Representation of the NIH Stroke Scale with Probabilistic Diffusion Weighted Imaging Lesion Atlas

K. W. Singleton[1], T. J. Schaewe[2], W. J. Boscardin[3], M. Luby[1], S. Warach[1], C. S. Kidwell[4], and J. R. Alger[2]

[1]Section on Stroke Diagnostics and Therapeutics, National Institute of Neurological Disorders and Stroke, NIH, Bethesda, MD, United States, [2]Department of Neurology, David Geffen School of Medicine at UCLA, Los Angeles, CA, United States, [3]Department of Biostatistics, School of Public Health, UCLA, Los Angeles, CA, United States, [4]Department of Neurology, Georgetown University, Washington, DC, United States

Introduction: Diffusion weighted imaging (DWI) is an indicator of ischemic brain damage in acute stroke. It has been used together with other acute MR imaging data for the prediction of final infarct volume. However knowledge of lesion volume alone is only modestly helpful in defining what acute or long term functional deficits might be present. The National Institutes of Health Stroke Scale (NIHSS) is an 11 category, 15 item neurologic examination scale that has become the most widely used measure for quantifying clinical deficits associated with ischemic stroke. The hypothesis for this study is that probabilistic brain atlas techniques can be used to formulate statistically significant relationships between functional deficits as measured with the NIHSS and DWI lesion characteristics in acute stroke patients.

Methods: Data for this project was prospectively gathered from a natural history imaging repository at one of the participating institutions. Consented patients agree to allow imaging and clinical data to be archived for use in future studies, such as the present one. A population of acute stroke cases was identified using the following inclusion criteria: acute ischemic stroke; brain MRI, including DWI, performed within 24 hours of symptom onset, and baseline NIHSS performed within an hour of the MRI. Only cases for which an expert had outlined the DWI lesion on all image slices were included. A total of 163 studies that met the criteria were identified. Mean age was 70, range 21-97. Baseline NIHSS was median 4, mean 6, range 0-27. The population included 88 middle cerebral artery, 22 posterior cerebral artery, 10 anterior cerebral artery, 20 basilar artery, 3 vertebral artery, 16 lenticulostriate, 2 anterior choroidal artery, 4 watershed, and 15 unclassified strokes.

All DWI data was co-registered to a brain template. Software was written to identify voxels with a significant relationship between presence of lesion on the image and an abnormal score on a particular NIHSS item. The software used a contingency table approach to estimate for each voxel the chi square statistic and the one-tailed probability (p) of obtaining a chi square statistic of equal value or greater (i.e. a null hypothesis of no relationship between DWI hyperintensity and abnormal NIHSS item score). Because a large number of voxels were studied, there is a need for a multiple comparisons correction to account for voxels that have low p as a result of chance. We adopted the technique of false discovery rate (FDR) control which is frequently used for an analogous purpose in functional MRI studies. An FDR threshold of 5% was chosen and voxels that met this criterion were saved into 1-p and z score significance maps for each NIHSS item (see Figure 1). The FDR technique did not yield significant voxels for all NIHSS items.

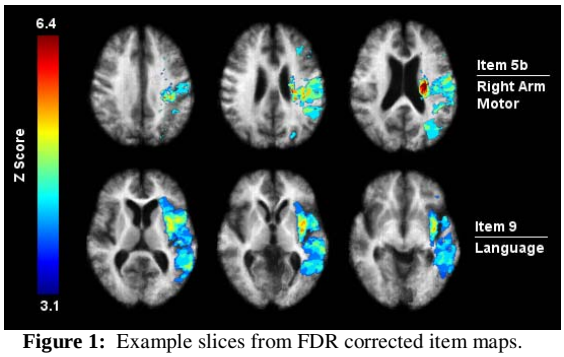

Figure 1: Example slices from FDR corrected item maps. Top row: Item 5b (motor arm) significant voxels correspond to the corticospinal tract. Bottom row: Item 9 (language) significant regions correspond to primary language areas.

Each expert segmented lesion from the study population (n=163) was compared individually to the item maps. The average 1-p and z scores over the lesion volume were computed for each item. The 1-p and z scores were evaluated using Receiver operating characteristic (ROC) curves generated using SPSS (see Figure 2). ROC curves compare the power of a model as a binary classifier as the detection threshold is varied. The optimal detection threshold for each item was chosen to correspond to the point on the ROC curve nearest to theoretically perfect detector performance (sensitivity = 1.0, specificity = 0.0; the upper left corner of the range of ROC values).

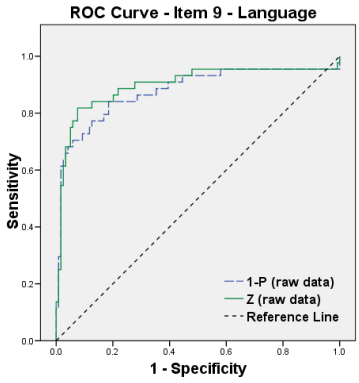

Figure 2: ROC curves for NIHSS item 9, Language. AUC for Z = 90%, AUC for 1-P = 88%

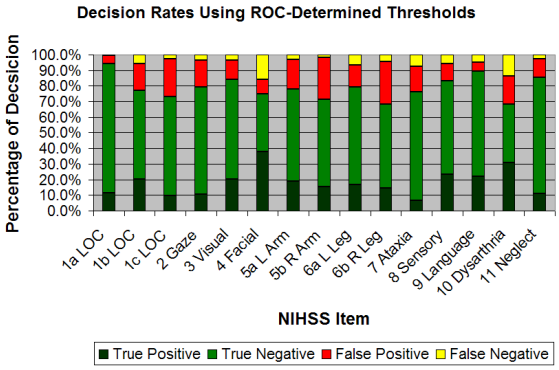

Figure 3: Decision rates for 163 cases using the optimal thresholds selected from ROC curves of raw data maps.

Results: Comparison of FDR corrected data and raw data showed only minor differences in indicator performace as measured by area under the ROC curve (AUC). Therefore, raw data was used to allow for comparison of outcome for all NIHSS items. AUC was greater than 70% for all items except for item 7, limb ataxia. Area under the curve was above 80% for 10 of 15 items. While 1-p and z scores performed similarly as indicators, z scores performed better in 12 of 15 items. Using the chosen optimal thresholds, correct decisions (true positive + true negative) were made in 79% of decisions (n=2445). Figure 3 provides a summary of performance for each stroke scale item.

Conclusions: We have developed the first neuroanatomic atlas to combine acute infarct location and volume information with NIHSS item information. Here we show that lesion overlap measures derived from this atlas may be used to detect individual NIHSS item deficits with a high degree of accuracy. Further improvements in performance are readily attainable. We observe that certain items, such as ataxia, have insufficient representation in the atlas. Incorporation of additional cases (n >> 163) into the atlas will address this issue. In addition, it is well known that perfusion deficits contribute to symptoms in acute stroke but are unrepresented in the current DWI-based atlas. We expect that integration of perfusion weighted imaging (PWI) will further refine the anatomic representation of acute functional deficit provided by the atlas.