the one-dimensional coverage measure, data in MISTIE III can potentially confirm that this measure does, in fact, explain some of the variability in NIHSS even after adjustment.

This analysis is proof of concept that more quantitative analyses can be done with CT data, specifically with regards to hemorrhagic stroke. There is a wealth of data in there and simple processes may yield large results in determining if a patient is likely to recover. Moreover, there are practical goals that may be easily implemented.

For example, using a rigid-body registration of a post-operative scan where

an intracranial catheter is placed into the hemorrhage to a pre-operative scan may allow researchers to view where the catheter was placed in view of the pre-operative hemorrhage. As dosing of the catheter or manual suctioning may remove some of the hemorrhage before the post-operative CT scan was done but after the operation, this gives a more accurate picture of catheter placement. As poor catheter placement may delay dosing or require replacement of the catheter, it is important to know if the removal of the hemorrhage occurred versus a bad placement.

Overall, CT image analysis has many interesting problems, both clinically and statistically. One large hurdle remains, however, compared to MRI scans: open data sources. Publicly-available MRI data are becoming increasingly popular whereas raw clinical CT scans are not as prevalent, if at all. Encouraging researchers to publish or release their de-identified CT data with meaningful clinical questions may allow the growth of the analysis of this imaging modality significantly.

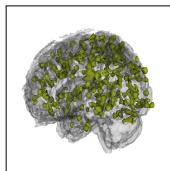
7.3 Conclusion

We have shown how tools for neuroimaging can be ported to R, which give them a larger audience to statisticians. We can use these tools to process CT images. These tools also can create predictors of intracerebral hemorrhage, which can accurately predict a hemorrhage mask from a CT scan. Using registration tools and a newly created CT template, we can use these images and hemorrhage masks to quantitatively describe where hemorrhages occur in a population and perform tests to determine which areas of the brain are

related to NIHSS if hemorrhage occurs there.

References

- Xie, Yihui (2015b). *knitr: A General-Purpose Package for Dynamic Report Generation in R*. URL: <http://yihui.name/knitr/>.
- Xie, Yihui (2015a). *Dynamic Documents with R and knitr*. 2nd. Boca Raton, Florida: Chapman and Hall/CRC. URL: <http://yihui.name/knitr/>.
- Xie, Yihui (2014). “knitr: A Comprehensive Tool for Reproducible Research in R”. In: *Implementing Reproducible Computational Research*. Ed. by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman and Hall/CRC. URL: <http://www.crcpress.com/product/isbn/9781466561595>.
- Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., Bravo, H. C., Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, K. D., Irizarry, R. A., Lawrence, M., Love, M. I., MacDonald, J., Obenchain, V., Ole’s, A. K., Pag’es, H., Reyes, A., Shannon, P., Smyth, G. K., Tenenbaum, D., Waldron, L., Morgan, and M. (2015). “Orchestrating high-throughput genomic analysis with Bioconductor”. In: *Nature Methods* 12.2, pp. 115–121. URL: <http://www.nature.com/nmeth/journal/v12/n2/full/nmeth.3252.html>.



John Muschelli

2120 Moyer St,
Baltimore MD 21231
610-291-7685
✉ muschellij2@gmail.com
<http://biostat.jhsph.edu/~jmuschel/>
Blog: A HopStat and Jump Away
Twitter: @StrictlyStat
GitHub: muschellij2
Skype: muschellij2

Research Interests

Neuroimaging, image segmentation, stroke, dynamic reports, statistical computing, machine learning.

Education

- 2012–Present **PhD Candidate, Biostatistics,**
Johns Hopkins Bloomberg School of Public Health, Baltimore, MD.
Expected graduation: May 2016
Areas of Study: Stroke CT image segmentation
Population-level stroke characterization
Gadolinium-enhancing brain lesion segmentation on MRI
Advisor: Professor Ciprian Crainiceanu
- 2008–2010 **Masters of Science, Biostatistics,**
Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, GPA: 3.80.
Area of Study: fMRI brain image data analysis
Thesis Topic: *An Iterative Approach to Hemodynamic Response Function Temporal Derivatives in Statistical Parametric Mapping for Functional Neuroimaging*
Advisor: Professor Brian Caffo
- 2004–2008 **Bachelors of Science, Biomathematics and Neuroscience,**
The University of Scranton, Scranton, PA, GPA: 3.87 (Summa Cum Laude).
Advisors: Professor Jakub Jasinski, Professor J. Timothy Cannon

Relevant Experience

- 2012–Present **Trainee, T32AG021334: Epidemiology and Biostatistics of Aging Training Grant,**
Mentors: Dr. Michelle Carlson, Dr. Ravi Varadhan.
- 2009–Present **Research Associate, Johns Hopkins Biostatistics Consulting Center, Baltimore, MD.**
Collaborated on statistical projects with senior consultants.
Weekly consulting for student research projects.
Report writing and analyzing data using statistical software: R, Stata.
- 2009–2014 **Data Analyst / Data Manager, Brain Injury Outcomes Division, Baltimore, MD.**
Increased turnaround time on data safety report (from weeks to hours) by using knitr, LaTeX, and dynamic documents.
Created a standardized database and processing pipeline for CT images.
Analyzed phase II and III trials for treatment of intracerebral hemorrhage
Data management and consultation of electronic case report form (eCRF) creation.

- 2010–2012 **Data Analyst**, *Laboratory for Neurocognitive and Imaging Research at Kennedy Krieger Institute*, Baltimore, MD.
Reduced manual steps in complex imaging study analysis using automation from programming.
Analysis of functional MRI (fMRI) imaging studies using Statistical Parametric Mapping.
Programming consultant: Matlab & R.
- 2008 **Intern**, *Analysis & Inference*, Swarthmore, PA.
Cooperated on statistical projects and conferred with clients about possible analysis options.
Report writing of analyses, data cleaning.
- 2007 **Research Intern**, *Dupont Stine-Haskell Laboratory*, Wilmington, DE.
Developed lab skills and techniques: cell culturing, making and sterilizing broth media, optical density readings, inoculations, quality control, cell counts, screening for fungicidal properties of compounds.

Teaching Experience

- 2015 **Instructor**, *ENAR, A Tutorial for Multisequence Clinical Structural Brain MRI*.
Co-developed and instructed a tutorial for 35 statisticians. Created 75% of all code and slides for presentation and presented for half of the 3-hour session.
- 2015 **Instructor**, *Coursera, Neurohacking with R*.
Co-developed a MOOC (massive open online course) for Coursera on neuroimage processing and statistical analysis completely within R. Developed 50% of code and slides for presentation and recorded lectures delivering slides.
- 2014–present **Co-Instructor**, *JHSPH, Introduction to R for Public Health Researchers*.
Co-developed a one-week, 8-hour-a-day course in the Winter and Summer Institutes at Johns Hopkins with Dr. Andrew Jaffe. Developed 25% of code and slides for presentation and recorded lectures delivering slides.

Teaching Assistant

All teaching assistantships were in the Department of Biostatistics at the Johns Hopkins Bloomberg of Public Health.

- 2015-2016 **Advanced Data Science I-II (PH.140.711-712)**, 1st-2nd term.
Instructors: Jeff Leek, PhD and Elizabeth Colantuoni, PhD
- 2014-2015 **Statistical Methods in Public Health IV (PH.140.624)**, 4th term.
Instructors: James Tonascia, Ph.D and Mark Van Natta, MHS
- 2014-2015 **Special topics: Statistical Consulting**, 1st-3rd term.
Instructors: Carol Thompson, MS and Elizabeth Colantuoni, PhD
- 2013-2014 **Methods in Biostatistics I-II (PH.140.651-652)**, 1st-2nd term.
Instructor: Ciprian Crainiceanu, PhD
- 2012-2013 **Methods in Biostatistics III-IV (PH.140.653-654)**, 3rd-4th term.
Instructor: Hongkai Ji, PhD
- 2012-2013 **Methods in Biostatistics I-II (PH.140.651-652)**, 1st-2nd term.
Instructor: Thomas Louis, PhD
- 2010-2011 **Statistical Methods in Public Health IV (PH.140.624)**, 4th term.
Instructors: James Tonascia, PhD and Mark Van Natta, MHS
- 2010-2011 **Statistical Methods in Public Health I (PH.140.621)**, 1st-3rd term.
Instructors: Marie Diener West, PhD and Karen Bandeen Roche, PhD

2009-2010 **Methods in Biostatistics I-II (PH.140.651-652)**, 1st-2nd term.
Instructor: Brian Caffo, PhD

Working Groups

- 2014-Present **Statistical and Applied Mathematical Sciences Institute (SAMSI) working group on Clinical Brain Imaging.**
- 2012-Present **Epidemiology and Biostatistics of Aging (EBA) Training Program Meeting**, Johns Hopkins University, Center on Aging and Health.
- 2014-Present **Penn Statistical Imaging and Visualization Endeavor (PennSIVE) Working Group**, University of Pennsylvania, Department of Biostatistics and Epidemiology.
- 2009-Present **Statistical Methods and Applications for Research in Technology (SMART) Working Group**, Johns Hopkins University, Department of Biostatistics.

Peer-Reviewed Publications

- 2015 **Muschelli, J.**, Ullman, N. L., Mould, W. A., Vespa, P., Hanley, D. F., Crainiceanu, C. M. “Validated automatic brain extraction of head CT images”. *NeuroImage* 114, pp. 379–385.
- Muschelli, J.**, Sweeney, E., Lindquist, M., Crainiceanu, C. “fslr: connecting the FSL software with R”. *R Journal* 7.1, pp. 163–175.
- Muschelli, J.**, Ullman, N. L., Sweeney, E. M., Eloyan, A., Martin, N., Vespa, P., Hanley, D. F., Crainiceanu, C. M. “Quantitative intracerebral hemorrhage localization”. *Stroke*, In Press.
- Choe, A. S., Jones, C. K., Joel, S. E., **Muschelli, J.**, Belegu, V., Caffo, B. S., Lindquist, M. A., Zijl, P. C., Pekar, J. J. “Reproducibility and temporal structure in weekly resting-state fMRI over a period of 3.5 years”. *PLOS ONE*, In Press.
- Sweeney, E. M., Shinohara, R. T., **Muschelli, J.**, Dewey, B. E., Reich, D. S., Crainiceanu, C. M., Schindler, M. K., Eloyan, A. “Relating multi-sequence longitudinal intensity profiles and clinical covariates in incident multiple sclerosis lesions”. *NeuroImage: Clinical*, In Press.
- Webb, A. J., Ullman, N. L., Morgan, T. C., **Muschelli, J.**, Kornbluth, J., Awad, I. A., Mayo, S., Rosenblum, M., Ziai, W., Aldrich, Zuccarello, F. M., John, S., Harnof, S., Lopez, G., Broaddus, W. C., Wijman, C., Vespa, P., Bullock, R., Haines, S. J., Cruz-Flores, S., Tuhrim, S., Hill, M. D., Narayan, R., Hanley, D. F. “Accuracy of the ABC/2 score for intracerebral hemorrhage systematic review and analysis of MISTIE, CLEAR-IVH, and CLEAR III”. *Stroke* 46.9, pp. 2470–2476.
- 2014 **Muschelli, J.**, Nebel, M. B., Caffo, B. S., Barber, A. D., Pekar, J. J., Mostofsky, S. H. “Reduction of motion-related artifacts in resting state fMRI using aCompCor”. *NeuroImage* 96, pp. 22–35.
- Muschelli, J.**, Sweeney, E., Crainiceanu, C. “brainR: interactive 3 and 4D images of high resolution neuroimage data”. *R Journal* 6.1, pp. 41–48.
- Muschelli, J.**, Betz, J., Varadhan, R. “Binomial regression in R”. *Handbook of Statistics: Computational Statistics with R* 32, pp. 257–309.

- Eloyan, A., Li, S., **Muschelli, J.**, Pekar, J. J., Mostofsky, S. H., Caffo, B. S. “Analytic programming with fMRI data: a quick-start guide for statisticians using R”. *PLOS ONE* 9.2, e89470.
- Nebel, M. B., Joel, S. E., **Muschelli, J.**, Barber, A. D., Caffo, B. S., Pekar, J. J., Mostofsky, S. H. “Disruption of functional organization within the primary motor cortex in children with autism”. *Human Brain Mapping* 35.2, pp. 567–580.
- 2013 Mould, W. A., Carhuapoma, J. R., **Muschelli, J.**, Lane, K., Morgan, T. C., McBee, N. A., Bistran-Hall, A. J., Ullman, N. L., Vespa, P., Martin, N. A., Awad, I., Zuccarello, M., Hanley, D. F. “Minimally invasive surgery plus recombinant tissue-type plasminogen activator for intracerebral hemorrhage evacuation decreases perihematomal edema”. *Stroke* 44.3, pp. 627–634.
- Mould, W., Carhuapoma, J., **Muschelli, J.**, Lane, K., Morgan, T., McBee, N., Bistran-Hall, A., Ullman, N., Vespa, P., Martin, N., Awad, I., Zuccarello, M., Hanley, D. F. “MISTIE investigators: minimally invasive surgery plus recombinant tissue-type plasminogen activator for intracerebral hemorrhage evacuation decreases perihematomal edema”. *Stroke* 44.3, pp. 627–634.
- 2012 Bundy, D. G., **Muschelli, J.**, Clemens, G. D., Strouse, J. J., Thompson, R. E., Casella, J. F., Miller, M. R. “Ambulatory care connections of medicaid-insured children with sickle cell disease”. *Pediatric Blood & Cancer* 59.5, pp. 888–894.
- Eloyan, A., **Muschelli, J.**, Nebel, M. B., Liu, H., Han, F., Zhao, T., Barber, A. D., Joel, S., Pekar, J. J., Mostofsky, S. H., others, “Automated diagnoses of attention deficit hyperactive disorder using magnetic resonance imaging”. *Frontiers in Systems Neuroscience* 6, p. 6.
- Hinson, H. E., Melnychuk, E., **Muschelli, J.**, Hanley, D. F., Awad, I. A., Ziai, W. C. “Drainage efficiency with dual versus single catheters in severe intraventricular hemorrhage”. *Neurocritical Care* 16.3, pp. 399–405.
- Jaffe, J., Melnychuk, E., **Muschelli, J.**, Ziai, W., Morgan, T., Hanley, D. F., Awad, I. A. “Ventricular catheter location and the clearance of intraventricular hemorrhage”. *Neurosurgery* 70.5, pp. 1258–1264.
- Webb, A. J., Ullman, N. L., Mann, S., **Muschelli, J.**, Awad, I. A., Hanley, D. F. “Resolution of intraventricular hemorrhage varies by ventricular region and dose of intraventricular thrombolytic the clot lysis: evaluating accelerated resolution of IVH (CLEAR IVH) program”. *Stroke* 43.6, pp. 1666–1668.
- Ziai, W. C., **Muschelli, J.**, Thompson, C. B., Keyl, P. M., Lane, K., Shao, S., Hanley, D. F. “Factors affecting clot lysis rates in patients with spontaneous intraventricular hemorrhage”. *Stroke* 43.5, pp. 1234–1239.
- 2011 Newell, D. W., Shah, M. M., Wilcox, R., Hansmann, D. R., Melnychuk, E., **Muschelli, J.**, Hanley, D. F. “Minimally invasive evacuation of spontaneous intracerebral hemorrhage using sonothrombolysis”. *Journal of Neurosurgery* 115.3, pp. 592–601.
- Niedner, M. F., Huskins, W. C., Colantuoni, E., **Muschelli, J.**, Harris, J. M., Rice, T. B., Brilli, R. J., Miller, M. R. “Epidemiology of central line-associated bloodstream infections in the pediatric intensive care unit”. *Infection Control* 32.12, pp. 1200–1208.

Submitted

- 2015 Fortin, J.-P., Sweeney, E. M., **Muschelli, J.**, Crainiceanu, C. M., Shinohara, R. T. “Removing inter-subject technical variability in magnetic resonance imaging studies”. *NeuroImage*, Submitted.
- Kickingereder, P., Götz, M., **Muschelli, J.**, Wick, A., Neuberger, U., Shinohara, R. T., Radbruch, A., Schlemmer, H.-P., Wick, W., Bendszus, M., Maier-Hein, K., Bonekamp, D. “Large-scale radiomic profiling of glioblastoma identifies an imaging signature for predicting and stratifying antiangiogenic treatment response”. *Journal of Clinical Oncology*, Submitted.
- Bundy, D. G., **Muschelli, J.**, Clemens, G., Strouse, J., Thompson, R., Casella, J., Miller, M. “Preventive care delivery to young children with sickle cell disease”. *Journal of Pediatric Hematology and Oncology*, In Press.

Talks and Presentations

- 2015 **SuBGELS: Subtraction-Based Gadolinium-Enhancing Lesion Segmentation**,
Hopkins Imaging Conference, Baltimore, MD, Poster.
- Automated Intracerebral Hemorrhage Segmentation of CT Scans**,
Joint Statistical Meeting (JSM), Seattle, WA, SPEED Talk and Poster.
- PItcHPERFECT: Primary Intracerebral Hemorrhage Prediction Employing Regression and Features Extracted from CT**,
Eastern North American Region (ENAR), Miami, FL, Poster.
- Quantitative Localization and Predictive Performance of Intracranial Hemorrhage**,
International Stroke Conference (ISC), Nashville, TN, Poster.
- Validated Automatic Brain Extraction of Head CT Images**,
Organization for Human Brain Mapping (OHBM), Honolulu, HI, Poster.
- 2014 **Validated Automatic Brain Extraction of Head CT Images**,
Hopkins Imaging Conference, Baltimore, MD, Talk and Poster.
- Reduction of motion-related artifacts in resting state fMRI using aCompCor**,
Hopkins Imaging Conference, Baltimore, MD, Poster.
Award: Top Poster
- 2013 **Visualizing Brain Imaging in Interactive 3D**,
ENAR, Orlando, FL, Talk.
- 2012 **Resting State Preprocessing and Motion Artifacts**,
Second Biennial Conference on Resting State, Magdeburg, Germany, Poster.
- Effects of preprocessing on motion-induced artifacts in resting state fMRI**,
Society for Neuroscience (SfN), New Orleans, LA, Poster.

Software

R Packages

All download counts are from RStudio CRAN logs and are accurate as of January 26, 2016.

CRAN	fslr: Wrapper functions for FSL (FMRIB Software Library) from Functional MRI of the Brain (FMRIB), Downloads: 6164.
	WhiteStripe: Whitestripe White Matter Normalization for Magnetic Resonance Images, Downloads: 3560.
	brainR: Helper functions to misc3d and rgl packages for brain imaging, Downloads: 6976.
	matlabr: R interface with calling MATLAB code without a server, Downloads: 903.
	spm12r: Wrapper Functions for SPM (Statistical Parametric Mapping) Version 12 from the Wellcome Trust Centre for Neuroimaging, Downloads: 586.
	rscopus: Scopus Database API Interface, Downloads: 571.
	diffR: Display Differences Between Two Files using the Codediff JavaScript Library, Downloads: 533.
GitHub	drammsr: Port of Deformable Registration via Attribute Matching and Mutual-Saliency Weighting (DRAMMS) Registration to R. extrantsr: Additional functionality and extensions to the ANTsR R package. dcm2niir: R wrapper for dcm2nii DICOM converter. googleCite: Scraper for Google Citations. processVISION: Scripts for Parsing XML from VISION database.

Shiny Web Applications

2015	Abandoned Cars in Baltimore Finder, https://jmuschelli.shinyapps.io/Abandoned_Baltimore_Car .
	Unofficial ENAR 2015 Itinerary Maker, https://muscchellij2.shinyapps.io/ENAR_2015 .
2014	Online DICOM TO NIfTI Converter, https://muscchellij2.shinyapps.io/dcm2nii . Cost of most common medical procedures at United States hospitals based on Centers for Medicare and Medicaid Services data, https://jmuschelli.shinyapps.io/Shiny_Health_Data .

Skills

Languages	Proficient: R, bash, Stata, MATLAB. Beginner: SAS, Python, C++, Visual Basic, JavaScript
Markup	TeX, L ^A T _E X, BIBTeX, TeXShop, WinEdt, knitr, HTML, CSS

Honors and Awards

- 2014 **SOURCE (Student Outreach Resource Center) Community Service Award.**
- 2011 **Member of the winning team of the ADHD 200 Competition:**
a competition to develop diagnostic classification tools for ADHD diagnosis based on imaging of the brain.
- 2004–2008 **Presidential Scholar (Full Tuition Scholarship).**
- 2004–2008 **Dean's List.**
- 2004 **Alpha Lambda Delta.**
- 2008 **Alpha Sigma Nu.**

Additional Experience

- 2015 **IdEar Team Member, Hackathon: MedHacks 1.0,** Showed as a proof of concept that ears could be used as biometric markers in a global health framework. The target was areas with poor to no registries of people in which pictures of ears could be used to distinguish community members when other demographic information was not unique. Implemented a MATLAB implementation of an SVM to classify people based on image of ear and scale invariant feature transform (SIFT) features. In top 10 of 30 teams.
- 2015 **Safer Baltimore Biking Team, Hackathon: JHU Data Science Hackathon,** Full description and product: <http://kbroman.org/jhudashbike/>. Team used open data from Baltimore City to determine road safety as measured by accidents, hazards (potholes), and accidents. Geocoded all hazards and helped develop leaflet final product (map).

Academic Service

- 2013–Present **Middle Manager,**
Thread/Incentive Mentoring Program.
Interfaced between executive-level staff and lower-level management. Organized monthly meetings, weekly progress updates and e-mails, and provided broad-scale mentorship for high-school students in the Baltimore City school district. Program title was "Grandparent", as it is a family-based, positive change model.
- 2014 **Organizer, Journal Club,**
JHSPH Department of Biostatistics.
Scheduled and organized a club for reading and discussing statistical papers.
- 2013-2015 **Founder/Organizer, Writing Accountability Group,**
JHSPH Department of Biostatistics.
Founded and organized a small group (6-8) of students, where the aim is to develop weekly goals for writing and publication.
- 2013-2014 **Founder/Organizer, Blogging Club,**
JHSPH Department of Biostatistics.
Founded and organized a club for student blogging.
- 2010–2013 **Manager,**
Thread/Incentive Mentoring Program.
Title was a "Head of Household"; mentored and tutored a student from Dunbar High School, teaching coursework, life skills, support as needed.