Early prediction of salvageable tissue with multiparametric MRI-based algorithms after experimental ischemic stroke

M. J. Bouts[1], I. A. Tiebosch[1], R. Zwartbol[1], E. Hoogveld[1], O. Wu[1,2], and R. M. Dijkhuizen[1]

[1]Image Sciences Institute, University Medical Center Utrecht, Utrecht, Netherlands, [2]Athinoula A. Martinos center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, United States

Introduction: Early identification of salvageable tissue is crucial in management and treatment of patients suffering from acute, ischemic stroke. MRI-based measurement of a 'perfusion-diffusion mismatch' is increasingly applied to identify potentially salvageable tissue, but it may overestimate tissue at risk.[1,2,3] It has been postulated that predictive algorithms, that combine multiple MRI parameters to calculate a single risk index, more favorably identify tissue at risk.[2] However only few studies have actually measured the ability of differentiating salvageable tissue from tissue irrevocably destined to infarct[3], and compared the potential of different algorithms. We therefore evaluated the ability of several predictive algorithms in differentiating between salvageable and irreversibly damaged tissue, by comparing acute measures of infarct risk with subsequent outcome in rats treated with recombinant tissue plasminogen activator (rt-PA).

Material & Methods: In this study twelve, male Wistar rats were included. All procedures were approved by our institution's animal care committee. Rats were subjected to unilateral middle cerebral artery occlusion (MCAo) by either insertion of an intraluminal filament (permanent MCAo; group 1; n=6) or by injection of an autologous blood clot (embolic MCAo; group 2; n=6). For all animals, MRI was executed immediately after surgery, and at 3 days (group 1), or at 1 and 7 days post-stroke (group 2). Immediately after the first MRI session (ca. 2 h post-stroke) group 2 animals received rt-PA (10.0 ml/kg) intravenously. Multi-echo multislice T_2-weighted, diffusion-weighted, and dynamic susceptibility contrast-enhanced MRI were conducted on a 4.7T horizontal bore MR spectrometer (Varian, Palo Alto, CA, USA). Quantitative maps of the T_2, trace of the apparent diffusion coefficient (ADC_{trace}), cerebral blood flow index (CBF_i), cerebral blood volume (CBV), mean transit time (MTT), and tracer arrival delay were calculated from the acquired data.[4] Parametric maps were spatially aligned and maps were normalized (and expressed as relative) to unaffected contralateral gray matter. Four different models were selected. Prediction was either based on identification of a separating hyperplane in parameter space (n=2), separating infarcted from non-infarcted voxels, or using an ensemble of multiple decision trees (n=2). A separating plane was either derived by linear regression using a generalized linear model (glm) or a support vector machine (svm)[5] which maximized the margin between infarcted and non-infarcted voxels using a non-linear Laplace transformation kernel. Adaptive boosting[5] (adaboost) sequentially trained and adaptively weighted the input and output of a predefined number of decision trees using a logistic loss function to obtain an estimate of infarction risk. A fourth method used a bootstrap aggregate of the data samples in combination with random feature selection to create decision trees which together provided an estimate of infarction risk (RandomForest (rf)).[6] All models were trained based on a balanced dataset. Training voxels were taken from the permanent MCAo group and infarcted areas were identified on 3 day follow-up (F/U) T_2 data. Abnormal tissue was identified ipsilaterally as a deviation of more than two times the standard deviation from normal contralateral values. Prediction performance was tested and compared using a jack-knifing approach. Probabilistic maps were thresholded at the optimal point of operation represented by the maximal Youden's J statistic (J-index=specificity+sensitivity-1). Prediction performance was expressed as the area under the receiver operating curve (AUC) and dice's similarity index (DSI). DSI > 0.7 was considered excellent agreement, for AUC and J-index values approaching a maximum of 1.0 represented better performance. Subsequently an aggregate model was trained and used to predict tissue outcome in group 2. Hemispheric lesion fraction (HLFs) (affected tissue volume divided by total hemispheric brain volume) from derived parameters were compared to F/U lesion fractions. Ipsilateral areas were subdivided in regions that i) became infarcted ('irreversibly injured' ('irreversible')), ii) were at risk but became normal at F/U ('salvageable'), and iii) remained normal ('normal'). 'Normal' areas were identified by ipsilateral risk values below risk values representing optimal point of operation. Statistical differences were determined using repeated measures ANOVA or Friedman test with post hoc Tukey testing; $p<0.05$ was considered significant.

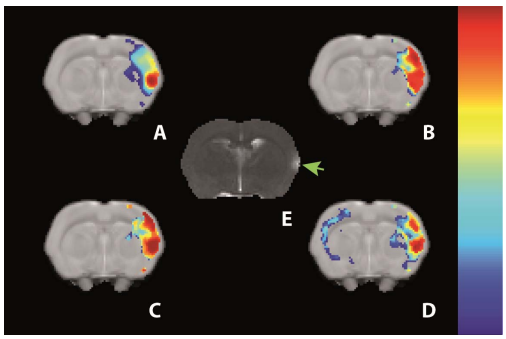

Figure 1: Acute post-stroke infarction risk maps calculated by glm (a), svm (b), adaboost (c), rf (d), overlaid on anatomical template. 7 day F/U T_2 map shows eventual infarct (arrow) (e). Risk maps were thresholded for optimal prediction. Color-coding indicates risk level (low (0.4 (blue)) to high (1.0 (red)).

Results: In group 1 median abnormal tissue ADC HLF (0.24; interquartile range (IQR):0.08-0.45) was significantly different from median perfusion (MTT) HLF (0.66; IQR:0.67-0.70; p=0.01). Tissue lesion volume had increased at F/U (0.62; IQR:0.40-0.73; p=0.05) to an area comparable to the acutely abnormal MTT volume (p=0.83). This progression was absent in group 2. In this group median acute ADC HLF (0.14; IQR:0.08-0.19) was equivalent to F/U HLFs (0.09; IQR:0.09-0.17; p=0.96). Acute MTT HLFs showed near significant difference (0.33; IQR: 0.27-0.37; p=0.05) from ADC HLFs. In contrast to acute MTT and F/U T_2 HLFs (p=0.002), acute ADC HLFs did not differ across groups (p=0.48). Training and evaluation resulted in 48 predictive maps expressing infarction risk. All predictive maps were thresholded at risk percentages representing maximum J-index (threshold: glm=41%; svm=47%; adaboost=59%; rf=41%). Comparison showed no significant difference in mean AUC (glm=0.89±.06; svm=0.91±.06; adaboost=0.92±.05; rf=0.92±.06), nor in overlap between areas of increased risk and infarcted areas on T_2 F/U (mean DSI: glm=0.72±.12; svm=0.73±.10; adaboost=0.71±.11; rf=0.73±.12). Subsequently, aggregate models were used to predict tissue outcome in group 2. All models produced maps of spatially differently distributed risk values. Glm showed increased AUC values compared to adaboost (p=0.05) and rf (p=0.09) (mean AUC glm=0.95±.02; svm=0.92±.04; adaboost=0.91±.02; rf=0.89±.07). A significant difference in risk values was observed for adaboost compared to glm and rf (p=0.004 and 0.03, resp.). Figure 1 shows an example of different distribution of risk values for glm (a), svm (b), adaboost (c), and rf (d) in a successfully treated animal as indicated by the minor cortical infarct on T_2 F/U (e). For all models the predicted risk in the 'irreversible' and 'salvageable' areas were significantly higher compared to values in 'normal' areas (p<0.001) (Figure 2). In addition, all models except adaboost (p=0.82) assigned significantly different risk values to the 'salvageable' area as compared to the 'irreversible' area. The most significant difference between these tissue conditions was determined with glm (0.17 ± .09, p=0.00003).

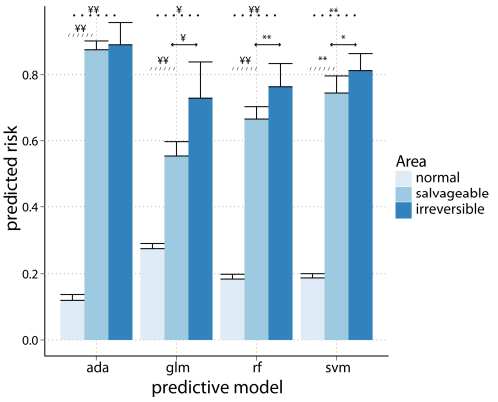

Figure 2: Predicted infarction risk for ipsilateral areas identified as 'normal', 'salvageable', and 'irreversible' in group 2 animals. (*p<0.01, **p<0.001, ¥ p<0.0001, ¥¥ p<0.00001)

Discussion: Four models for tissue outcome prediction were evaluated for their potential to identify salvageable tissue in a rat stroke model. All evaluated models performed equally well in predicting infarction after permanent MCAo. After acute embolic MCAo, the models showed significant differences in distribution of risk values. Non-linear models as adaboost and svm produced more homogeneous risk maps compared to glm. However, heterogeneity contributed to improved discernment of salvageable tissue that could be rescued by rt-PA treatment.

References: [1] Heiss, WD et al. *Int J Stroke*.5:290-5 (2010). [2] Østergaard L, et al. *Curr Opin Neurol.* 22:54-9 (2009). [3] Wu, O et al. *Brain*, 129:2384-93 (2006). [4] Dijkhuizen RM, et al. *J Cereb Blood Flow Metab.* 21: 964-71 (2001). [5] Bishop. CM. Pattern recognition and machine learning, *Springer Science+business Media* (2007). Breiman, L. *Machine learning* 45:5-32 (2001).