

# Intra-arterial therapy for basilar artery thrombosis: the role of machine learning in outcome prediction

*EJMINT* Original Article, 2014: **1449000234** (2nd December 2014) H Asadi, R Dowling, B Yan, P Mitchell

# Abstract

INTRODUCTION: Although only 10% of strokes are due to involvement of the posterior circulation, they can be associated with a very poor prognosis. Therefore, accurately predicting stroke outcome from a set of variables may identify high-risk patients and guide treatment approaches, leading to decreased morbidity. In this study, our aim was to design and compare different machine learning methods, capable of predicting the outcome of endovascular intervention in acute posterior circulation ischaemic stroke. METHODS: We conducted a retrospective analysis of a prospectively collected database of acute posterior circulation ischaemic stroke treated by endovascular intervention. Using SPSS®, MATLAB® and Rapidminer®, classical statistics as well as artificial neural network and support vector algorithms were applied to design a supervised machine capable of classifying these predictors into potential good and poor outcomes, as defined by 30 day mRS. RESULTS: We included 50 consecutive acute posterior circulation ischaemic stroke patients treated by endovascular technique. All the available demographic, procedural and clinical factors were included in the machine. The final confusion matrix of the neural network, demonstrated an overall congruency of ~90% between the target and output classes, with a relatively favourable overall receiving operative characteristic. However, after optimisation, the support vector machine had a relatively better performance, with a root mean squared error of 2.432 (SD: ±0.584). CONCLUSION: Consistent with the findings for anterior circulation, we showed promising accuracy of outcome prediction in posterior circulation strokes, suggesting that a robust machine learning system can potentially help in prognostication of acute posterior circulation stroke. Keywords: stroke, cerebrovascular occlusion, brain infarction - posterior circulation, brain infarction - anterior circulation, predictors.

## Introduction

#### Stroke and Endovascular Treatment

Stroke is a major global public health issue and is considered the third most costly health condition in developed countries [1]. Approximately 800,000 cases of stroke are reported in the USA per annum, leading to 200,000 deaths, and accounts for almost 1 of every 16 deaths [2,3]. For those who survive, it is the most common cause of adult disability in the modern world [2,4] requiring expensive long term rehabilitation care [2,5-7] amounting to costs estimated at over 60 billion dollars per year in the USA alone [2,5,8]. More than 80% of stroke cases are ischaemic, with the remainder being haemorrhagic [2].

**B Yan** — Melbourne Brain Centre, Department of Medicine, Royal Melbourne Hospital, University of Melbourne, Parkville, Victoria, Australia

Peter Mitchell (Corresponding Author) — Radiology Department, Royal Melbourne Hospital, Melbourne, Australia *Email:* peter.mitchell@mh.org.au

H Asadi – Melbourne Brain Centre, Department of Medicine, Royal Melbourne Hospital, University of Melbourne, Parkville, Victoria, Australia

**R Dowling** — Melbourne Brain Centre, Department of Medicine, Royal Melbourne Hospital, University of Melbourne, Parkville, Victoria, Australia



Urgent reperfusion of the ischaemic brain is the primary treatment aim, either by intravenous thrombolysis or by endovascular interventional techniques [9]. These treatments focus on vascular recanalisation and restoration of blood flow to the ischaemic tissue [10]. Although there are varying estimates of the number of patients who potentially benefit from endovascular intervention, there is likely to be an increased number of patients treated using these techniques [2,11,12].

Initial treatments focussed on intra-arterial thrombolysis, proposed to be safe up to 6 hours post onset [1,13]; however, rapid mechanical clot extraction with decreased time to cerebral reperfusion has obvious appeal and is theoretically ideal for platelet poor, fibrin rich, well organised cardiogenic emboli, refractory to mechanical lysis [2]. Therefore, subsequent development of various mechanical thrombectomy devices has gained much interest with the likely advantage of faster recanalisation and potential lower rate of haemorrhagic transformation; possibly leading to an extended time window for stroke intervention [1].

However, despite recanalisation success rates of more than 80%, randomised controlled trials such as Interventional Management of Stroke (IMS) - III [1,14] have still failed to show a significant improvement in the clinical outcome, evaluated by 90 days modified Rankin Scale (mRS) score [2,10,15-21]. The SYNTHESIS trial also failed to show any superiority of endovascular intervention or even combined endovascular and intravenous thrombolysis over traditional intravenous tPA [22-24]. This conundrum was further complicated when the MR-RESCUE trial demonstrated not only that embolectomy was no better than standard care, but also that a favourable penumbral pattern on imaging does not necessarily indicate patients who would benefit from endovascular therapy [25].

This discrepancy between the IMS-III, SYNTHESIS and MR-RESCUE outcomes and what may have been intuitively expected, is likely related to the multiple potential pitfalls in the design of these trials, which could influence the interpretation of the results [1,26,27]. The most commonly hypothesised factor is that patient selection was neither targeted to those who failed IV thrombolysis, nor to those with large vessel occlusion or large clot burden  $\geq$  8mm, who are usually not responsive to chemical treatment alone, since vascular imaging was not required prior to inclusion into the studies [1,26-28]. These limitations could certainly influence the accuracy of the studies in evaluation of the clot retrieval techniques. On the other hand, stent-retrievers, now acknowledged as more effective devices, were included only very late into the studies such as IMS-III, with fewer than 1% of cases treated using Solitaire. Since the release of these preliminary results, at least 6 additional devices have started premarket testing [1,26-28].

To complicate matters further, it appears that the situation is different for posterior circulation involvement. Although these cause only 6-10% of large vessel strokes, posterior circulation occlusions have a relatively different course and failure of recanalisation, in particular in comatose patients or those with basilar trunk involvement, and results in a very poor prognosis [2,29].

Surprisingly, the BASICS (Basilar Artery International Cooperation Study) did not show a definite superiority for intra-arterial intervention over intravenous thrombolysis [2,30], and the overall outcome is quite variable in patients who are treated with intra-arterial or intravenous thrombolysis, in particular depending on the therapeutic delay [2,31]. On the other hand, some trials have already demonstrated recanalisation success rates of more than 50% for intra-arterial techniques, with relatively good outcome [2,32]. However, randomised control trials are restricted and limited by the lower incidence of posterior



circulation strokes, and the results are potentially influenced by the heterogeneity of both the presentations and the causes; at this stage, the rationale for aggressive treatment is mainly based on anecdotal evidence [2,15,33,34].

Overall, the major obstacle in endovascular intervention of the posterior circulation ischaemic stroke is to establish a set of criteria identifying those patients who may benefit from intervention, whilst avoiding potential unwanted catastrophic treatment related complications.

There is currently level I evidence that NIHSS (National Institute of Health Stroke Score) [35,36] is a quick and relatively simple guide to estimate the extent and severity of a stroke, and probably correlates with the clinical outcome [35]. It is, however, unable to measure the size of established infarction separate from the salvageable parenchyma; therefore, it is unable to predict potential outcome after endovascular intervention accurately. This is consistent with the well known fact that multiple factors contribute to and influence recanalisation success, including the extent and site of the vascular occlusion; the overall outcome also depends on patient demographic factors as well as clinical setting such as the time from onset, duration and the severity of the presenting neurological insult [10].

The complexity of the all of these factors involved makes prediction of the final outcome difficult. On the other hand, accurately predicting the outcome from a set of predictive variables is an important aspect of clinical work, which can assist in identifying high-risk patients and guide treatment approaches, thus potentially decreasing morbidity and mortality. Such a model in prediction of the outcome not only may be crucial in prognosis, but can also have future roles in patient selection for the variety of the treatment options available and the relevant studies.

## Prognostic Modeling and Machine Learning

The usual approach to analyse stroke outcomes data is to develop logistic regression models; however, machine learning algorithms have been proposed as an alternative in particular for large-scale multiinstitutional data, with the advantage of easily incorporating newly available data to improve prediction performance [37,38].

Machine Learning algorithms can be applied and trained under two different scenarios: supervised and unsupervised. In supervised scenarios, the predicted outputs are known and used to train the models. In unsupervised machines, the desired output is unknown and the objective is to discover structure in the data, not to generalise a mapping from inputs to outputs.

Two of the most commonly used machine learning methods include artificial neural network and support vector machine. These models are trained supervised, with neural network algorithms capable of unsupervised training as well.

Although the technical details of these algorithms are beyond the scope of this article, a summary of them follows:

Artificial neural network is a mathematical and computational model that is inspired by the structure and functional aspects of biological neural systems [38,39]. It consists of interconnected nodes, processing information using a connectionist computational approach. The central connectionist principle proposes



that complex neurological and mental phenomena can be described by an interconnected network of simple uniform units [39], adaptively changing their structure based on external or internal information, which flows during the learning phase, forming a robust dynamic system modelling the complex relationships between inputs and outputs or patterns in data [38-40].

From the different topological types of neural networks, the commonly used feed-forward is a network where connections between the units do not form a directed cycle or loop, and the information moves in only one direction—forward—from the input nodes through the hidden nodes to the output nodes [39]. Back propagation algorithm is a supervised learning method divided into propagation and weight update phases, which are repeated until the performance of the network is good enough, while the output values are compared with the correct answer to compute the value of some predefined error-function [38,39]. This calculated error is then fed back through the network, adjusting the weights of each connection accordingly in order to reduce the error function [39]. Repeating this process usually eventually converges to some state where the error of the calculations is minimised, at which point the network is considered trained for a certain target function [38].

In comparison to the artificial neural network, the support vector machine works very differently. It takes a set of input data and predicts which of the different possible classes comprises the input, making it a non-probabilistic linear classifier. A set of training data is given and marked as belonging to one of the categories. An SVM training algorithm builds a model that assigns new data into one category or the other.

The example data points are initially mapped as points in space so that the examples of the separate categories are divided by a clear gap that is as wide as possible and then unknown data is represented in that same space and predicted to belong to a category based on which side of the gap they fall [38-40]. In doing so, the algorithm constructs a hyperplane or a set of hyperplanes in an infinite-dimensional space, which can be used for classification, regression or other tasks. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data points of any class [38-40]. This gap is called functional margin and in general the larger the margin the lower the generalisation error of the classifier [38-40].

Whereas the original problem may be stated in a finite dimensional space, it often happens that the sets to discriminate are not linearly separable in that space. For this reason, it was proposed that the original finite-dimensional space be mapped into a much higher-dimensional space, presumably making the separation easier in that space [38-40]. To keep the computational load reasonable, the mapping is designed to ensure that dot products may be computed easily in terms of the variables in the original space, by defining them in terms of a kernel function K (x, y) selected to suit the problem. The hyperplanes in the higher dimensional space are defined as the set of points whose inner product with a vector in that space is constant [38-40].

## Our Study

Following the promising results from modelling of the anterior circulation stroke [41] and using the same techniques, we aimed to design a prognostic model for the endovascular intervention in acute posterior circulation ischaemic stroke using machine learning algorithms. We decided to focus on the posterior circulation strokes separately and model them independently to avoid potential inadvertent underlying inhomogeneities.



We compared and assessed these 2advanced methods in terms of their capability for predicting outcome.

#### Method

This is a retrospective study on a prospectively collected completely de-identified clinical database, which received approval from the Ethics Committee at our institution; our Review Board waived the need for consent (HREC: QA2011100).The technical details are provided below to facilitate reproducibility for other datasets if available.

Demographics and clinical details of 50 patients who presented with acute posterior circulation stroke to our institution and underwent endovascular treatment over a period of approximately 5 years were extracted from a prospectively maintained stroke database.

Patients were screened for relevant comorbidities at the time of presentation including: diabetes mellitus, hypertension, hypercholesterolaemia, atrial fibrillation, history of ischaemic heart disease and previous cerebral stroke or transient ischaemic attack. Neurological examination was performed for all of patients prior to any intervention and baseline National Institute of Health Stroke Scores were recorded in the database.

From the initial diagnostic angiogram occluded vessels were identified. In the case of multiple sequential occlusions, the proximal vessel was used as a data point and, depending on the extent and segments involved, the artery was categorised as first and second occlusion.

Some of the patients also had IV-tPA prior to endovascular intervention. Different endovascular recanalisation devices were used, including the Solitaire stent-retriever and MERCI device.

In addition to mechanical thrombectomy, some cases also received intra-arterial chemical thrombolytic agents and, if present, associated or post-recanalisation hemodynamically significant stenoses were also treated with angioplasty or stent insertion.

After treatment of the occluded artery(s), recanalisation success was assessed using Thrombolysis in Cerebral Infarction (TICI) Score by the blinded consensus of the treating neurointerventionalists. TICI score in conjunction with the number of attempts for recanalisation, procedure duration, and time of onset to recanalisation, as well as patient general anesthesia status, were all recorded.

All procedural or delayed post-procedural complications were also recorded, including arterial perforation and puncture site haematoma or pseudoaneurysm.

Post-procedure CT scans of the brain at 24-36h were all assessed by neuroradiologist and neurointerventionists for the presence of acute stroke and intracranial haemorrhage. Intracranial haemorrhagic transformations were divided into clinically silent or symptomatic and then classified into different categories.

Procedural outcome was assessed using mRS, measured 90 days after onset. A final dichotomised good and bad outcome was also recorded for the patients as per mRS, with less than or equal 2considered as good.



First, using SPSS® (IBM Corporation), a Standard Linear Model was designed, using forward-stepwise as the model selection method and Information Criterion (AICC) as the criteria for entry. Potential predictors of the mRS as the outcome measure were identified and a prediction model was formed and compared with the observed outcome for validation.

Supervised machine learning was then attempted. Initially using MATLAB® (MathWorks Inc.) and its Neural Network Toolbox, a two-layer Feed-Forward network with sigmoid hidden and linear output neurons was designed. The data was then randomly divided into 70, 15 and 15 percent subsets and the network was trained using Levenberg-Marquardt algorithm, validated and tested using the mRS as outcome, with the performance of the model monitored using Mean Squared Error. Prediction errors were also depicted on a histogram. In addition, for comparison, the network was also trained using the dichotomised mRS,>2 or  $\leq 2$ , to evaluate a binary classifier for potential good and poor outcomes.

For the seven scale mRS network, linear regressions were also performed between the observed and estimated outcome over the training, validation and test datasets independently using Theil-Sen estimator. However, with the dichotomised model being a binary classifier, Receiver Operating Characteristic(ROC) curves were calculated to illustrate the performance of the system over each dataset as its discrimination threshold is varied. Additionally, confusion matrices or contingency tables were also calculated, allowing better representation of the performance of the network.

The designed network and its calculated weighting matrix was then saved to be imported into the Simulink Toolbox of MATLAB® (MathWorks Inc.) for outcome prediction of the future data. Subsequently, to assess the capabilities of other supervised machine learning systems, the dataset with scaled and dichotomised mRS were imported into the data-mining program, Rapidminer® (Rapid-I Inc.). The filtered data was then given to the input training port of a nested cross-validation operand, with the relative number of validation of 10% and a shuffled sampling type as well as "Leave One Out". The cross-validation operand consisted of two components: training and testing. The testing component contained a Support Vector Machine, with ANOVA Kernel, which is defined by raised to power "d" of summation of "exp(-g (x-y))" where "g" is gamma and "d" is the degree: "g" and "d" were set to be 1 and 2 in our machine. The size of the cache for kernel evaluations was set to be 200 megabytes. The complexity constant ("C"), which sets the tolerance for misclassification, was set to 0. The convergence epsilon, which is an optimiser parameter specifying the iterations stop point, was set to 0.001, with maximum iteration set to 100,000. In our machine, the loss function positive and negative complexity constant was set to 1.0. The insensitivity constant, epsilon as well as the epsilon for positive and negative deviations, were all set to 0.

The model calculated in this machine was passed onto the testing component of the parent x-validation operand and then applied on to the test dataset. The performance of the machine was monitored by a classic performance monitor operand and was reported as the mean squared error as well as its root. In addition, accuracy of the machine was assessed by aggregation of a hidden confusion matrix constructed by evaluating different models on different test sets. The designed model was finally incorporated into an apply operand ready for the prediction of the outcome of the future patients.



# Results

The median of patients' baseline NIHSS was 12, with average of 20.6 (SD:  $\pm$  15.5).Only 7 of our patients also had IV-tPA prior to endovascular intervention. The remainder of the patients did not receive tPA due to a variety of contraindications, mainly time of onset to treatment delay. As expected for posterior circulation, the majority of the procedures (88%) were performed under general anaesthesia.

The Solitaire stent-retriever was used in the majority (19) of the mechanically thrombectomised cases, and the MERCI device was used in 4 cases. In some cases instrumentation was repeated up to 5 times to improve recanalisation. More than half of the patients (31 cases) received urokinase and 16 patients had intra-arterial tissue plasmin activator infusion as adjunct intra-arterial chemical thrombolytic agents.

Overall, recanalisation was relatively successful with TICI 2b or 3 demonstrated on the final angiographic run in approximately 50% of cases. Significant associated and post-recanalisation arterial stenosis was also noted in some cases, with 15 patients requiring angioplasty, and 12 patients eventually stented. The average duration of procedures was97.2min (SD:  $\pm$  43), and time of onset to recanalisation was on average 660min (SD:  $\pm$  374).

Immediate procedural complications were uncommon, with only 1 case of arterial perforation and no puncture site haematoma or pseudoaneurysms noted. Sixteen patients (32%) were diagnosed with intracranial haemorrhage on the delayed post-procedural CT, with a spectrum of locations and severities, from subarachnoid haemorrhage to asymptomatic or large intra-parenchymal bleeds.

The average of mRSat90 days was 3.9 (SD:  $\pm$  2.4), with median of 5 and mode of 6.

## Standard Modelling

The information criterion and accuracy of the proposed linear model were calculated as 53.4 and 55.1% respectively. The most influential predictor was baseline NIHSS, with relative predictive value of ~0.8.

#### Artificial Neural Network

The best validation performance was 6.3 and 0.03, at epoch 2 and 27 for seven scale and dichotomised mRS models respectively.

Gradient of 5.228 x 10-11 and 2.333 x 10-4 were calculated at epoch 6 and 33for seven scale and dichotomised mRS models respectively. Error histograms were calculated as the difference between the target and output which are equivalent of observed and estimated outcome, from the training, test and validation datasets, for seven scale and dichotomised mRS models.

Using Theil-Sen estimator, the root of the Coefficient of Determination was calculated as 0.96, 0.26 and 0.18 for each subset respectively. However, overall network estimated and observed outcome for the whole dataset demonstrate a relatively good linear correlation with an R of approximately 0.72 in a linear regression.

There was favourable overall ROC curve; however, the test curve is very poor, with the estimated area under curve (AUC) of 0.35. The contingency table, with each column representing the instances of the



predicted outcome and each row demonstrating the observed outcome, confirmed acceptable model sensitivity and specificity.

#### Support Vector Machine

For the scaled mRS outcome, our support vector machine had a good performance with a mean squared error of 6.257 (SD:  $\pm$  3) and the estimated root at 2.432 (SD:  $\pm$  0.584). Also, the system accuracy was assessed by "mikro", calculated as 2.5. On the other hand, system's MSE and "mikro" were calculated as 0.222 (SD:  $\pm$  0.096), and 0.471, respectively, in prediction of the dichotomised outcome, with the precision of 68% predicting poor outcomes and an overall precision of approximately 70% with a model accuracy of66% and an AUC of 0.5.

Best performance of the scaled mRS model was in prediction of the patients with poor prognosis with mRS of 6, with a class precision up to48%; however, for the remainder, prediction was quite imprecise in the scaled compared to dichotomised model. Estimated machine performance improved even further when the cross validation operand set to work with "Leave Out One" sampling rather than "Shuffled", with a MSE of 6.112 (SD:  $\pm$  7.307) and rMSE of 1.984 (SD:  $\pm$  1.475) for the scaled mRS outcome and0.263 (SD:  $\pm$  0.317) and rMSE of 0.408 (SD:  $\pm$  0.310)for the dichotomised model. The "mikro" indicator of accuracy, was calculated as 2.472 and 0.512, with the "Leave Out One" sampling for the scaled mRS and dichotomised outcome predictor machines, respectively.

# Conclusions

We showed that despite a small dataset, a modest prediction accuracy of 70% was attainable, and there is the likely potential of further improving prediction by incorporating larger multicentre datasets.

There has been recent interest in adopting machine learning techniques in the prediction of the outcome of stroke patients. A recent study submitted for publication has shown promising results in modelling the outcome in patients with anterior circulation ischaemic stroke and with relatively good accuracy and precision in prediction of the final mRS using artificial neural network modelling and support vector machine algorithms. One study has proposed spatial regularisation of the diffusion-weighted images acquired at the acute stage using support vector machine with a Graph encoding the voxels' proximity, and found significant accuracy in prediction of the motor outcome at 90 days, showing that poor motor outcome is associated with the changes in the corticospinal bundle and white matter tracts originating from the premotor cortex [42].

Another study has proposed use of machine learning in individualised stroke treatment decision making by accurate identification of the extent of salvageable tissue on MRI in rats based on measurement of a perfusion-diffusion mismatch and calculation of infarction probability. This study compared generalised linear model (GLM), generalised additive model, support vector machine, adaptive boosting and random forest, proposing that assessment of the heterogeneity of infarction probability with imaging based algorithms enables estimation of the extent of potentially salvageable tissue after acute ischaemic stroke [43].

Conversely, congruent with anterior circulation ischaemic stroke, attempts to prove the effectiveness of the invasive posterior circulation stroke treatments have shown inconsistent results. However, more than



ever before, endovascular treatments of acute posterior circulation ischaemic stroke are evolving to mainstream management in particular given the mortality particularly for those patients with no other available therapeutic options, e.g. contraindication for IV thrombolysis and large vessel occlusions [10,44-48].

To our knowledge there is no comprehensive multifactorial study in humans attempting to apply machine learning algorithms in acute posterior circulation ischaemic stroke outcome prediction after invasive endovascular management. Undoubtedly, numerous factors including extensive clinical heterogeneity, can influence the final stroke outcome with varying significance and mechanisms, making conventional modelling challenging and potentially inaccurate. However, machine learning models, which are relatively independent of the unknown potential underlying interactions between these factors, are probably able to simulate the eventual result of such a complex system. Such models may be of use not only for prognosis and in predicting outcomes under different circumstances, but hopefully in the near future to assist in clinical decision-making, in particular identifying those patients who may benefit from a variety of possible treatment options, including more aggressive management such as endovascular interventions.

## Limitations

Parallel to the all abovementioned advantages of the machine learning algorithms, there are important underlying assumptions and limitations that should be noted. Although these models can be accurate and perhaps useful in answering the primary question, these complex algorithms require large training datasets to improve their performance, with the true underlying relationships between influential factors remaining undiscovered to the user [37,38,40].

This inherent need for large training datasets may affect the accuracy of the machines in studies like the current study when only representative training data is used. In addition, with no clear understanding of the true predictors, an overcorrected conservative design may lead to the models being over-fitted by irrelevant demographics or clinical factors, thus increasing the random error and covering the desired signal with noise, a phenomenon which may explain the poor ROC curve for the test group in this study. To avoid this, techniques like cross-validation, regularisation, pruning or Bayesian model comparison can be used to indicate the tipping point when further training no longer results in a better performance or alternatively decision tree learning methods can be employed, providing more interpretable models [37,38,40].

#### Future work

Putting the underlying methodological and computational complexities aside, our long term goal is to design an easy to use online system which facilitates relative prediction of the clinical outcome based on demographics and clinical findings, which can be used as a guide to therapeutic decision making.

Such a system has the potential for fine adjustment from the continuous training provided via handling large-scale national or international multi-institutional users, with the advantage of easily incorporating newly available data to improve prediction performance.



#### **Conflict of Interest**

We declare that we have no conflict of interest.



#### References

- 1. Ferrell AS, Britz GW (2013) Developments on the horizon in the treatment of neurovascular problems. *Surg Neurol Int* 4: S31-37.
- 2. Blackham KA, Meyers PM, Abruzzo TA, Albuquerque FC, Fiorella D, et al. (2012) Endovascular therapy of acute ischemic stroke: report of the Standards of Practice Committee of the Society of NeuroInterventional Surgery. *J Neurointerv Surg* 4: 87-93.
- 3. Lloyd-Jones D, Adams R, Carnethon M, De Simone G, Ferguson TB, et al. (2009) Heart disease and stroke statistics--2009 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation* 119: 480-486.
- 4. Sacco RL, Benjamin EJ, Broderick JP, Dyken M, Easton JD, et al. (1997) American Heart Association Prevention Conference. IV. Prevention and Rehabilitation of Stroke. Risk factors. *Stroke* 28: 1507-1517.
- 5. Hacke W, Furlan AJ, Al-Rawi Y, Davalos A, Fiebach JB, et al. (2009) Intravenous desmoteplase in patients with acute ischaemic stroke selected by MRI perfusion-diffusion weighted imaging or perfusion CT (DIAS-2): a prospective, randomised, double-blind, placebo-controlled study. *Lancet Neurol* 8: 141-150.
- 6. Adelman SM (1981) The National Survey of Stroke. Economic impact. Stroke 12: 169-87.
- 7. Taylor TN, Davis PH, Torner JC, Holmes J, Meyer JW, et al. (1996) Lifetime cost of stroke in the United States. *Stroke* 27: 1459-1466.
- 8. Rosamond W, Flegal K, Friday G, Furie K, Go A, et al. (2007) Heart disease and stroke statistics-2007 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation* 115: e69-171.
- 9. Barreto AD, Alexandrov AV (2012) Adjunctive and alternative approaches to current reperfusion therapy. *Stroke* 43: 591-598.
- Agrawal A, Golovoy D, Nimjee S, Ferrell A, Smith T, et al. (2012) Mechanical thrombectomy devices for endovascular management of acute ischemic stroke: Duke stroke center experience. *Asian J Neurosurg* 7: 166-170.
- 11. Cloft HJ, Rabinstein A, Lanzino G, Kallmes DF (2009) Intra-arterial stroke therapy: an assessment of demand and available work force. *AJNR Am J Neuroradiol* 30: 453-458.
- 12. Hirsch JA, Yoo AJ, Nogueira RG, Verduzco LA, Schwamm LH, et al. (2009) Case volumes of intraarterial and intravenous treatment of ischemic stroke in the USA. *J Neurointerv Surg* 1: 27-31.



- Furlan A, Higashida R, Wechsler L, Gent M, Rowley H, et al. (1999) Intra-arterial prourokinase for acute ischemic stroke. The PROACT II study: a randomized controlled trial. Prolyse in Acute Cerebral Thromboembolism. JAMA 282: 2003-2011.
- 14. Khatri P, Hill MD, Palesch YY, Spilker J, Jauch EC, et al. (2008) Methodology of the Interventional Management of Stroke III Trial. *Int J Stroke* 3: 130-137.
- (2009) The penumbra pivotal stroke trial: safety and effectiveness of a new generation of mechanical devices for clot removal in intracranial large vessel occlusive disease. Stroke 40: 2761-2768.
- Kulcsar Z, Bonvin C, Pereira VM, Altrichter S, Yilmaz H, et al. (2010) Penumbra system: a novel mechanical thrombectomy device for large-vessel occlusions in acute stroke. AJNR Am J Neuroradiol 31: 628-633.
- 17. Tenser MS, Amar AP, Mack WJ (2011) Mechanical thrombectomy for acute ischemic stroke using the MERCI retriever and penumbra aspiration systems. *World Neurosurg* 76: S16-23.
- 18. Jeromel M, Milosevic ZV, Kocijancic IJ, Lovric D, Svigelj V, et al. (2013) Mechanical revascularization for acute ischemic stroke: a single-center, retrospective analysis. *Cardiovasc Intervent Radiol* 36: 338-345.
- 19. Bosel J, Hacke W, Bendszus M, Rohde S (2012) Treatment of acute ischemic stroke with clot retrieval devices. *Curr Treat Options Cardiovasc Med* 14: 260-272.
- 20. Cohen JE, Gomori JM, Leker RR, Moscovici S, Ramirez-Denoriega F, et al. (2012) Recanalization with stent-based mechanical thrombectomy in anterior circulation major ischemic stroke. *J Clin Neurosci* 19: 39-43.
- 21. Day JS, Hurley MC, Chmayssani M, Rahme RJ, Alberts MJ, et al. (2011) Endovascular stroke therapy: a single-center retrospective review. *Neurosurg Focus* 30: E10.
- 22. Ciccone A, Valvassori L (2013) Endovascular treatment for acute ischemic stroke. *N Engl J Med* 368: 2433-2434.
- 23. Ciccone A, Valvassori L, Nichelatti M (2011) SYNTHESIS expansion: design of a nonprofit, pragmatic, randomized, controlled trial on the best fast-track endovascular treatment vs. standard intravenous alteplase for acute ischemic stroke. *Int J Stroke* 6: 259-265.
- 24. Ciccone A, Valvassori L, Nichelatti M, Sgoifo A, Ponzio M, et al. (2013) Endovascular treatment for acute ischemic stroke. *N Engl J Med* 368: 904-913.
- 25. Kidwell CS, Jahan R, Gornbein J, Alger JR, Nenov V, et al. (2013) A trial of imaging selection and endovascular treatment for ischemic stroke. *N Engl J Med* 368: 914-923.



- 26. Riedel CH, Zimmermann P, Jensen-Kondering U, Stingele R, Deuschl G, et al. (2011) The importance of size: successful recanalization by intravenous thrombolysis in acute anterior stroke depends on thrombus length. *Stroke* 42: 1775-1777.
- 27. Saqqur M, Uchino K, Demchuk AM, Molina CA, Garami Z, et al. (2007) Site of arterial occlusion identified by transcranial Doppler predicts the response to intravenous thrombolysis for stroke. *Stroke* 38: 948-954.
- 28. Zaidat OO, Lazzaro MA, Gupta R, Yavagal DR, Rasmussen PA, et al. (2012) Interventional Management of Stroke III trial: establishing the foundation. *J Neurointerv Surg* 4: 235-237.
- 29. Levy EI, Firlik AD, Wisniewski S, Rubin G, Jungreis CA, et al. (1999) Factors affecting survival rates for acute vertebrobasilar artery occlusions treated with intra-arterial thrombolytic therapy: a metaanalytical approach. *Neurosurgery* 45: 539-545; discussion 545-538.
- 30. Schonewille WJ, Wijman CA, Michel P, Rueckert CM, Weimar C, et al. (2009) Treatment and outcomes of acute basilar artery occlusion in the Basilar Artery International Cooperation Study (BASICS): a prospective registry study. *Lancet Neurol* 8: 724-730.
- 31. Lindsberg PJ, Mattle HP (2006) Therapy of basilar artery occlusion: a systematic analysis comparing intra-arterial and intravenous thrombolysis. *Stroke* 37: 922-928.
- 32. Smith WS (2007) Intra-arterial thrombolytic therapy for acute basilar occlusion: pro. *Stroke* 38: 701-703.
- 33. Smith WS, Sung G, Starkman S, Saver JL, Kidwell CS, et al. (2005) Safety and efficacy of mechanical embolectomy in acute ischemic stroke: results of the MERCI trial. *Stroke* 36: 1432-1438.
- 34. Nogueira RG, Yoo AJ, Buonanno FS, Hirsch JA (2009) Endovascular approaches to acute stroke, part 2: a comprehensive review of studies and trials. *AJNR Am J Neuroradiol* 30: 859-875.
- 35. Gonzalez RG, Copen WA, Schaefer PW, Lev MH, Pomerantz SR, et al. (2013) The Massachusetts General Hospital acute stroke imaging algorithm: an experience and evidence based approach. J Neurointerv Surg 5 Suppl 1: i7-12.
- 36. Jauch EC, Saver JL, Adams HP, Jr., Bruno A, Connors JJ, et al. (2013) Guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 44: 870-947.
- 37. Warwick K (2004) March of the machines : the breakthrough in artificial intelligence. Urbana: University of Illinois Press. x, 307 p. p.
- 38. Alpaydin E (2010) Introduction to machine learning. Cambridge, Mass.: MIT Press. xl, 537 p. p.
- 39. Russell SJ, Norvig P, Davis E (2010) Artificial intelligence : a modern approach. Upper Saddle River: Prentice Hall. xviii, 1132 p. p.



- 40. Warwick K (2012) Artificial intelligence : the basics. New York: Routledge. viii, 183 p. p.
- 41. Asadi H DR, Yan B, Mitchell P (2013) Machine Learning for Outcome Prediction of Acute Ischemic Stroke Post Intra-arterial therapy. PLoS One.
- 42. Cuingnet R, Rosso C, Lehericy S, Dormont D, Benali H, et al. (2010) Spatially regularized SVM for the detection of brain areas associated with stroke outcome. *Med Image Comput Comput Assist Interv* 13: 316-323.
- 43. Bouts MJ, Tiebosch IA, van der Toorn A, Viergever MA, Wu O, et al. (2013) Early identification of potentially salvageable tissue with MRI-based predictive algorithms after experimental ischemic stroke. *J Cereb Blood Flow Metab* 33: 1075-1082.
- 44. Alexandrov AV, Grotta JC (2002) Arterial reocclusion in stroke patients treated with intravenous tissue plasminogen activator. *Neurology* 59: 862-867.
- 45. Grotta JC, Welch KM, Fagan SC, Lu M, Frankel MR, et al. (2001) Clinical deterioration following improvement in the NINDS rt-PA Stroke Trial. *Stroke* 32: 661-668.
- 46. (1997) Intracerebral hemorrhage after intravenous t-PA therapy for ischemic stroke. The NINDS t-PA Stroke Study Group. *Stroke* 28: 2109-2118.
- 47. Rai AT, Raghuram K, Carpenter JS, Domico J, Hobbs G (2013) Pre-intervention cerebral blood volume predicts outcomes in patients undergoing endovascular therapy for acute ischemic stroke. *J Neurointerv Surg* 5 Suppl 1: i25-32.
- 48. Roth C, Papanagiotou P, Behnke S, Walter S, Haass A, et al. (2010) Stent-assisted mechanical recanalization for treatment of acute intracerebral artery occlusions. *Stroke* 41: 2559-2567.



# Figures





Figure 2 - Relative importance of predictors



Figure 3 - Comparison between predicted and observed outcome













#### Figure 5 - Error Histogram, for seven scale (left) and dichotomised mRS models (right)







**Figure 7** - Over all ROC curve and confusion matrix for the dichotomised outcome network, on the left and right respectively



Figure 8 - Work is in progress to design a system capable of proposing a dichotomised outcome for each patient with and without endovascular intervention





# Tables

 Table 1 - Demographics and gender ratio: Age and gender distribution of the patients:

Age		Gender	
Mean	65		Female = 15
Median	67.5		Male = 35
Mode	70		
Std. Deviation	14.85		
Minimum	28		
Maximum	88		

#### Table 2 - Distribution of the occluded vessels

Artony	Occlusion		
Artery	1st	2nd	
Vertebral	7	0	
Basilar	42	1	

#### Table 3 - ICH classification

Classification of	Infarctio	on Haemorrhagic Transformation		
Туре		Name	Definition	
Asymptomatic	HI-1	Haemorrhagic infarction type 1	Small petechiae along the margins of the infarct	
	HI-2	Haemorrhagic infarction type 2	More confluent petechiae within the infarcted area but without space-occupying effect	
	PH-1	Parenchymal haemorrhage type 1	Haematoma in ≤30% of the infarcted area with some slight space-occupying effect	
	PH-2	Parenchymal haemorrhage type 2	Dense haematoma in >30% of the infarcted area with substantial space-occupying effect or as any haemorrhagic lesion outside the infarcted area.	
Symptomatic		Symptomatic intracranial haemorrhage	parenchymal haemorrhage type 2 (PH-2) with neurological deficit	
Others	IVH	Interventricular Haemorrhage		
Others	SAH	Subarachnoid Haemorrhage		

#### Table 4 - Haemorrhagic Transformation

Туре	No.	
	HI-1	2
Asymptomatic	HI-2	4
Asymptomatic	PH-1	0
	PH-2	4
Symptomatic		4
Others	IVH	1
Oulers	SAH	1



#### Table 5 - 90 days mRS

Modified Rankin Score (mRS)		
Mean	3.9	
Median	5	
Mode	6	
Std. Deviation	2.4	
Minimum	0	
Maximum	6	

Table 6 - Model precision of support vector machine in prediction of the dichotomised intervention outcome

Outcome	True Good	True Bad	Class Precision
Predicted Good	14	8	63.64%
Predicted Bad	9	19	67.86%
Class Recall	60.87%	70.37%	