


## ORIGINAL ARTICLE

# Machine-learning-based outcome prediction in stroke patients with middle cerebral artery-M1 occlusions and early thrombectomy

Janne Hamann<sup>1</sup> | Lisa Herzog<sup>1,2,3</sup> | Carina Wehrli<sup>1,4</sup> | Tomas Dobrocky<sup>5</sup> |  
Andrea Bink<sup>4</sup> | Marco Piccirelli<sup>4</sup> | Leonidas Panos<sup>6</sup> | Johannes Kaesmacher<sup>5,7</sup> |  
Urs Fischer<sup>6</sup> | Christoph Stippich<sup>4</sup> | Andreas R. Luft<sup>1</sup> | Jan Gralla<sup>5</sup> | Marcel Arnold<sup>6</sup> |  
Roland Wiest<sup>5</sup> | Beate Sick<sup>2,3</sup> | Susanne Wegener<sup>1</sup> 

<sup>1</sup>Department of Neurology and Clinical Neuroscience Center, University Hospital Zurich and University of Zurich, Zurich, Switzerland

<sup>2</sup>Epidemiology, Biostatistics and Prevention Institute (EBPI), University of Zurich, Zurich, Switzerland

<sup>3</sup>Institute of Data Analysis and Process Design, ZHAW Winterthur, Winterthur, Switzerland

<sup>4</sup>Department of Neuroradiology and Clinical Neuroscience Center, University Hospital Zurich and University of Zurich, Zurich, Switzerland

<sup>5</sup>Diagnostic and Interventional Neuroradiology, University Hospital of Berne, Berne, Switzerland

<sup>6</sup>Department of Neurology, University Hospital of Berne, Berne, Switzerland

<sup>7</sup>Department of Diagnostic, Interventional and Pediatric Radiology, University Hospital of Berne, Berne, Switzerland

## Correspondence

Susanne Wegener, Department of Neurology, University Hospital Zurich, Frauenklinikstrasse 26, 8091 Zurich, Switzerland.  
Email: Susanne.Wegener@usz.ch

## Funding information

Support by the Swiss National Science Foundation (SNSF PP00P3\_170683) and Clinical Research Priority Program stroke (University of Zurich) is acknowledged. The funders have no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. All researchers are independent from funders and all authors had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

## Abstract

**Background and purpose:** Clinical outcomes vary substantially among individuals with large vessel occlusion (LVO) stroke. A small infarct core and large imaging mismatch were found to be associated with good recovery. The aim of this study was to investigate whether those imaging variables would improve individual prediction of functional outcome after early (<6 h) endovascular treatment (EVT) in LVO stroke.

**Methods:** We included 222 patients with acute ischemic stroke due to middle cerebral artery (MCA)-M1 occlusion who received EVT. As predictors, we used clinical variables and region of interest (ROI)-based magnetic resonance imaging features. We developed different machine-learning models and quantified their prediction performance according to the area under the receiver-operating characteristic curves and the Brier score.

**Results:** The rate of successful recanalization was 78%, with 54% patients having a favorable outcome (modified Rankin scale score 0–2). Small infarct core was associated with favorable functional outcome. Outcome prediction improved only slightly when imaging was added to patient variables. Age was the driving factor, with a sharp decrease in likelihood of favorable functional outcome above the age of 78 years.

**Abbreviations:** AUC, area under the curve; CI, confidence interval; EVT, endovascular treatment; IVT, intravenous thrombolysis; LVO, large vessel occlusion; MCA, middle cerebral artery; MR, magnetic resonance; MRI, magnetic resonance imaging; mRS, modified Rankin Scale; NIHSS, National Institutes of Health Stroke Scale; OR, odds ratio; PtC, patient characteristics; rCBF, regional cerebral blood flow; rCBV, regional cerebral blood volume; RF, random forest; ROC, receiver-operating characteristic; ROI, region of interest; THRIVE, Totalled Health Risks in Vascular Events; TICl, Thrombolysis In Cerebral Infarction scale; Tmax, time to maximum; VISTA, Virtual International Stroke Trials Archive.

Janne Hamann and Lisa Herzog contributed equally

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *European Journal of Neurology* published by John Wiley & Sons Ltd on behalf of European Academy of Neurology

**Conclusions:** In patients with MCA-M1 occlusion strokes referred to EVT within 6 h of symptom onset, infarct core volume was associated with outcome. However, ROI-based imaging variables led to no significant improvement in outcome prediction at an individual patient level when added to a set of clinical predictors. Our study is in concordance with current practice, where imaging mismatch or collateral readouts are not recommended as factors for excluding patients with MCA-M1 occlusion for early EVT.

**KEYWORDS**

machine learning, stroke outcome prediction

## INTRODUCTION

Ischemic stroke treatment aims at recanalization of the occluded vessel to restore blood flow as fast as possible [1]. With the introduction of stent retrievers, a breakthrough in large vessel occlusion (LVO) stroke management was achieved in 2015, when five randomized controlled trials showed greater efficacy of endovascular treatment (EVT) compared with intravenous thrombolysis (IVT) alone [2]. However, successful recanalization of LVO stroke does not always translate into good clinical recovery. Despite high recanalization rates, only approximately 50% of patients with LVO have a favorable functional outcome [3]. Currently, reliable predictors of functional outcome for patients undergoing EVT are lacking, but are needed to understand the effect of treatment on the individual patient.

Imaging mismatch with estimates of a small infarct core and significant area of hypoperfusion (at least 120% of the core) have been incorporated into previous trials as selection criteria for likely treatment success (EXTEND-IA, DEFUSE 3) [4,5]. In general, the evidence of imaging mismatch for functional outcome prediction has so far relied on association studies showing that the means of conditional outcome distributions depend on imaging mismatch variables [6]. However, a different mean does not necessarily imply a clear separation of the conditional outcome distributions, which is the prerequisite for achieving reliable outcome predictions [7]. At present, there are no prediction models integrating imaging predictors including imaging mismatch in stroke patients subjected to EVT.

Our goal was to develop a robust machine-learning model to improve functional outcome prediction in LVO stroke patients, integrating clinical and imaging variables. All patients had middle cerebral artery (MCA)-M1 occlusions and were subjected to EVT within 6 h of symptom onset; therefore, all were deemed good treatment candidates. Using this homogenous cohort, we intended to overcome interactions of occlusion site and treatment regimen in outcome prediction for LVO stroke and to reveal the most relevant factors indicating favorable functional outcome after early EVT.

## METHODS

### Patient data

We used encrypted clinical and imaging data from a cohort of ischemic stroke patients treated with EVT between January 2012 and August 2017 at the Inselspital Berne, Switzerland. The study was performed according to the ethical guidelines of the Canton of Berne, with approval from the local ethics committee (no. 231/14). Patients were included in the analysis if: (i) diagnosis of ischemic stroke due to an MCA-M1 occlusion was established by magnetic resonance (MR) angiography and MR perfusion; (ii) modified Rankin Scale (mRS) score at 3 months was documented; (iii) EVT was attempted; and (iv) perfusion and diffusion images were complete and of sufficient quality for analysis. Patients were excluded if previous territorial infarction was evident on magnetic resonance imaging (MRI) or if additional occlusions other than clot extension to the internal carotid artery or distal MCA branches were revealed by angiography. Patient characteristics that were available for analyses included demographic information, previous medication, risk factors, past ischemic events, as well as baseline stroke admission information (Table 1). Treatment decisions were made according to routine clinical guidelines of the Bernese Stroke Center. The study was reported according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for observational studies.

### Imaging data

Stroke imaging was performed using predefined stroke protocols with 1.5-T or 3-T MRI systems from one vendor (Siemens Healthineers, Erlangen, Germany). The full stroke protocol is described in Appendix S1A.

All image analysis was performed blinded to clinical outcome. The dataset used for predictive analysis was restricted to diffusion-weighted and perfusion images. Imaging data were post-processed with Olea Sphere's Acute Stroke Care Plug-in (Olea Sphere™

TABLE 1 Patient characteristics

	All (n = 222)	Favorable (n = 119)	Unfavorable (n = 103)	p
Age, years	73.54 (19.99)	68.80 (18.23)	79.20 (17.70)	<0.001
Women, n(%)	134 (60)	72 (61)	62 (60)	1.000
Risk factors, n(%)				
Diabetes mellitus	39 (18)	15 (13)	24 (23)	0.051
Atrial fibrillation	92 (42)	39 (33)	53 (51)	0.006
Hypertension	148 (67)	69 (58)	79 (77)	0.004
Dyslipidemia	133 (60)	75 (63)	58 (57)	0.410
Smoker	49 (24)	35 (31)	14 (16)	0.013
Coronary heart disease	35 (16)	19 (16)	16 (16)	1.000
Peripheral arterial disease	8 (4)	3 (3)	5 (5)	0.474
Past ischemic events	30 (14)	15 (13)	15 (15)	0.698
Previous medication, n(%)				
Antiplatelet therapy	70 (32)	30 (25)	40 (39)	0.031
Oral anticoagulation	21 (9)	10 (8)	11 (11)	0.648
Statin therapy	49 (22)	25 (21)	24 (23)	0.747
Antihypertensive therapy	130 (59)	61 (51)	69 (67)	0.020
On admission, n(%)				
Independent before stroke	184 (92)	104 (96)	80 (88)	0.032
NIHSS baseline	13 (9)	12 (7)	15 (9.5)	0.005
Systolic blood pressure, mmHg	153 (36)	147 (29)	160 (35)	0.004
Diastolic blood pressure, mmHg	82 (25)	81 (22)	84.5 (27.5)	0.585
Glucose, mmol/L	6.5 (2.02)	6.2 (1.7)	6.75 (2.5)	0.010
HbA1c, %	5.8 (0.6)	5.8 (0.5)	5.8 (0.9)	0.650
LDL, mmol/L	2.41 (1.41)	2.53 (1.34)	2.36 (1.38)	0.252
HDL, mmol/L	1.37 (0.57)	1.33 (0.56)	1.44 (0.6)	0.423
Triglycerides, mmol/L	1.32 (0.84)	1.33 (0.77)	1.27 (0.85)	0.825
C-reactive protein, mg/L	3 (5)	3 (3)	4 (8)	0.015
International Normalized Ratio	1.01 (0.07)	1.01 (0.05)	1.02 (0.07)	0.247
Infarct side (left), n(%)	99 (45)	54 (45)	45 (44)	0.892
IVT, n(%)	103 (46)	58 (49)	45 (44)	0.501
Onset to imaging, min	132 (210.5)	128 (187)	149.5 (243)	0.383
Onset to groin puncture, min	216 (231)	210.5 (190.75)	230 (271)	0.540
General anesthesia, n(%)	112 (52)	58 (51)	54 (53)	0.786
Collateralization status, n(%)				
Good	115 (53)	69 (59)	46 (46)	0.124
Moderate	78 (36)	39 (33)	39 (39)	
Poor	25 (11)	10 (8)	15 (15)	
Perfusion imaging, n(%)				
Volume core	19.90 (27.77)	16.50 (20.82)	23.83 (33.77)	0.543
Volume hypoperfusion	137.78 (72.25)	132.32 (69.94)	144.10 (74.68)	0.142
Volume of tissue at risk	117.88 (66.06)	115.82 (60.98)	120.26 (71.70)	0.692
rCBF core	11.83 (7.05)	13.04 (6.76)	10.43 (7.15)	0.002
rCBF penumbra	18.49 (7.59)	19.43 (7.93)	17.40 (7.06)	0.057
rCBV core	1.62 (0.93)	1.74 (0.90)	1.48 (0.95)	0.037
rCBV penumbra	3.11 (1.44)	3.11 (1.32)	3.11 (1.57)	0.728

(Continues)

TABLE 1 (Continued)

	All (n = 222)	Favorable (n = 119)	Unfavorable (n = 103)	p
After treatment				
TICI score, n(%)				
0	10 (5)	2 (2)	8 (8)	0.011
1	10 (5)	2 (2)	8 (8)	
2a	27 (12)	13 (11)	14 (13)	
2b	75 (33)	39 (32)	36 (35)	
3	100 (45)	63 (53)	37 (36)	
NIHSS score after 24 h	6 (10)	4 (4.5)	11 (9.75)	<0.001
ICH (SITS-MOST) , n(%)	5 (2)	1 (1)	4 (4)	0.185

Note: Data are frequencies (percentages), as indicated, for categorical variables and median (interquartile range) for continuous variables of clinical data and patient characteristics. For perfusion imaging, mean (standard deviation) of core and penumbral volumes (cc), as well as their regional blood flow (ml/100 ml/min) and regional blood volume (ml/100 ml) are shown. Differences in patients with favorable (modified Rankin scale [mRS] score 0–2) and unfavorable outcome (mRS score 3–6) were tested with Fisher's (categorical) and Wilcoxon's (continuous) test using a significance level of 5%.

Abbreviations: HbA1c, glycated hemoglobin; HDL, high-density lipoprotein; ICH, intracerebral hemorrhage; IVT, intravenous thrombolysis; LDL, low-density lipoprotein; NIHSS, National Institutes of Health Stroke Scale; rCBF, regional cerebral blood flow; rCBV, regional cerebral blood volume; TICI, Thrombolysis In Cerebral Infarction scale.

3.0.13, La Ciotat, France). Quantitative (regional cerebral blood flow [rCBF], regional cerebral blood volume [rCBV]), time to maximum [Tmax], and mean transit time) and qualitative/semi-quantitative (time to peak and temporal maximum intensity projections) maps were calculated from perfusion images. Automatic volume segmentation was used to define the core (with an apparent diffusion coefficient threshold at values  $<600 \times 10^{-6} \text{ mm}^2/\text{s}$ ) and penumbral volumes (with a predefined Tmax threshold at values  $>6 \text{ s}$ ; Figure 1a,b).

We manually drew regions of interest (ROIs) for the lentiform nucleus (MCA territory) and thalamus (mostly posterior cerebral artery territory) in one slice, where structures were best demarcated on the raw perfusion images (Figure 1c). Additionally, a total of seven slices were automatically segmented into five ROIs in each hemisphere using the Sector ROI template (Figure 1d). Here, we discarded slices that were two slices below and six slices above the level of the manual ROIs, because they were prone to artefacts. The ROIs in each slice corresponded to anterior cerebral artery territory, MCA territory (M ant, M med and M post), and posterior cerebral artery territory (Post) of the stroke (S) and contralateral hemisphere. For each perfusion map, we pooled the perfusion values in each ROI across the seven slices by calculating a weighted mean, adjusting for the number of pixels per ROI.

## Statistical and association analyses

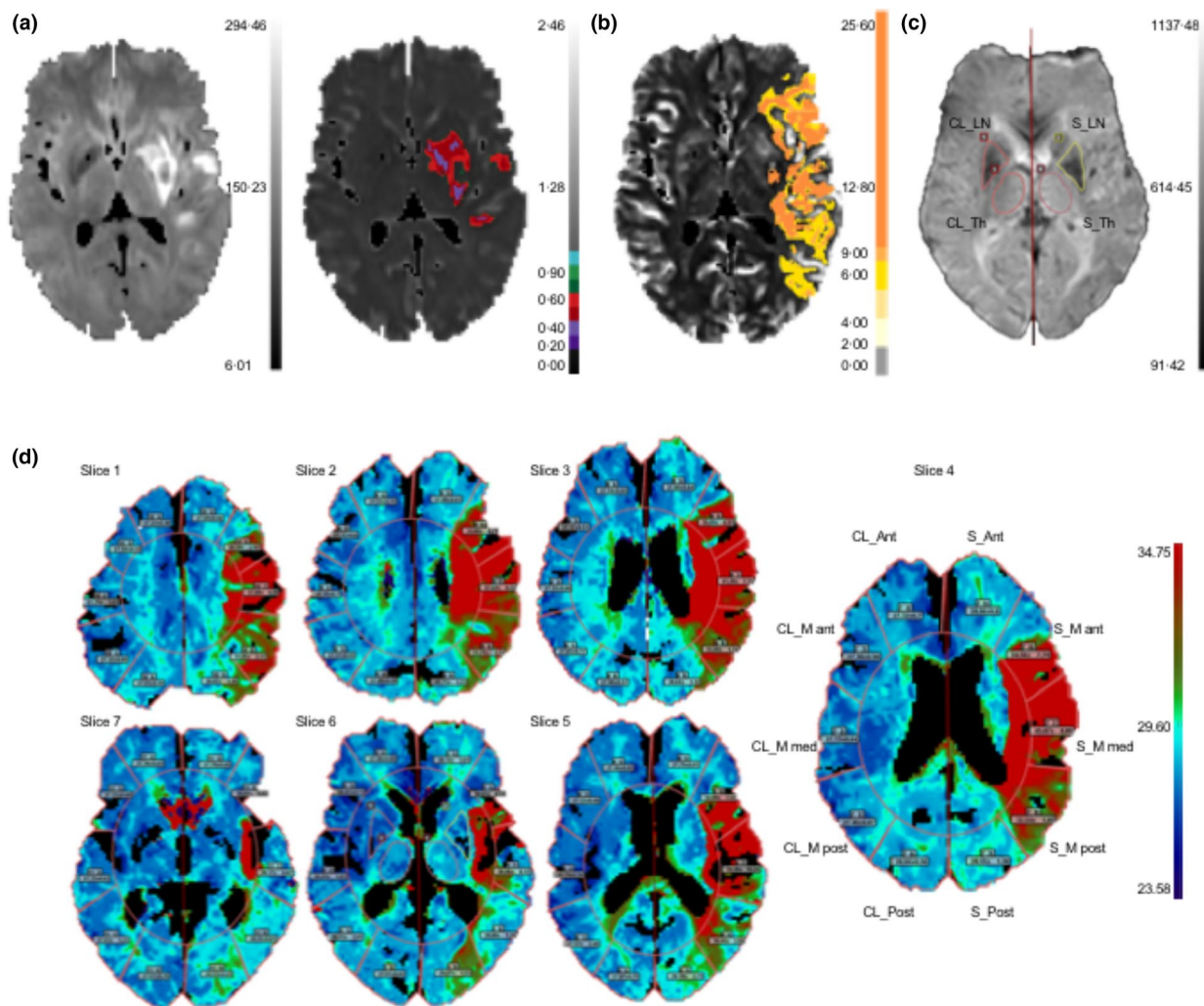
Statistical analysis was performed in open source software R (V3.5.2) [8].

To benchmark our dataset and findings, we first evaluated if infarct core and perfusion imaging were associated with functional outcome after EVT. Therefore, we followed a recent meta-analysis of 309 patients of the HERMES collaboration [6]. We estimated the effect of

infarct core and volume mismatch on favorable functional outcome (mRS scores 0–2 vs. 3–6) and functional improvement (mRS score 0–6) using a binary or an ordinal logistic regression model adjusted for age, sex, National Institutes of Health Stroke Scale (NIHSS) score at baseline, onset-to-treatment time, thrombolysis and infarct size.

## Prediction models

Next, we developed prediction models for favorable functional outcome (mRS scores 0–2 vs. 3–6) assessing the predictive power of different patient characteristics and imaging variables and a combination of both. Therefore, we defined five predictor sets. The first two sets incorporated only imaging variables from (i) the stroke hemisphere (MRI(S)) or (ii) the entire brain (MRI). The third set (iii) contained only patient characteristics (PtC). Two more predictor sets incorporated both PtC and imaging variables (iv) of the stroke hemisphere (PtC + MRI(S)) or (v) the entire brain (PtC + MRI). All predictors were taken after clinical and MRI assessment but prior to thrombectomy (see Table S1 for an overview of the predictors). To test whether prediction would improve if thrombectomy success (Thrombolysis In Cerebral Infarction [TICI] scale) and the early outcome variable 24-h NIHSS score were considered as variables, we computed all models again after adding these variables (Figure S1). The imaging variables in the respective predictor set were extracted from six perfusion maps (rCBF, rCBV, Tmax, mean transit time, time to peak, temporal maximum intensity projections) and, for each perfusion map, ROI-based features including core and parameter mismatch were summarized (Figure 1). That is, the following analyses were repeated seven times while using in each iteration the imaging variables from one of the perfusion maps as well as a combination of all imaging variables from all perfusion maps (All).



**FIGURE 1** Workflow for image-post-processing with Olea Sphere. (a) Diffusion-weighted image and automatic segmentation of core volume (right) with apparent diffusion coefficient  $< 600 \times 10^{-6} \text{ mm}^2/\text{s}$  and (b) penumbral volume with a time-to-maximum threshold of  $> 6 \text{ s}$ . (c) Manual segmentation of the lentiform nucleus (LN) and thalamus (Th) on stroke (S) and contralateral (CL) side. (d) Automatic region of interest segmentation of anterior cerebral artery territory (S\_Ant, CL\_Ant), middle cerebral artery territory (S\_M ant/med/post, CL\_M ant/med/post), and posterior cerebral artery territory (S\_Post, CL\_Post) in seven slices

Prediction models were evaluated in a fivefold cross-validation setting. The dataset was split into five test sets and data not contained in the respective test set were used for training and validation. In each of the five iterations, a random forest (RF) was learned for variable selection. Variable importance was assessed by ordering the variables with respect to the mean decrease in accuracy and by subsequently selecting the most important predictors for outcome prediction (Figure S4). For the predictor sets MRI(S) and MRI, all variables were used ( $n = 7$ ). For the remaining predictor sets, the number of predictors was reduced to the 10 most important ones. Additionally, we evaluated an “Expert” predictor set including variables which were considered most relevant by experts based on literature and clinical judgment: age; NIHSS score on admission; systolic blood pressure; risk factors (hypertension, diabetes, smoking, previous ischemic event); preceding IVT; onset to groin puncture time; collateralization status; perfusion value of the medial MCA territory (S\_M med); and volume of core and tissue at risk [9,10].

Three prediction models were used to analyse the automatically selected or the expert-selected predictors; an RF, a logistic regression (logistic regression) with L1 regularization, and a neural network [11]. The trained models were applied to predict the corresponding test set. As the baseline benchmark for our prediction models, we considered Total Health Risks in Vascular Events (THRIVE) score [12]. Detailed information about the models and the analysis procedure is provided in Appendix S1B and Figure S1, respectively.

Overall performance was evaluated by computing the area under the curve (AUC) value of the receiver-operating characteristic (ROC) curve as well as Brier scores across all test sets. Whereas AUC values indicate the model's ability to discriminate the classes, Brier scores measure the agreement between predicted and observed values. To test for significant differences between predictor sets, we performed 5% significance tests using bootstrapping to compare the AUCs of correlated ROC curves. To visualize the influence of a

selected variable on predicted outcome, we produced partial dependence plots.

## RESULTS

### Patient and imaging data

We screened 578 ischemic stroke patients with MCA-M1 occlusion and included 222 in the analysis (Figure 2). The distribution of all available patient characteristics with regard to favorable and unfavorable outcomes are shown in Table 1. The median (interquartile range) age was 73.54 (19.99), and 134 patients (60%) were women. The median (interquartile range) NIHSS score on admission was 13 (9). Successful recanalization (TICI scores 2b-3) was achieved in 175 patients (78%). Of the 222 patients, 119 (53.6%) had a favorable functional outcome after 3 months.

Outcome assessed by mRS score at 3 months was 0 in 32 (14.41%), 1 in 48 (21.62%), 2 in 39 (17.58%), 3 in 32 (14.41%), 4 in 32 (14.42%), 5 in five (2.25%), and 6 in 34 patients (15.32%). Patients with favorable outcome were younger, more often had no history of hypertension or atrial fibrillation, and had a lower NIHSS score at baseline. Onset-to-imaging or onset-to-groin-puncture times were similar between patients with favorable and unfavorable outcome, as was collateral status and frequency of IVT application.

The mean (standard deviation) core, hypoperfused area and tissue-at-risk volumes were 19.90 (27.77) cc, 137.78 (72.25) cc,

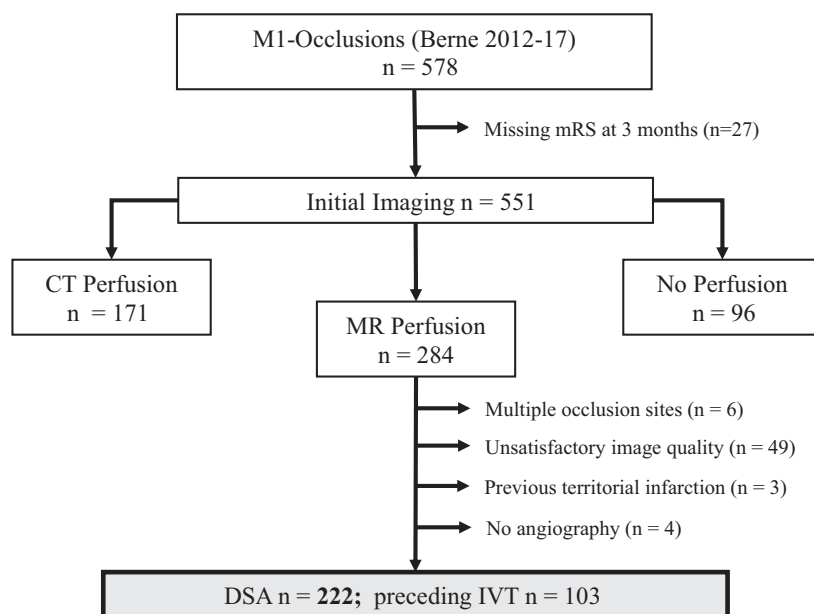
117.88 (66.06) cc, respectively. The only imaging parameters that were significantly different between the two groups were the mean relative rCBF and rCBV values in the infarct core.

### Association analyses

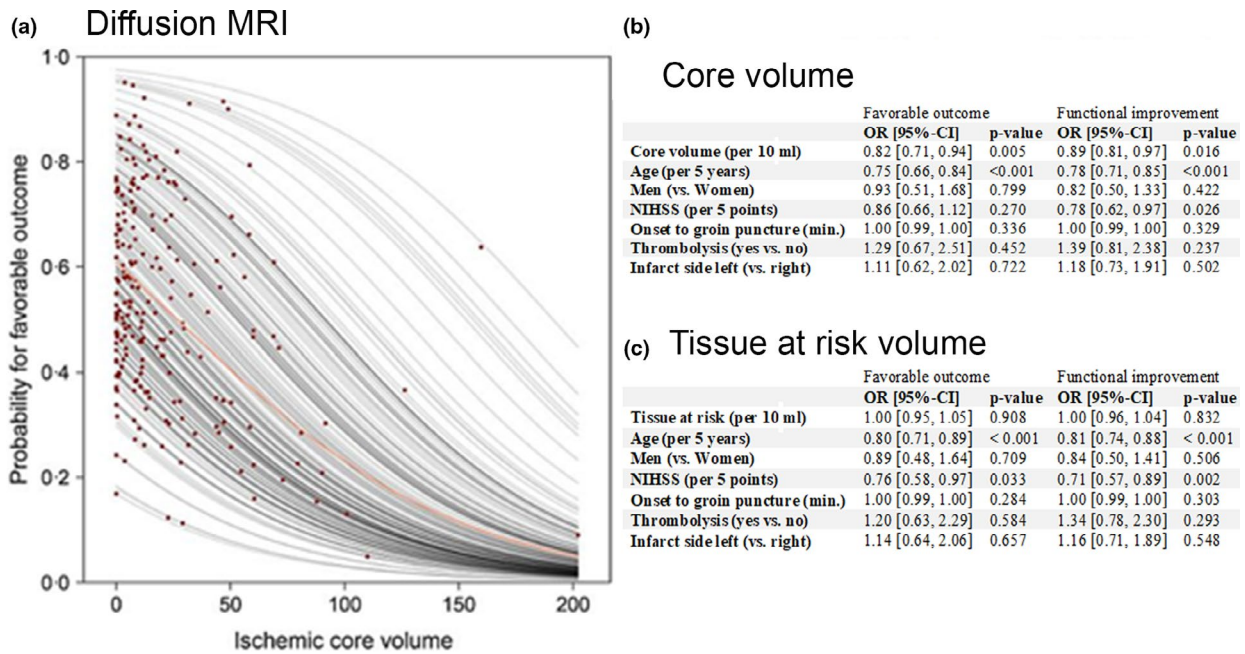
The results of the association analysis between ischemic core, volume mismatch and 90-day mRS score are summarized in Figure 3 and were in good agreement with previous studies [6]. Increasing core volume was significantly associated with a decrease in odds of favorable outcome (odds ratio of 0.82 with 95% confidence interval (CI) [0.71, 0.94] per 10 ml) and functional improvement (odds ratio of 0.89 with 95%-CI [0.81, 0.97] per 10 ml). Tissue at risk was not associated with outcome. Age was associated with a decreased chance of favorable outcome and functional improvement.

### Prediction models

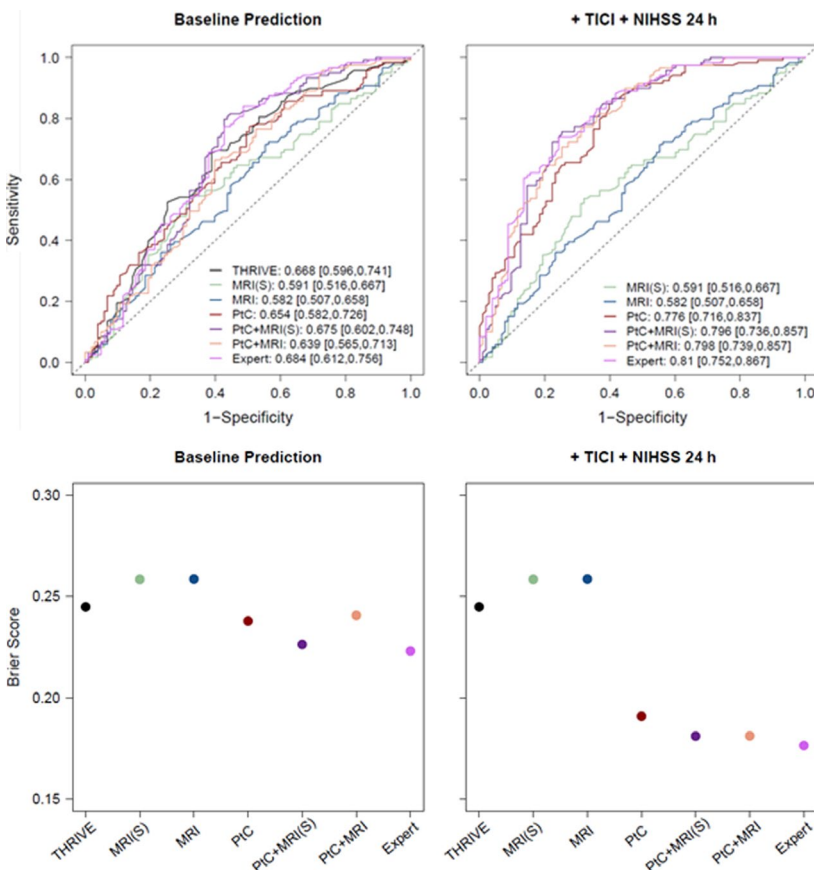
Figure 4 visualizes the model performance assessed according to ROC curves and the corresponding AUC values, together with the Brier scores for favorable functional outcome prediction based on the automatically selected variables from the predictor sets, the expert-selected predictors and the THRIVE score. For the sake of simplicity, we only show the RF as a prediction model; results obtained with the logistic regression and neural network were similar,



**FIGURE 2** Flowchart for patient selection. We screened 578 ischemic stroke patients with middle cerebral artery (MCA)-M1 occlusion and included 222 into the analysis. We excluded 27 patients due to missing modified Rankin Scale (mRS) values at the 3-month follow-up time point ( $n = 551$ ). In this study, we analysed only patients with acute stroke magnetic resonance (MR) imaging including perfusion imaging ( $n = 284$ ). A total of 62 patients were further excluded either due to insufficient perfusion data quality, for example, because of excessive head motion and other artefacts or missing MR perfusion sequence data ( $n = 49$ ), additional large vessel occlusion other than extension of the occlusion to the internal carotid artery, or more distal MCA occlusions ( $n = 6$ ), prior territorial infarction ( $n = 3$ ), or missing angiography ( $n = 4$ ). CT, computed tomography; MT, mechanical thrombectomy; DSA, digital subtraction angiography; IVT, intravenous thrombolysis.



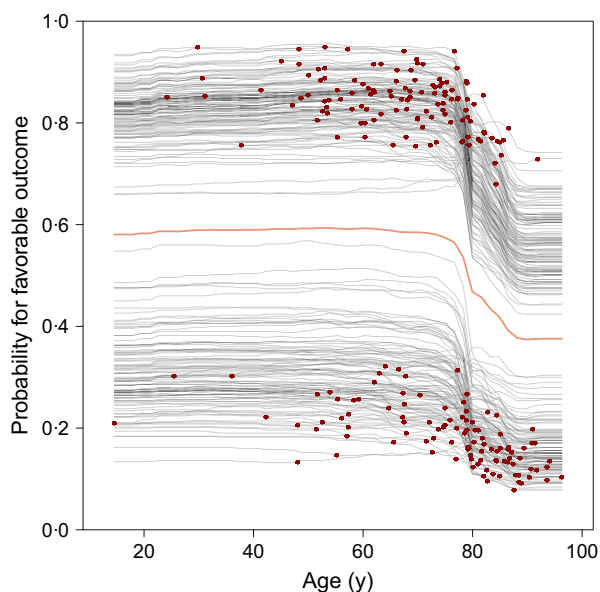
**FIGURE 3** Association of core and mismatch volumes with functional outcome. (a) Partial dependence plot derived from binary logistic regression analysis illustrating the probability for favorable outcome at 3 months (modified Rankin Scale [mRS] scores 0–2) in relation to ischemic core volume derived from diffusion-weighted imaging. Each red dot indicates an individual patient. (b and c) show the results of a binary logistic regression for favorable functional outcome (mRS scores 0–2 vs. 3–6) and an ordinal logistic regression (proportional odds model) for functional improvement (mRS scores 0–6) investigating the association between infarct core (b) and tissue at risk (c) on the outcome. CI, confidence interval; NIHSS, National Institutes of Health Stroke Scale; OR, odds ratio. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 4** Receiver-operating characteristic curves and corresponding area under the curve (AUC) values (upper row) together with Brier scores (lower row) of the random forest classifier for predictor sets with imaging variables from all perfusion maps. Five predictor sets -magnetic resonance imaging features from the stroke hemisphere (MRI(S)), from the entire brain (MRI), patient characteristics (PtC), and a combination of both (PtC + MRI(S) and PtC + MRI) - were compared with the Total Health Risks in Vascular Events (THRIVE) score and the expert predictor set (Expert). The dotted line in the upper panel indicates the random model. In case of perfect discrimination, the AUC value would reach its maximum of one and the Brier score its minimum of 0. Models are compared at time point baseline prediction (left column) and after adding Thrombolysis In Cerebral Infarction scale (TIC1) and National Institutes of Health Stroke Scale (NIHSS).

highlighting the robustness of our findings. In addition, we visualized the predictor sets with imaging variables from all perfusion maps, which yielded better results than individual perfusion maps in most settings. However, results based on the imaging variables from the other perfusion parameters were similar. A summary of all results is provided in Figures S2 and S3.

An AUC of 0.654 with 95%-CI [0.582, 0.726] and an average Brier score of 0.245 was achieved when using clinical (PtC) predictors. Imaging information alone performed relatively poorly as a predictor of functional outcome. Adding image information including infarct core and mismatch to patient variables yielded similar or only slightly better outcome predictions in terms of AUC [95%-CIs] (AUC 0.654 [0.582, 0.726] for PtC vs. AUC 0.675 [0.602, 0.748] for PtC + MRI(S), AUC 0.639 [0.565, 0.713] for PtC + MRI and AUC 0.684 [0.612, 0.756] for the expert predictor set) and Brier score (0.245 for PtC vs. 0.226 for PtC + MRI(S), 0.240 for PtC + MRI and 0.223 for the expert predictor set). Yet, no significant differences between predictor sets were observed when adding imaging variables to patient characteristics (see Appendix S1C for a comparison of predictor sets). When we compared the RF with predictor set PtC to the THRIVE score, we observed similar results. However, in terms of overall prediction performance our proposed prediction models outperformed THRIVE score (Brier scores of 0.293 for THRIVE vs. 0.245 for PtC). As expected, prediction was significantly better when recanalization success (TICI grade) and the early clinical outcome measure NIHSS score after 24 h were added (For the best model, the RF with the expert predictor set, we achieved an AUC



**FIGURE 5** Partial dependence plots of the random forest model in the predictor set PtC + MRI (with regional cerebral blood volume as imaging feature). Each curve represents the change in outcome prediction for one patient, when changing the variable of interest (x-axis) and holding the remaining covariables constant. The red dots highlight the actual observed values. The thick, coloured line marks the average dependence curve

of 0.684 with 95%-CI [0.612, 0.756] vs. an AUC of 0.81 with 95%-CI [0.752, 0.867] and Brier scores of 0.223 vs. 0.176).

In all experiments, the predicted probability of a favorable outcome was strongly dependent on age (for a summary of the most important predictors see Figure S4). In addition, the partial dependence plots allowed us to elucidate the impact of the individual variables on outcome prediction (Figure 5). We had one very consistent finding across all predictor sets and models using any of the perfusion parameters: at an age of approximately 78 years, there was a drastic drop in favorable outcomes of approximately 20%, independent of other patient characteristics.

## DISCUSSION

In stroke patients with MCA-M1 occlusion referred to thrombectomy within 6 h of symptom onset, functional outcome is variable. In our cohort, favorable outcome at 3 months (mRS score 0–2) was reached in 54% of patients and was associated with a decrease in ischemic core volume. In this homogeneous group of patients receiving early (<6 h) EVT, adding imaging variables including ischemic core volume and imaging mismatch to clinical predictors improved the prediction of a stroke patient's outcome only slightly. The driving factor for functional outcome prediction was age. Considering the wide range of outcomes after stroke even in patients with similar vascular occlusions, prognostic research is an evolving field. Several prognostic scores have been established in the general stroke population [13–18], in stroke patients treated with IVT [19,20], and in patients eligible for endovascular treatment (e.g., THRIVE, HIAT, MR PREDICTS) [9,12,21]. Some of these scores were validated and compared recently in a large dataset including hemorrhagic and ischemic stroke ( $n = 10,777$ ) from the Virtual International Stroke Trials Archive (VISTA). However, even when the highly heterogeneous VISTA patient cohort with stroke in any vascular territory independent of vessel occlusion status was used (thus enhancing accurate prediction), the resulting AUCs for outcome prediction were only between 0.6 and 0.79 [22], similar to our findings.

In our models for outcome prediction, we incorporated patient baseline variables and MRI estimates of the infarct core and hypoperfused area from the stroke hemisphere as well as whole-brain MRI readouts. AUCs for different prediction models using varying combinations of potential predictors were between 0.6 and 0.7, and similar to the THRIVE score [23]. In terms of overall prediction performance, our models outperformed THRIVE score, as indicated by a smaller Brier score. The robustness of our findings is supported by similar results for different prediction models, predictor sets and perfusion variables.

To derive our prediction models, we applied machine-learning methods, which have recently succeeded the development of novel outcome prediction tools in other fields of cardiovascular disease [24,25]. Expectations are high that these highly efficient algorithms are capable of outperforming classic score-based prediction tools in medicine [24,26]. While machine-learning was



applied to stroke outcome prediction analyses, so far, none of these studies have incorporated diffusion- or perfusion-weighted imaging variables [27–29].

Older age (>78 years) was an important determinant of outcome in our study. Although age is known to be a critical factor for recovery after stroke [30], the abrupt change in the probability of a good outcome observed in the present study has not been described before.

A limitation of this study is that it had a retrospective, single-center cohort design, with data collected over 5 years (2012–2017), a period during which thrombectomy methods have evolved. We selected only patients deemed to be good thrombectomy candidates, which resulted in a homogenous cohort of patients, but limits conclusions about treatment effects or extended time windows (>6 h) of treatment. The outcomes observed in the present cohort are very similar to data from randomized trials pooled by the HERMES collaboration [2]. However, since we analyzed data from treated patients only, we cannot make inferences about treatment effect (such as EVT vs. no EVT).

In conclusion, for patients with LVO stroke referred to thrombectomy within 6 h of symptom onset, functional disability at 3 months after EVT can be predicted moderately well using patient baseline variables. Increasing age had the strongest predictive ability for unfavorable outcome in these patients. ROI-derived imaging variables such as infarct core, area of hypoperfusion or imaging mismatch indicated an enhancement in outcome prediction in these LVO stroke patients treated early, however, no significant improvement was observed when we added the imaging information. Our results support current treatment guidelines for thrombectomy within early time windows, which maintain that large core and lack of imaging mismatch should not be used to exclude patients from treatment because their recovery may be better than anticipated. It is likely that MRI may contain additional valuable information for outcome prediction in LVO stroke patients, since the available ROI-based imaging variables might not capture all the relevant image information. Novel approaches such as deep-learning-based methods promise to outperform standard image analysis techniques in a wide range of applications [31]; therefore, we plan to perform a follow-up study using deep-learning approaches to further improve functional outcome prediction in LVO stroke patients after early EVT.

#### ACKNOWLEDGMENTS

None.

#### CONFLICT OF INTEREST

J.H., L.H., C.D.W., T.D., A.B., M.P., L.P., U.F., C.S., A.R.L., J.G., M.A., B.S. and R.W. report no conflicts of interest. J.K. has received travel grants from Pfizer and Stryker and academic grants from the SAMW/Bangerter Foundation and the Swiss Stroke Society. S.W. has received academic grants from the Swiss National Science Foundation, the UZH (Clinical Research Priority Program Stroke), the Swiss Heart foundation and the Olga Mayenfisch foundation.

#### AUTHOR CONTRIBUTIONS

J.H. prepared the first draft of the manuscript and performed patient clinical and image data analysis. L.H. did the statistical analyses, data visualization and contributed in designing the statistical analysis plan and in writing the manuscript. C.D.W. contributed to image data analysis. M.P. assisted with image processing. A.B. contributed to planning and conduction of image data analysis. T.D., L.P., J.K., U.F., J.G., M.A. and R.W. participated in patient enrolment, data collection, and data analysis. A.R.L. and C.S. contributed to data analysis and interpretation. B.S. designed the statistical plan and data analysis. S.W. conceived the study, contributed to data analysis and drafting of the manuscript. All authors critically reviewed the report, and approved the final version.

#### DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request. Code for statistical analysis and visualization is made available on Github ([https://github.com/liherz/stroke\\_outcome\\_prediction\\_ML](https://github.com/liherz/stroke_outcome_prediction_ML)).

#### ORCID

Susanne Wegener  <https://orcid.org/0000-0003-4369-7023>

#### REFERENCES

1. Phipps MS, Cronin CA. Management of acute ischemic stroke. *BMJ*. 2020;368:l6983.
2. Goyal M, Menon BK, van Zwam WH, et al. Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. *Lancet*. 2016;387:1723-1731.
3. Campbell BCV, Donnan GA, Lees KR, et al. Endovascular stent thrombectomy: the new standard of care for large vessel ischaemic stroke. *Lancet Neurol*. 2015;14:846-854.
4. Campbell BC, Mitchell PJ, Investigators E-I. Endovascular therapy for ischemic stroke. *N Engl J Med*. 2015;372:2365-2366.
5. Albers GW, Marks MP, Kemp S, et al. Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. *N Engl J Med*. 2018;378:708-718.
6. Campbell BCV, Majoie C, Albers GW, et al. Penumbra imaging and functional outcome in patients with anterior circulation ischaemic stroke treated with endovascular thrombectomy versus medical therapy: a meta-analysis of individual patient-level data. *Lancet Neurol*. 2019;18:46-55.
7. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ*. 2009;338:b375.
8. Team RC. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2014.
9. Venema E, Mulder M, Roozenbeek B, et al. Selection of patients for intra-arterial treatment for acute ischaemic stroke: Development and validation of a clinical decision tool in two randomised trials. *BMJ*. 2017;357:j1710.
10. Corso G, Bottacchi E, Tosi P, et al. Outcome predictors in first-ever ischemic stroke patients: a population-based study. *Int Sch Res Notices*. 2014;2014:904647.
11. Friedman JH, Tibshirani R, Hastie T. *The elements of statistical learning*. New York, NY: Springer, 2008.
12. Flint AC, Cullen SP, Faigeles BS, Rao VA. Predicting long-term outcome after endovascular stroke treatment: the totaled health risks in vascular events score. *AJNR Am J Neuroradiol*. 2010;31:1192-1196.

13. Ntaios G, Faouzi M, Ferrari J, Lang W, Vemmos K, Michel P. An integer-based score to predict functional outcome in acute ischemic stroke: the astral score. *Neurology*. 2012;78:1916-1922.
14. Reid JM, Dai D, Delmonte S, Counsell C, Phillips SJ, MacLeod MJ. Simple prediction scores predict good and devastating outcomes after stroke more accurately than physicians. *Age Ageing*. 2017;46:421-426.
15. O'Donnell MJ, Fang J, D'Uva C, et al. The plan score: a bedside prediction rule for death and severe disability following acute ischemic stroke. *Arch Intern Med*. 2012;172:1548-1556.
16. Saposnik G, Kapral MK, Liu Y, et al. Iscore: a risk score to predict death early after hospitalization for an acute ischemic stroke. *Circulation*. 2011;123:739-749.
17. Konig IR, Ziegler A, Bluhmki E, et al. Predicting long-term outcome after acute ischemic stroke: a simple index works in patients from controlled clinical trials. *Stroke*. 2008;39:1821-1826.
18. Kwok CS, Potter JF, Dalton G, et al. The soar stroke score predicts inpatient and 7-day mortality in acute stroke. *Stroke*. 2013;44:2010-2012.
19. Strbian D, Meretoja A, Ahlhelm FJ, et al. Predicting outcome of iv thrombolysis-treated ischemic stroke patients: the dragon score. *Neurology*. 2012;78:427-432.
20. Saposnik G, Guzik AK, Reeves M, Ovbiagele B, Johnston SC. Stroke prognostication using age and nih stroke scale: Span-100. *Neurology*. 2013;80:21-28.
21. Hallevi H, Barreto AD, Liebeskind DS, et al. Identifying patients at high risk for poor outcome after intra-arterial therapy for acute ischemic stroke. *Stroke*. 2009;40:1780-1785.
22. Quinn TJ, Singh S, Lees KR, Bath PM, Myint PK, Collaborators V. Validating and comparing stroke prognosis scales. *Neurology*. 2017;89:997-1002.
23. Kastrup A, Brunner F, Hildebrandt H, et al. Thrive score predicts clinical and radiological outcome after endovascular therapy or thrombolysis in patients with anterior circulation stroke in everyday clinical practice. *Eur J Neurol*. 2017;24:1032-1039.
24. Than MP, Pickering JW, Sandoval Y, et al. Machine learning to predict the likelihood of acute myocardial infarction. *Circulation*. 2019;140:899-909.
25. Al'Aref SJ, Anchouche K, Singh G, et al. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *Eur Heart J*. 2019;40:1975-1986.
26. Patel UK, Anwar A, Saleem S, et al. Artificial intelligence as an emerging technology in the current care of neurological disorders. *J Neurol*. 2019. <https://doi.org/10.1007/s00415-019-09518-3>.
27. van Os HJA, Ramos LA, Hilbert A, et al. Predicting outcome of endovascular treatment for acute ischemic stroke: potential value of machine learning algorithms. *Front Neurol*. 2018;9:784.
28. Asadi H, Dowling R, Yan B, Mitchell P. Machine learning for outcome prediction of acute ischemic stroke post intra-arterial therapy. *PLoS One*. 2014;9:e88225.
29. Monteiro M, Fonseca AC, Freitas AT, et al. Using machine learning to improve the prediction of functional outcome in ischemic stroke patients. *IEEE/ACM Trans Comput Biol Bioinform*. 2018;15:1953-1959.
30. Lee WI, Mitchell P, Dowling R, Yan B. Clinical factors are significant predictors of outcome post intra-arterial therapy for acute ischaemic stroke: a review. *J Neuroradiol*. 2013;40:315-325.
31. Kamnitsas K, Ledig C, Newcombe VFJ, et al. Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. *Med Image Anal*. 2017;36:61-78.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Hamann J, Herzog L, Wehrli C, et al. Machine-learning-based outcome prediction in stroke patients with middle cerebral artery-M1 occlusions and early thrombectomy. *Eur J Neurol*. 2021;28:1234-1243. <https://doi.org/10.1111/ene.14651>