# PItcHPERFECT: Primary Intracranial Hemorrhage Probability Estimation using Regression and Features Extracted from CT

John Muschelli[a,*], Elizabeth M. Sweeney[a], Paul Vespa[c], Daniel F. Hanley[b], Ciprian M. Crainiceanu[a]

[a]*Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA*
[b]*Department of Neurology, Division of Brain Injury Outcomes, Johns Hopkins Medical Institutions, Baltimore, MD, USA*
[c]*Department of Neurosurgery, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA*

**Abstract**

*Keywords:* CT, ICH Segmentation

## 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 [38]. The volume of ICH has been consistently demonstrated to be an important diagnostic predictor of stroke severity, long-term functional outcome, and mortality [8, 17, 42]. ICH volume change is also common primary outcome [1, 2, 24, 33] and secondary outcome [1, 25, 26] in clinical trials. Moreover, the location of the ICH has been shown to affect functional outcome in patients with stroke [9, 37].

ICH volume can be rapidly measured using techniques such as the ABC/2 method [8]. 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 [20]. As this method only requires 3 measurements, this method can be done rapidly.

Although ABC/2 is is widely used, Divani et al. [11] 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. [43] found that ABC/2 measurements at a clinical site, 81% of the 4,369 scans were within 5 milliters (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 [17, 42]. ABC/2 also been shown to consistently over-estimate infarct volume [30], and can have significant inter-rater variability [18]. 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) Other methods have been presented for automated methods for estimating ICH from CT scans [14, 22, 23, 31, 32]. These methods include fuzzy clustering [22, 31], simulated annealing [23], 3-dimensional (3D) mathematical morphology operations [32], and template-based comparisons [14]. We wish to create an algorithm that can estimate the probability of ICH at a voxel-level, the volume of ICH, and the level of uncertainty in these

*Principal Corresponding Author
*Email addresses:* jmusche1@jhu.edu (John Muschelli), emsweene1@jhu.edu (Elizabeth M. Sweeney), PVespa@mednet.ucla.edu (Paul Vespa), dhanley@jhmi.edu (Daniel F. Hanley), ccrainic@jhsph.edu (Ciprian M. Crainiceanu)

estimates. We will compare our predicted ICH maps to the gold standard – manual segmentation. Moreover, we wish to provide a complete pipeline of analysis from raw images to binary hemorrhage masks and volume estimates.

## 2. Methods

### *2.1. Data*

### *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) and ICES (Intraoperative CT-Guided Endoscopic Surgery) stroke trials [25]. 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. [27]). 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.

### *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 [6, 39]; 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.

### *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 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 [35]). 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 [34], using the `fslr` [28] package to call functions from the FSL [19] neuroimaging software (version 5.0.4), and the `ANTsR` package to call functions from the ANTs (Advanced Normalization Tools) neuroimaging software [4].

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) [40], a function of FSL, using a previously published validated CT-specific brain extraction protocol [29].

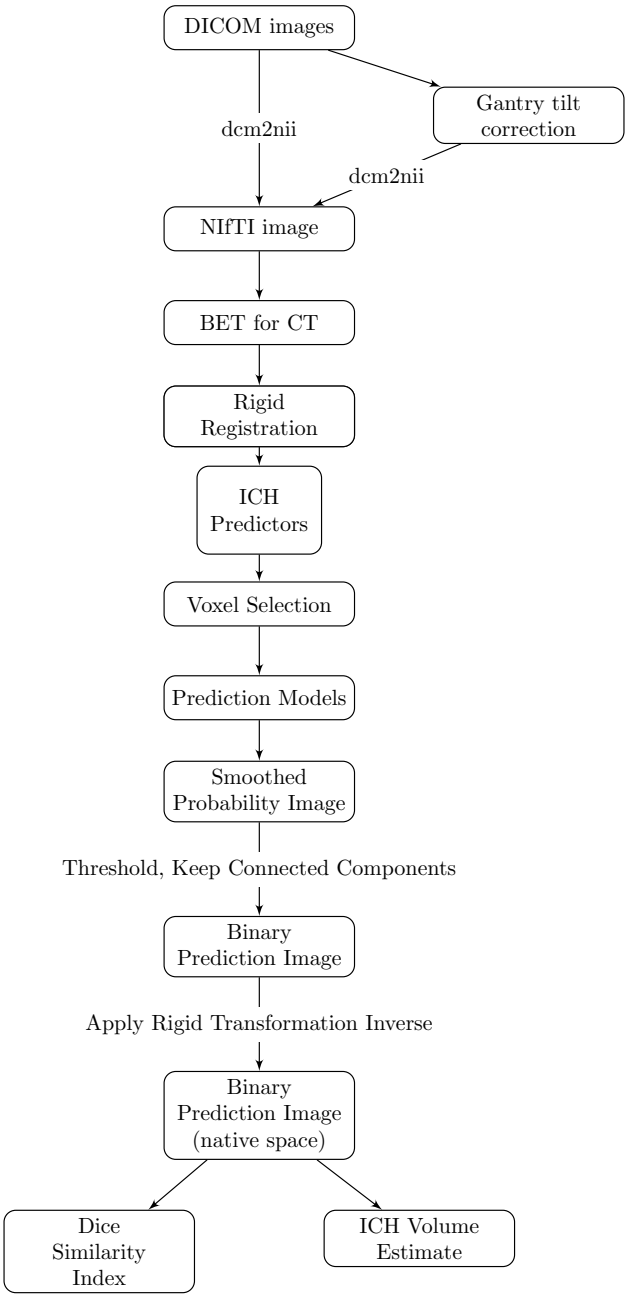


Figure 1: **Processing Pipeline**. Images in DICOM (Digital Imaging and Communications in Medicine) format were gantry tilt corrected if necessary and converted to NIfTI (Neuroimaging Informatics Technology Initiative) format using `dcm2nii`. After NIfTI conversion, the brain extraction tool (BET) was 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 native space. The ICH volume and the Dice Similarity Index, an overlap measure, were calculated compared to the true estimate from the manual segmentation.

*2.5. Image Registration*

Rorden et al. [36] 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. 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 operations are done in MNI space, unless otherwise specified.

*2.6. Voxel Selection*

Each brain mask was eroded by a box kernel ($3mm \times 3mm \times 1$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 ICH 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.

*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.

*2.7.1. CT voxel intensity information*

The raw voxel intensity value in HU was included, 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.

*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 $k = 1, \ldots, 27$. We created the voxel neighborhood mean intensity ($\bar{x}(v)$):

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

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 skewness, 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 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:

$$p_{\text{thresh}}(v) = \frac{1}{N(v)} \sum_{k \in N(v)} I\{40 \leq x_k(v) \leq 80\} \tag{2}$$

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 as these are less likely to be ICH. Voxels not within the 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:

$$p_0(v) = \frac{1}{N(v)} \sum_{k \in N(v)} I\{x_k(v) = 0\} \tag{3}$$

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

$$\bar{I}_0(v) = I\{p_0(v) > 0\} \tag{4}$$

*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:

$$z_o(v) = \frac{x(v) - \bar{x}(v, o)}{\sigma(v, o)} \tag{5}$$

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).

*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 [5]. We used a 4-tissue class segmentation, and combined the top 2 probability values into one class. We used this probability image as a predictor. 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.

*2.7.5. Contralateral Difference Images*

As most hemorrhages are constrained to one side of the brain, the contralateral side should have 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:

$$f(v) = x(v) - x(v^*) \tag{6}$$

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

*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 which 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.

*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) [3]. 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 ($z_{\text{template}}$) 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 [14].

*2.8. Model Creation*

We chose 10 scans from 10 patients to perform exploratory data analysis, model fitting, and estimation of model cutoffs, denoted as training data. Of the 102 remaining scans, we split the data into 51 validation scans and 51 test scans.

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 $z_{\text{axial}}$, $z_{\text{coronal}}$, $p_{\text{thresh}}$, 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. All models were fit with all predictors.

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.

*2.9. Models*

Using the sampled training voxels, we created predictions for the probability of ICH using a 1) logistic regression model, 2) logistic regression model with the LASSO 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 [41] using the `glmnet` package [13]. The tuning parameter ($\lambda$) for the penalization was chosen using 10-fold cross-validation (of only the training voxels), with the cost function of misclassification rate. The parameter was chosen using the largest value such that the error is within 1 standard error of the minimum, chosen for out-of-sample performance stability.

The generalized additive model (GAM) [15, 16] 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 [44, 45].

We fit a random forest [7] using the `randomForest` package in R [21], using the default pruning parameters and number of trees (`ntree`=500, `mtry`=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 would like a binary prediction 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) [10] between the prediction and true value. The DSI is measure of overlap insensitive to values where neither the true segmentation or predicted segmentation were considered ICH, and will be used as a performance measure when comparing models.

After thresholding the smoothed image using this probability cutoff, we discarded regions with fewer than 100 (0.1 milliliters) connected voxels. We then transformed this binary mask back to 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 [12].

For each scan in the validation data, this prediction process was performed and each each scan has a corresponding binary prediction image.

### 2.11. Measuring and Testing ICH Prediction Performance

For all measures, we used the binary prediction masks from the validation scans. We measured performance for each model using the DSI and also estimated the ICH volume of the prediction. Specificity and overall accuracy are inflated by the fact that most of the voxels within the brain are non-ICH. 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. For DSI and correlation, higher values indicate better agreement with the manual segmentation. For RMSE, lower indicates better agreement.

## 3. Results

### 3.1. Dice Similarity Index

In Figure 2, we show the DSI distributions from the validation 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. Overall, the results indicate good overlap between the predicted ICH segmentation and manual ICH segmentation.

### 3.2. ICH Volume Estimation

In Figure 3, 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.

### 3.3. Model Choice

Overall, all models perform adequately for ICH segmentation. Some failures exist, but the algorithm using the

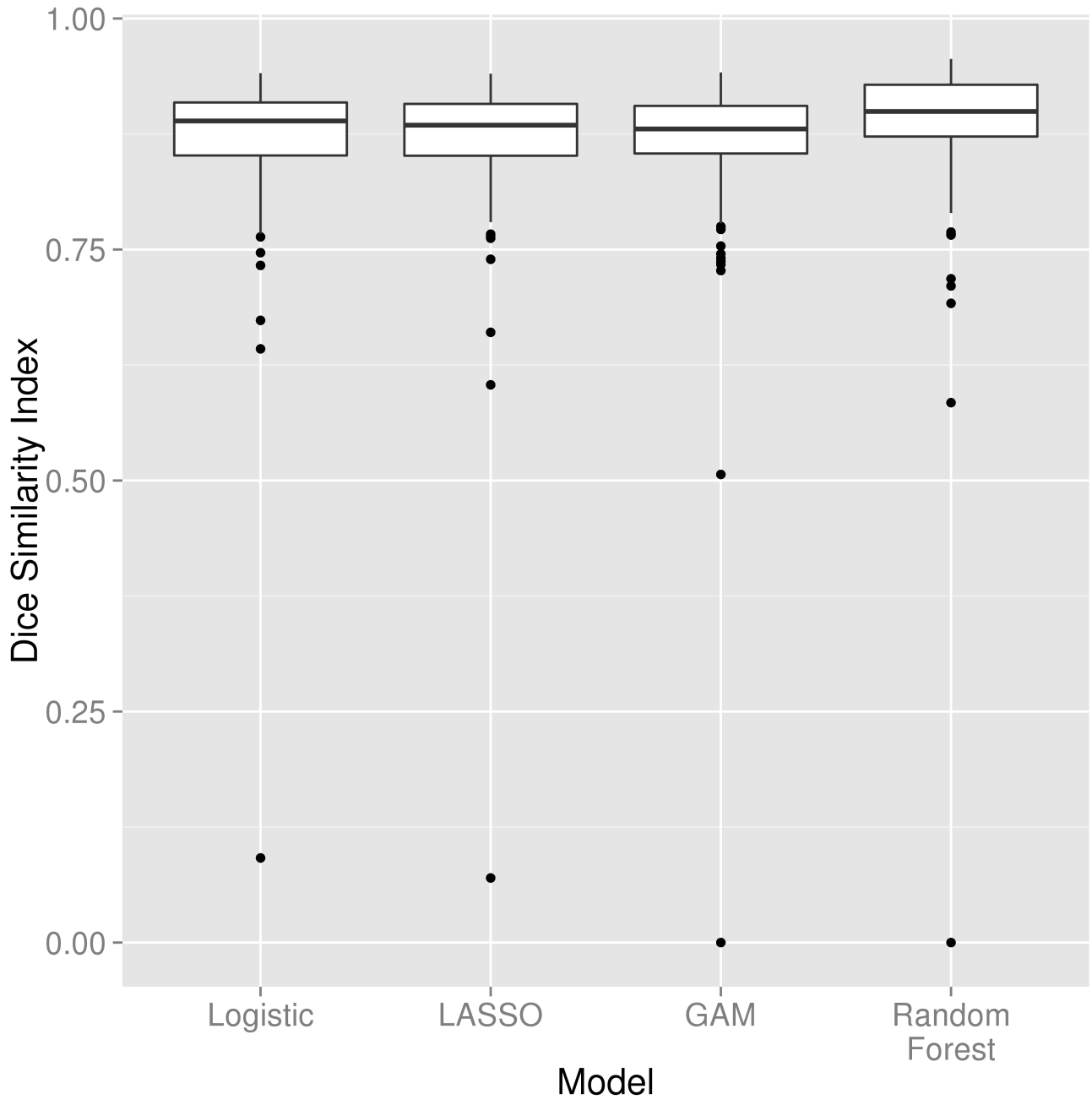


Figure 2: **Distribution of Dice Similarity I**

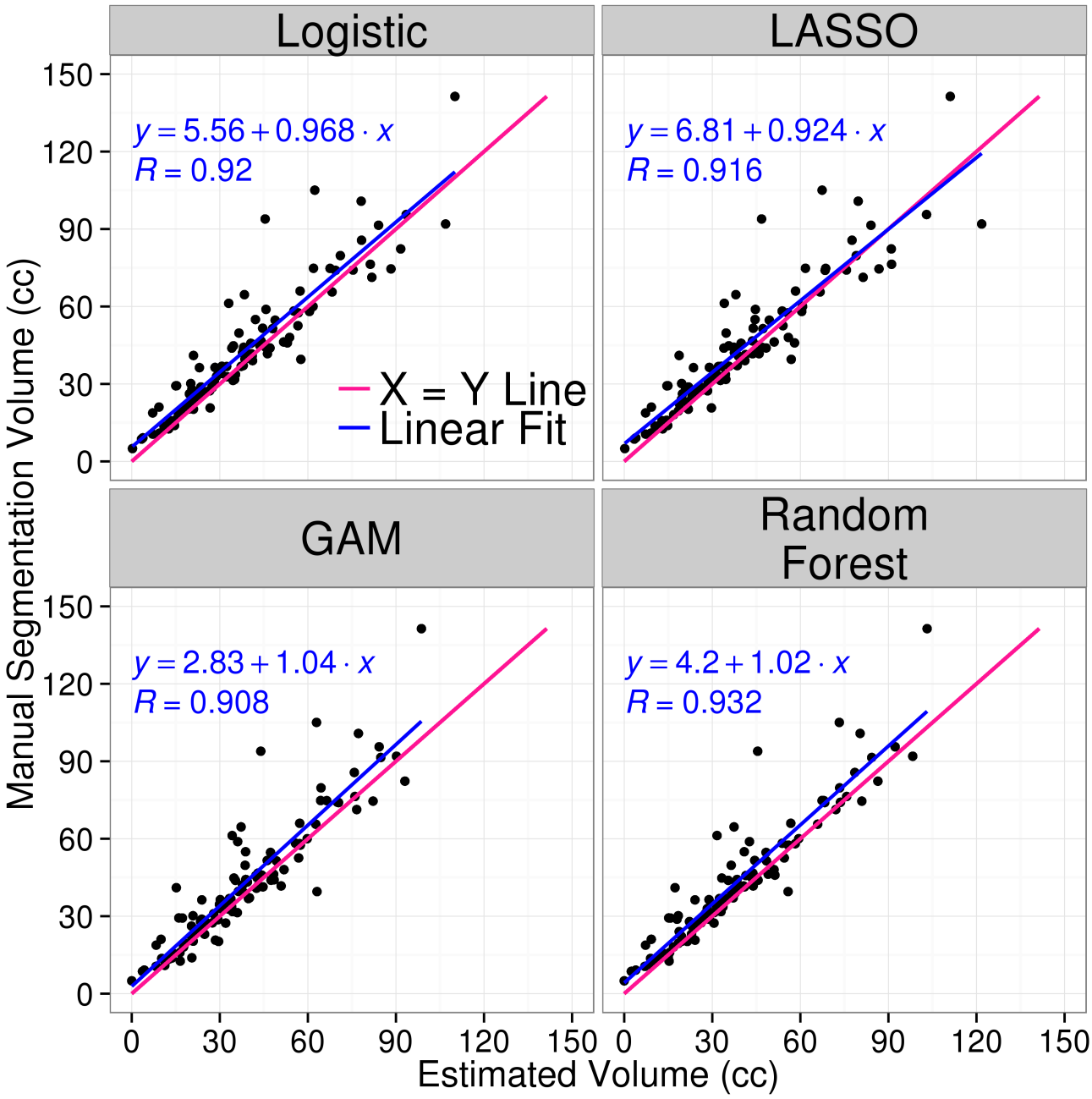
Logistic
LASSO
GAM
Random Forest
y = 5.56 + 0.968 · x
R = 0.92
y = 6.81 + 0.924 · x
R = 0.916
y = 2.83 + 1.04 · x
R = 0.908
y = 4.2 + 1.02 · x
R = 0.932
X = Y Line
Linear Fit
Manual Segmentation Volume (cc)
Estimated Volume (cc)
0
30
60
90
120
150

Figure 3

## 4. Discussion

We have presented a fully automated of hemorrhage segmentation for intracranial hemorrhage. We create a set of predictors that attempt to capture the features relevant to distinguishing ICH versus non-ICH areas. The models use allow for a relatively fast and intuitive estimation of the probability of ICH. We have incorporated different aspects of a CT scan, each representing intuitive information. Different pipelines were used to determine which processing methods could potentially improve ICH prediction.


## Acknowledgments

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).

## References


[1] Craig S. Anderson et al. "Intensive blood pressure reduction in acute cerebral haemorrhage trial (INTERACT): a randomised pilot trial". In: *The Lancet Neurology* 7.5 (2008), pp. 391–399.

[2] Craig S. Anderson et al. "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 (Feb. 1, 2010), pp. 307–312.

[3] B. B. Avants et al. "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 (Feb. 2008), pp. 26–41.

[4] Brian B. Avants et al. "A reproducible evaluation of ANTs similarity metric performance in brain image registration". In: *NeuroImage* 54.3 (Feb. 1, 2011), pp. 2033–2044.

[5] Brian B Avants et al. "An open source multivariate framework for n-tissue segmentation with evaluation on public data". In: *Neuroinformatics* 9.4 (2011), pp. 381–400.

[6] M Bergstrm et al. "Variation with time of the attenuation values of intracranial hematomas". In: *Journal of computer assisted tomography* 1.1 (Jan. 1977), pp. 57–63.

[7] Leo Breiman. "Random forests". In: *Machine learning* 45.1 (2001), pp. 5–32.

[8] J. P. Broderick et al. "Volume of intracerebral hemorrhage. A powerful and easy-to-use predictor of 30-day mortality." In: *Stroke* 24.7 (July 1, 1993), pp. 987–993.

[9] M. Castellanos et al. "Predictors of good outcome in medium to large spontaneous supratentorial intracerebral haemorrhages". In: *Journal of Neurology, Neurosurgery & Psychiatry* 76.5 (May 1, 2005), pp. 691–695.

[10] Lee R. Dice. "Measures of the amount of ecologic association between species". In: *Ecology* 26.3 (1945), pp. 297–302.

[11] Afshin A. Divani et al. "The ABCs of accurate volumetric measurement of cerebral hematoma". In: *Stroke* 42.6 (2011), pp. 1569–1574.

[12] *Frequently Asked Questions for FLIRT*. `http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT/FAQ`.

[13] Jerome Friedman, Trevor Hastie, and Rob Tibshirani. "Regularization Paths for Generalized Linear Models via Coordinate Descent". In: *Journal of statistical software* 33.1 (2010), pp. 1–22.

[14] Cline R. Gillebert, Glyn W. Humphreys, and Dante Mantini. "Automated delineation of stroke lesions using brain CT images". In: *NeuroImage: Clinical* 4 (2014), pp. 540–548.

[15] Trevor Hastie and Robert Tibshirani. "Generalized additive models". In: *Statistical science* (1986), pp. 297–310.

[16] Trevor J. Hastie and Robert J. Tibshirani. *Generalized additive models*. Vol. 43. CRC Press, 1990.

[17] J. Claude Hemphill et al. "The ICH Score A Simple, Reliable Grading Scale for Intracerebral Hemorrhage". In: *Stroke* 32.4 (Apr. 1, 2001), pp. 891–897.

[18] Haitham M. Hussein et al. "Reliability of Hematoma Volume Measurement at Local Sites in a Multicenter Acute Intracerebral Hemorrhage Clinical Trial". In: *Stroke* 44.1 (Jan. 1, 2013), pp. 237–239.

[19] Mark Jenkinson et al. "FSL". In: *NeuroImage* 62.2 (Aug. 15, 2012), pp. 782–790.

[20] Rashmi U. Kothari et al. "The ABCs of Measuring Intracerebral Hemorrhage Volumes". In: *Stroke* 27.8 (Aug. 1, 1996), pp. 1304–1305.

[21] Andy Liaw and Matthew Wiener. "Classification and Regression by randomForest". In: *R News* 2.3 (2002), pp. 18–22.

[22] Sven Loncaric, Dubravko Cosic, and Atam P. Dhawan. "Hierarchical segmentation of CT head images". In: *Proc IEEE EMBS. doi* 10 (1996), p. 1109.

[23] Sven Loncaric et al. "Quantitative intracerebral brain hemorrhage analysis". In: *Medical Imaging'99*. International Society for Optics and Photonics, 1999, pp. 886–894.

[24] Stephan A. Mayer et al. "Recombinant Activated Factor VII for Acute Intracerebral Hemorrhage". In: *New England Journal of Medicine* 352.8 (Feb. 24, 2005), pp. 777–785.

[25] T. Morgan et al. "Preliminary findings of the minimally-invasive surgery plus rtPA for intracerebral hemorrhage evacuation (MISTIE) clinical trial". In: *Cerebral Hemorrhage*. Springer, 2008, pp. 147–151.

[26] T Morgan et al. "Preliminary report of the clot lysis evaluating accelerated resolution of intraventricular hemorrhage (CLEAR-IVH) clinical trial". In: *Cerebral Hemorrhage*. Springer, 2008, pp. 217–220.

[27] W. Andrew Mould et al. "Minimally Invasive Surgery Plus Recombinant Tissue-type Plasminogen Activator for Intracerebral Hemorrhage Evacuation Decreases Perihematomal Edema". In: *Stroke* 44.3 (Mar. 1, 2013), pp. 627–634.

[28] John Muschelli et al. "fslr: Connecting the FSL Software with R". In: *R Journal* 7.1 (2015), pp. 163–175.

[29] John Muschelli et al. "Validated automatic brain extraction of head CT images". In: *NeuroImage* (2015).

[30] Salvador Pedraza et al. "Reliability of the ABC/2 Method in Determining Acute Infarct Volume". In: *Journal of Neuroimaging* 22.2 (2012), pp. 155–159.

[31] K. N. Bhanu Prakash et al. "Segmentation and quantification of intra-ventricular/cerebral hemorrhage in CT scans by modified distance regularized level set evolution technique". In: *International Journal of Computer Assisted Radiology and Surgery* 7.5 (Sept. 1, 2012), pp. 785–798.

[32] Noel Prez et al. "Set of methods for spontaneous ICH segmentation and tracking from CT head images". In: *Progress in Pattern Recognition, Image Analysis and Applications*. Springer, 2007, pp. 212–220.

[33] Adnan I. Qureshi et al. "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 (Dec. 1, 2011), pp. 428–435.

[34] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2015.

[35] Chris Rorden and Matthew Brett. "Stereotaxic Display of Brain Lesions". In: *Behavioural Neurology* 12.4 (2000), pp. 191–200.

[36] Christopher Rorden et al. "Age-specific CT and MRI templates for spatial normalization". In: *NeuroImage* 61.4 (July 16, 2012), pp. 957–965.

[37] Natalia S. Rost et al. "Prediction of Functional Outcome in Patients With Primary Intracerebral Hemorrhage The FUNC Score". In: *Stroke* 39.8 (Aug. 1, 2008), pp. 2304–2309.

[38] Ramandeep Sahni and Jesse Weinberger. "Management of intracerebral hemorrhage". In: *Vascular Health and Risk Management* 3.5 (Oct. 2007), pp. 701–709.

[39] Eric E. Smith, Jonathan Rosand, and Steven M. Greenberg. "Imaging of Hemorrhagic Stroke". In: *Magnetic Resonance Imaging Clinics of North America* 14.2 (May 2006), pp. 127–140.

[40] Stephen M. Smith. "Fast robust automated brain extraction". In: *Human Brain Mapping* 17.3 (2002), pp. 143–155.

[41] Robert Tibshirani. "Regression Shrinkage and Selection via the Lasso". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 58.1 (1996), pp. 267–288.

[42] Stanley Tuhrim et al. "Volume of ventricular blood is an important determinant of outcome in supratentorial intracerebral hemorrhage". In: *Critical care medicine* 27.3 (1999), pp. 617–621.

[43] Alastair J. S. Webb et al. "Accuracy of the ABC/2 Score for Intracerebral Hemorrhage Systematic Review and Analysis of MISTIE, CLEAR-IVH, and CLEAR III". In: *Stroke* 46.9 (Sept. 1, 2015), pp. 2470–2476.

[44] Simon N. Wood. "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 (2011), pp. 3–36.

[45] Simon N. Wood, Yannig Goude, and Simon Shaw. "Generalized additive models for large data sets". In: *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 64.1 (Jan. 1, 2015), pp. 139–155.

## 5. Appendix

### *5.1. 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 $k$, $k = 1, \ldots 21$.

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

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

| term | estimate |
|---|---|
| Intercept | 1.0077 |
| Neighborhood mean | 0.0514 |
| Neighborhood sd | 0.0001 |
| Neighborhood skew | 0.0650 |
| Neighborhood kurtosis | -0.3515 |
| Image intensity (HU) | -0.1723 |
| Threshold (>= 40 and <= 80) | -0.1514 |
| Z-score plane 1 | -0.6322 |
| Z-score plane 2 | -0.2493 |
| Z-score plane 3 | 1.0369 |
| Winsorized z-score (20\% trim) | 0.5475 |
| Percentage thresholded neighbors | 2.0612 |
| Atropos probability image | 0.1497 |
| Percent of zero neighbors | -9.1802 |
| Indicator of any zero neighbors | 0.0708 |
| Distance to Image centroid | -0.0870 |
| Gaussian smooth (5mm^3) | -0.0507 |
| Gaussian smooth (10mm^3) | 0.5499 |
| Gaussian smooth (20mm^3) | -0.3897 |
| Z-score-to-template image | 1.4598 |
| Contralateral difference | 0.0326 |

The specification for the functional form of the model fit with the LASSO penalty, is the same, but optimizes the following criteria:

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