Published October 28, 2024 as 10.3174/ajnr.A8547

ORIGINAL RESEARCH

Data-Driven Prognostication in Distal Medium Vessel Occlusions Using Explainable Machine Learning

Mert Karabacak*, Burak Berksu Ozkara*, Tobias D. Faizy, Trevor Hardigan, Jeremy J. Heit, Dhairya A. Lakhani, Konstantinos Margetis, J Mocco, Kambiz Nael, Max Wintermark and Vivek S. Yedavalli

ABSTRACT

BACKGROUND AND PURPOSE: Distal medium vessel occlusions (DMVOs) are estimated to cause acute ischemic stroke (AIS) in 25-40% of cases. Prognostic models can inform patient counseling and research by enabling outcome predictions. However, models designed specifically for DMVOs are lacking.

MATERIALS AND METHODS: This retrospective study developed a machine learning model to predict 90-day unfavorable outcome [defined as a modified Rankin Scale (mRS) score of 3-6] in 164 primary DMVO patients. A model developed with the TabPFN algorithm utilized selected clinical, laboratory, imaging, and treatment data with the Least Absolute Shrinkage and Selection Operator feature selection. Performance was evaluated via 5-repeat 5-fold cross-validation. Model discrimination and calibration were evaluated. SHapley Additive Explanations (SHAP) identified influential features. A web application deployed the model for individualized predictions.

RESULTS: The model achieved an area under the receiver operating characteristic curve of 0.815 (95% CI: 0.79-0.841) for predicting unfavorable outcome, demonstrating good discrimination, and a Brier score of 0.19 (95% CI: 0.177-0.202), demonstrating good calibration. SHAP analysis ranked admission National Institutes of Health Stroke Scale (NIHSS) score, premorbid mRS, type of thrombectomy, modified thrombolysis in cerebral infarction score, and history of malignancy as top predictors. The web application enables individualized prognostication.

CONCLUSIONS: Our machine learning model demonstrated good discrimination and calibration for predicting 90-day unfavorable outcomes in primary DMVO strokes. This study demonstrates the potential for personalized prognostic counseling and research to support precision medicine in stroke care and recovery.

ABBREVIATIONS: DMVO = Distal medium vessel occlusion; AIS = acute ischemic stroke; mRS = modified Rankin Scale; SHAP = SHapley Additive Explanations; NIHSS = National Institutes of Health Stroke Scale; ST = stroke thrombectomy; TRIPOD = Transparent Reporting of Multivariable Prediction Models for Individual Prognosis or Diagnosis; CT = computed tomography; CTP = CT perfusion; MRI = magnetic resonance imaging; CTA = CT angiography; DVT = deep vein thrombosis; PE = pulmonary emboli; TIA = transient ischemic attack; BMI = body mass index; ALP = alkaline phosphatase; ALT = alanine transaminase; AST = aspartate aminotransferase; NCCT-ASPECTS = Alberta Stroke Program Early CT Score; IVT = intravenous thrombolysis; mTICI = modified thrombolysis in cerebral infarction; ER = emergency room; kNN = k-nearest neighbor; LASSO = Least Absolute Shrinkage and Selection Operator; PDPs = partial dependence plots; ROC = receiver operating characteristic; PRC = precision-recall curve; AUROC = area under the ROC curve; AUPRC = area under the PRC; CI = confidence interval.

From the Department of Neurosurgery (M.K., T.H., K.M., J.M.), and Department of Diagnostic, Molecular and Interventional Radiology (B.B.O.), Mount Sinai Health System, New York, NY, USA; Neuroendovascular Division, Department of Radiology (T.D.F.), University Medical Center Münster, Münster, Germany; Departments of Radiology and Neurosurgery (J.J.H.), Stanford University, Palo Alto, CA, USA; Department of Neuroradiology (D.A.L.), Rockefeller Neuroscience Institute, West Virginia University, Morgantown, WV, USA; Department of Radiology (K.N.), University of California San Francisco, San Francisco, CA, USA; Department of Neuroradiology (M.W.), MD Anderson Cancer Center, Houston, TX, USA; Russell H. Morgan Department of Radiology and Radiological Sciences (V.S.Y.), Johns Hopkins Medicine, Baltimore, MD, USA.

Jeremy J. Heit is a consultant for Medtronic and MicroVention; and a member of advisory boards for iSchemaView, Balt, and MicroVention. J Mocco reports stock holdings in Endostream; stock holdings in Viz.ai; stock options in Neuroradial Technologies; stock options in Echovate; compensation from CVAid for consultant services; compensation from Penumbra Inc for other services; stock options in Whisper; compensation from RIST Neurovascular, Inc, for consultant services; stock options in Neurotechnology Investors, LLC, Managers; stock holdings in Q'Apel; grants from PCORI; compensation from Stryker for other services; compensation from MicroVention for other services; stock holdings in Vastrax; stock holdings in Imperative Care Inc; compensation from Synchron for consultant services; stock holdings in Blinktbi; stock holdings in Cerebrotech; stock options in Sim&Cure; stock holdings in Spinaker; stock options in Songbird; stock options in E8; compensation from Perflow for consultant services; stock options in Tulavi; stock options in Borvo; stock holdings in Viseon Inc; stock options in Radical; and stock options in Neurolutions. Kambiz Nael is a consultant for Olea Medical and Brainomix. Vivek S. Yedavalli is a consultant for iSchemaView. The disclosures presented have no relevance to the topic discussed in this paper.

Please address correspondence to Vivek S. Yedavalli, MD, MS, Russell H. Morgan Department of Radiology and Radiological Sciences, Johns Hopkins Hospital, 1800 Orleans Street, Baltimore, MD, 21287, USA; vyedava1@jhmi.edu.

*Equally contributed authors.

SUMMARY SECTION

PREVIOUS LITERATURE: Distal medium vessel occlusions (DMVOs) account for a significant portion of acute ischemic strokes. While prognostic models for large vessel occlusion stroke outcomes exist, those specifically designed for DMVOs are lacking. Previous studies

Copyright 2024 by American Society of Neuroradiology.

have explored machine learning approaches for predicting outcomes in acute ischemic stroke, but these models are not tailored to the unique characteristics of DMVOs. One prior study developed machine learning models to predict NIHSS shift in DMVO patients, but did not address longer-term functional outcomes measured by the modified Rankin Scale (mRS).

KEY FINDINGS: Our machine learning model, utilizing the TabPFN algorithm, achieved good discrimination with an area under the receiver operating characteristic curve of 0.815 (95% CI: 0.79-0.841) and calibration with a Brier score of 0.19 (95% CI: 0.177-0.202) in predicting 90-day unfavorable outcomes in DMVO patients. The model identified admission NIHSS score, premorbid mRS, thrombectomy type, mTICI score, and malignancy history as top predictors.

KNOWLEDGE ADVANCEMENT: This study presents the first prognostic machine learning model specifically designed for predicting mRS outcomes in DMVO patients. By integrating clinical, laboratory, imaging, and treatment data, our model provides a tool for personalized prognostication in this important stroke subtype, potentially informing clinical decision-making and research strategies.

INTRODUCTION

Introduction should be placed here. Please write a brief introduction to the paper that outlines the Background/Purpose in further detail. Distal medium vessel occlusions (DMVOs), also known as medium vessel occlusions, are estimated to cause acute ischemic stroke (AIS) in 25-40% of cases.[1] They are most commonly defined as occlusions in non-co-dominant M2, M3, or M4 segments of the middle cerebral artery (MCA), anterior cerebral artery, posterior cerebral artery, posterior inferior cerebellar artery, superior cerebellar artery, or anterior inferior cerebellar artery.[2] The clinical manifestations of DMVOs are heterogeneous, and the optimal imaging modality for diagnosis has yet to be defined.[3,4] DMVOs represent an emerging application for stroke thrombectomy (ST), thus it is an active area of research.[1,5] While ST has started to be employed for DMVOs in clinical practice, a better understanding of ST's efficacy and safety in DMVOs is necessary, as findings from retrospective studies and meta-analyses have been conflicting.[6]

These challenges underscore the need for accurate prognostic models to anticipate disease progression and outcomes in patients with DMVOs, regardless of whether they receive ST, medical therapy, or both. While prognostic scales, nomograms, and machine learning approaches have been extensively studied for predicting outcomes in AIS, models tailored specifically for DMVOs are lacking in the literature.[7–9] Machine learning models to predict the National Institutes of Health Stroke Scale (NIHSS) shift of patients with DMVOs were developed in a previously published study.[10] In the current study, we aimed to address the need for accurate individual-level predictions of longer-term functional outcomes in primary DMVO patients. To do so, we developed a novel model based on a modified Prior-Data Fitted Network architecture that leverages clinical, laboratory, imaging, and treatment variables to predict unfavorable outcome, defined as a modified Rankin Scale (mRS) score of 3-6—a standard measure in clinical trials.[11]

MATERIALS AND METHODS

This was a retrospective cohort study using machine learning to predict unfavorable functional outcome, defined as an mRS score of 3-6 at 90 days, in primary DMVO patients. Patients were dichotomized into two outcome groups [favorable (mRS 0-2) versus unfavorable (mRS 3-6)], and a binary classifier was developed to predict the outcome of interest.[11] The study adhered to the Transparent Reporting of Multivariable Prediction Models for Individual Prognosis or Diagnosis-Artificial Intelligence (TRIPOD+AI).[12] The data processing pipeline is depicted in Figure 1 as a summary of our methodology.

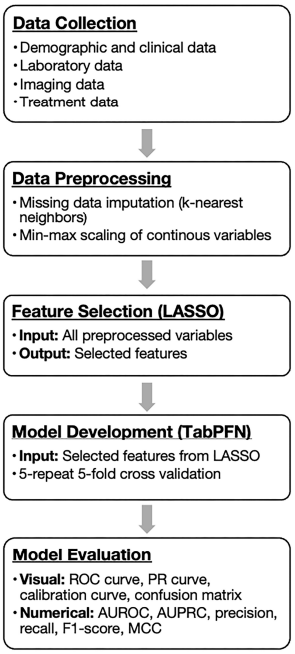


FIG 1. Data processing pipeline (LASSO: Least Absolute Shrinkage and Selection Operator, ROC: receiver operating characteristic, PR: precision-recall, AUROC: area under the receiver operating characteristic curve, AUPRC: area under the precision-recall curve, MCC: Matthews Correlation Coefficient).

Ethical Approval

This study was carried out in accordance with the Helsinki Declaration (as revised in 2013). The Johns Hopkins Hospital Institutional Review Board approved the study. The requirement for individual informed consent was waived due to the retrospective nature of the study.

Study Population

We utilized data from two comprehensive stroke centers, Johns Hopkins Hospital and Johns Hopkins Bayview Medical Center. Consecutive patients admitted between January 1, 2017, and October 16, 2022, were screened for eligibility. A DMVO was defined as an arterial occlusion involving anterior cerebral artery, M2-M4 MCA, posterior cerebral artery, posterior inferior cerebellar artery, anterior inferior cerebellar artery, or superior cerebellar artery.[2] The diagnosis of AIS was made based on clinical examination and confirmed with computed tomography (CT) or magnetic resonance imaging (MRI) of the brain. Patients were included if they met the following criteria: (I) admission within 24 hours of symptom onset; (II) age ≥18 years; (III) confirmed primary diagnosis of DMVO using CT angiography (CTA) or CT perfusion (CTP). Locations of the vessel occlusions found on either CTA or CTP were indeed further confirmed by digital subtraction angiography (DSA) in patients who underwent thrombectomy. Patients were excluded if outcome data were incomplete or if the DMVO was secondary to an iatrogenic embolus from the endovascular treatment of a different occlusion.

Demographic and Clinical Data

Demographic and clinical data were retrospectively extracted from electronic medical records. The following variables were collected: age, sex, race, smoking status, comorbidities (diabetes, dyslipidemia, hypertension, heart disease, atrial fibrillation, chronic kidney disease, sleep apnea, peripheral vascular disease), prior deep vein thrombosis (DVT) or pulmonary emboli (PE), prior stroke or transient ischemic attack (TIA), history of malignancy, antiplatelet use, body mass index (BMI), admission vitals (systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate, and oxygen saturation), admission NIHSS score, premorbid (pre-stroke) mRS score, stroke etiology, and 90-day mRS score.

Laboratory Data

Peripheral venous blood samples were collected from all patients upon arrival at the emergency department per institutional standard stroke protocol. Samples were processed and analyzed uniformly using the same methods at the clinical laboratories of the two hospitals. The following admission laboratory parameters were retrospectively retrieved: electrolytes (sodium, potassium, chloride, calcium, phosphorus, magnesium), carbon dioxide, glucose, blood urea nitrogen (BUN), creatinine, albumin, total protein, liver tests [total bilirubin, alkaline phosphatase (ALP), alanine transaminase (ALT), aspartate aminotransferase (AST)], complete blood count (hematocrit, hemoglobin, white blood cell count, platelet count), and coagulation studies (partial thromboplastin time, international normalized ratio).

Imaging Data

Both centers employed helical scanners from the Siemens SOMATOM Flash and/or Drive (Siemens Healthineers, Erlangen, Germany) to perform comprehensive baseline CT imaging. The imaging parameters applied in this study aligned with those detailed in an earlier published study.[13]

All patients' non-contrast CTs, CTAs, and CTPs were evaluated by a board-certified neuroradiologist (VSR, with nine years of experience in neuroradiology), and imaging notes were used to collect imaging data. RapidAI (iSchemaView, Menlo Park, CA) was utilized to assist in the interpretation of CTP findings. This assessment was carried out in tandem with examining all available imaging and clinical data. The same neuroradiologist also confirmed and gathered the presence of any DMVO, the baseline non-contrast computed tomography Alberta Stroke Program Early CT Score (NCCT-ASPECTS), the occluded vessel, the laterality of the occlusion, MCA dot sign, and the occurrence of hemorrhagic transformation.

Treatment Data

All thrombectomy procedures were performed by one of four credentialed interventional neuroradiologists or endovascular neurosurgeons. Device selection was at the discretion of the operator and limited to Food and Drug Administration-cleared options. The following treatment variables were collected: intravenous thrombolysis (IVT) administration, stroke thrombectomy performance, thrombectomy modality utilized, final reperfusion grade per modified thrombolysis in cerebral infarction (mTICI) score as assessed by the interventionist, and the number of thrombectomy passes. Additionally, the following time intervals (measured in minutes) were recorded: from groin puncture to first pass, groin puncture to recanalization, first pass to recanalization, last known well to emergency room (ER) arrival, symptom onset to ER arrival, ER to CT scan, last known well to CT, ER to groin puncture, ER to IVT bolus, and ER to final recanalization.

Data Preprocessing and Feature Selection

To prevent exclusion bias, imputation methods were used for missing data. Of the continuous variables, 25 had at least one missing value. After excluding one variable missing in > 25% of patients (symptom onset to arrival at the ER), the k-nearest neighbor (kNN) algorithm (k = 5) imputed missing values by leveraging data from the whole dataset.[14] The kNN approach fills in missing data using values from the 5 most similar cases. For categorical variables, 7 had missing values. No variable was excluded because no variable was missing in > 25% of patients, and the missing values were imputed using the mode.

Feature selection was performed to determine the variables most relevant for predicting outcomes from the preprocessed dataset. The Least Absolute Shrinkage and Selection Operator (LASSO) regression algorithm (alpha=0.01) was used for this purpose.[15] LASSO performs both variable selection and regularization to improve prediction accuracy. It adds a penalty proportional to the absolute size of the model coefficients. The degree of penalization is controlled by a tuning parameter lambda. As lambda increases, more coefficients are shrunk toward zero, effectively removing less predictive features.

We implemented LASSO feature selection within each cross-validation fold of which the details are given below. First, the training data in each fold was min-max scaled to normalize features. A LASSO model was then fitted to the scaled training data in each fold to select impactful features with non-zero model coefficients. The selected features were recorded per fold. Finally, features chosen in >50% of folds were selected as the final input variables in order to reduce variability.

Model Development and Evaluation

We utilized TabPFN, a modified Prior-Data Fitted Network architecture, for model development. TabPFN employs a meta-learning framework to enable adaption to new, unseen data by learning from diverse datasets.[16] Prior-data fitted networks like TabPFN are pre-trained on synthetic data to approximate Bayesian inference on real-world data.[16] The pre-training enables TabPFN to capture complex patterns in tabular data and transition smoothly to new datasets.

Model performance was evaluated using a 5-repeat 5-fold stratified cross-validation framework. In each repeat, the data was split into 5 roughly equal folds with a different random split, balancing outcome class ratios (stratification) to guarantee class balance across folds. The use of 5 repeats in this framework serves to increase the robustness of our performance estimates, reduce the impact of any particular random data split, mitigate potential overfitting to a single partition, and provide a more comprehensive assessment of model stability. This approach allows for more reliable estimation of confidence intervals for our performance metrics. Within each fold during every repeat, the initial training set (80% of data) was further segmented into a final training subset (70% of full data) and a validation subset (10% of full data). This resulted in a final 70:10:20 ratio for training to validation to hold-out testing. The validation subsets allowed for sigmoid calibration to align predicted risks with observed outcomes. Model discrimination, calibration, and accuracy were then evaluated on the held-out test folds.

The calibrated TabPFN model generated predictions and probability estimates on each test fold across the 5 cross-validation repeats. Overall performance was evaluated by aggregating results across all folds and repeats. Cross-validation enabled reliable assessment of generalizable predictive performance. To improve interpretability, SHapley Additive ExPlanations (SHAP) were utilized to determine relative feature importance. The SHAP plot displayed selected features hierarchically, with the most influential at the top. Additionally, partial dependence plots (PDPs) showed the isolated effect of individual features on predicted outcomes. PDPs illustrate the isolated effect of a single feature on the model's predicted output, revealing the extent to which individual features influence the predictions. The model code is available in the study GitHub repository (https://github.com/mertkarabacak/DMVO-mRS) for full transparency.

Model performance was evaluated graphically using a receiver operating characteristic (ROC) curve, which displays a binary classifier's ability to discriminate between positive and negative classes; a precision-recall curve (PRC) illustrating the tradeoff between precision and recall; a calibration plot for visually assessing the agreement between predicted probabilities and observed outcomes; and a confusion matrix that aggregates predictions across all folds and repeats to show the number of true positives, true negatives, false positives, and false negatives. Numerically, we computed precision, recall, F1-score, Matthews Correlation Coefficient, the area under the ROC curve (AUROC), the area under the PRC (AUPRC), and Brier score. A 95% confidence interval (CI) for each metric was calculated using a bootstrap approach with 1000 resampled datasets. This involved sampling with replacement from the original dataset to generate 1000 new samples. The CI was determined by finding the 2.5th and 97.5th percentiles of the bootstrapped metric mean values.

Web Application

We developed a web application to enable healthcare professionals and researchers to generate individualized predictions using our model. The application was deployed via Hugging Face, a platform for sharing machine learning models. Our implementation code is publicly available on the same platform for full transparency. The web application's functionality is demonstrated in Supplementary Video 1. It can be accessed at the following URL: https://huggingface.co/spaces/MSHS-Neurosurgery-Research/DMVO-mRS.

RESULTS

Initially, 212 patients who met all the inclusion criteria were identified. Forty-eight were excluded due to missing 90-day mRS data, leaving 164 patients for analysis. The group with a favorable outcome (90-day mRS 0-2) included 90 patients, while the unfavorable outcome group (90-day mRS 3-6) comprised 74 patients. The median age was 71 years. ST was performed in 43 (47.8%) and 41 (55.4%) patients in the favorable and unfavorable outcome groups, respectively. IVT was administered to 34 (37.8%) patients and 22 (29.7%) patients in favorable and unfavorable outcome groups, respectively. Online Supplemental Data summarizes key cohort characteristics, including demographics and selected variables. Supplementary Table 1 provides full details on all baseline clinical, laboratory, and imaging parameters.

Upon executing the LASSO regression algorithm, the following features (n = 16) were determined to be the most pertinent features to predict the outcome of interest and used for model development from the initial feature set (n = 64): age, current or former smoker, diabetes, hypertension, DVT or PE, history of malignancy, antiplatelet use, admission hemoglobin, admission BMI, admission NIHSS, premorbid mRS, MCA dot sign, occlusion laterality, IVT, type of thrombectomy, and mTICI.

The model achieved strong predictive performance with a precision of 0.711 (95% CI: 0.634-0.765), recall of 0.628 (95% CI: 0.553-0.765), F1-score of 0.656 (95% CI: 0.585-0.708), accuracy of 0.724 (95% CI: 0.696-0.752), MCC of 0.45 (95% CI: 0.39-0.503), and Brier

score of 0.19 (95% CI: 0.177-0.202). The AUROC was 0.815 (95% CI: 0.79-0.841), indicating good discrimination. The AUPRC of 0.808 (95% CI: 0.781-0.832) also demonstrated strong precision-recall performance (Table 1). The model ROC curve, PRC, calibration curve, and confusion matrix are displayed in Figures 2A-D, respectively. Figure 3 shows relative feature importance by SHAP values. The top five predictors were admission NIHSS score, premorbid mRS score, type of thrombectomy, mTICI score, and history of malignancy. To demonstrate the isolated impact of key variables, PDPs (Supplementary Figure 1) displayed effects for the nine most influential features.

Performance Metric	Metric Value (95% CI)
Precision	0.711 (0.634, 0.765)
Recall	0.628 (0.553, 0.694)
F1 Score	0.656 (0.585, 0.708)
Accuracy	0.724 (0.696, 0.752)
Matthew's Correlation Coefficient	0.45 (0.390, 0.503)
AUROC	0.815 (0.790, 0.841)
AUPRC	0.808 (0.781, 0.832)
Brier Score	0.19 (0.177, 0.202)

Table 1: Model performance (CI: confidence interval, AUROC: area under the receiver operating characteristic curve, AUPRC: area under the precision-recall curve).

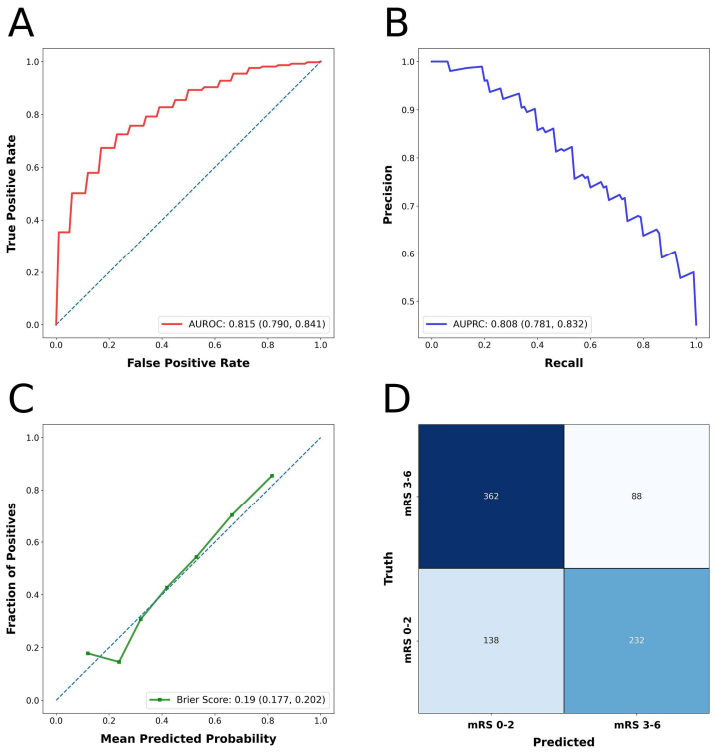


FIG 2. The model's A) receiver operating characteristic curve, B) precision-recall curve, C) calibration curve, and D) confusion matrix (AUROC: area under the receiver operating characteristics curve, AUPRC: area under the precision-recall curve, mRS: modified Rankin scale).

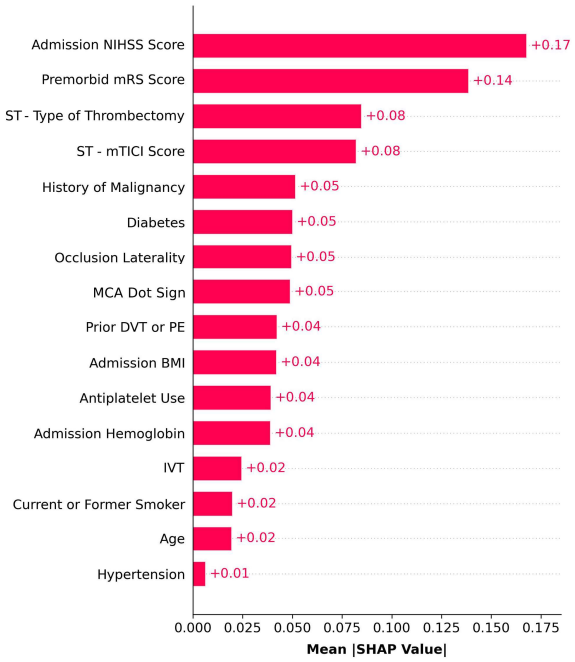


FIG 3. SHapley Additive ExPlanations (SHAP) plot of the model sorting features by their relative importance (NIHSS: National Institutes of Health Stroke Scale, mRS: modified Rankin scale, ST: Stroke thrombectomy, mTICI: modified thrombolysis in cerebral infarction, MCA: middle cerebral artery, DVT: deep vein thrombosis, PE: pulmonary embolism, BMI: body mass index, IVT: intravenous thrombolysis).

DISCUSSION

While results from ongoing trials will be critical for determining the efficacy of ST for DMVOs, ST is already being utilized clinically for many DMVO patients.[1] Reliable prognostication tools may be especially valuable given the current lack of consensus around optimal DMVO treatment approaches. These tools facilitate several applications: enabling informed discussions around likely prognosis, serving as a quality check to prompt protocol reassessment when outcomes are worse than expected, and streamlining patient stratification for research and clinical trials.

This study demonstrates the potential of machine learning models to improve prognostic predictions for DMVO patients by developing a practical tool to forecast unfavorable functional outcomes at 90 days. A unique aspect was integrating the model into a user-friendly web application to provide clinicians with personalized prognostic assessments. Our model achieved an AUROC of 0.815 for predicting unfavorable outcome (mRS score of 3-6) regardless of treatment approach. Additionally, the model demonstrated good calibration, with a Brier score of 0.19 and a near-ideal calibration curve (Figure 2C). With these promising discrimination and calibration results, our study showed machine learning could enable valuable individualized prognostication for DMVOs using key clinical and imaging parameters. Moreover, to our knowledge based on the literature review, this represents the first prognostic model specifically designed for predicting mRS outcomes in a DMVO population.

Our methodology enables precise outcome predictions for DMVO patients while also improving the interpretability of those forecasts. The SHAP feature importance plot (Figure 3) provides a global explanation of overall model behavior by revealing general patterns in how key variables relate to outcomes across the full dataset. Additionally, the SHAP plots integrated into our web application deliver local explanations that give granular insights into how unique predictions are impacted by certain variables in specific cases. This functionality allows for a personalized understanding of an individual prediction's drivers, which has not been readily accessible in most prior models. The local SHAP visualizations not only enhance interpretability but also bolster our model's credibility when combined with clinical judgment. By enabling clinicians to review the variables influencing each prediction, SHAP plots facilitate expert evaluation of model behavior and outputs. This synergy between data-driven insights and human expertise can improve the acceptance of model-based predictions, underscoring the approach's potential to meaningfully inform prognostication.

In our study, we used LASSO feature selection within each cross-validation fold, effectively reducing the number of covariates by keeping only those with high predictive power. This rigorous selection process resulted in fewer covariates, emphasizing model accuracy and stability over complexity. According to the global explanations provided by the SHAP analysis, the admission NIHSS score was the most important feature of our model. The NIHSS is a quantitative assessment of neurological deficits associated with stroke that has been shown to be reliable within and between raters and to have significant predictive power for functional outcomes.[17] Prior research indicates patients with very high or very low NIHSS scores often have predictable trajectories, with low scores suggesting probable good recovery and high scores indicating likely poor outcomes. Patients at the extreme ends are also less likely to exhibit a large observable treatment effect.[18,19] Though some prognostic scales utilize the NIHSS score as a predictor, none have been designed specifically for DMVOs.[7,8] Aligning with prior work, we found the admission NIHSS score informative for predicting 90-day functional status in our DMVO cohort. As DMVOs represent an AIS subtype, this logical association is consistent and emphasizes potential generalizability. Pre-stroke functional status, assessed by premorbid mRS score, also strongly contributed to predictions. Previous research suggests higher premorbid disability levels are associated with worse outcomes and mortality, even for patients receiving ST.[20–23] Similarly, in our cohort, 6 of 7 patients with

premorbid mRS scores >3, indicating significant prior disability, had died by the 90-day follow-up, versus only 30 deaths among 156 patients with premorbid mRS scores ≤3. Therefore, the baseline functionality level provided a useful predictive signal regarding the probability of favorable recovery in our study.

The type of ST technique utilized ranked as the third most important predictor in our model. A recent systematic review found heterogeneity amongst ongoing DMVO trials regarding the thrombectomy approach - some mandated stent retrievers while others allowed operator discretion across techniques.[5] These studies may reveal insights into how and why the specific thrombectomy modality impacts outcomes and contributes significant predictive signal, as seen in our model's feature importance ranking. Final reperfusion grade, as assessed by the mTICI score, also featured prominently as the fourth most influential factor. The mTICI scale represents an enhancement from the original TICI system designed specifically for cerebral circulation.[24] It is the primary scale recommended for evaluating reperfusion therapy in patients undergoing ST, as it is tailored for cerebral circulation, has high inter-rater reliability, and is a strong predictor of clinical outcomes.[25] Regarding the prognostic impact of the mTICI score in patients with DMVOs treated with ST, some studies have shown its positive effect on favorable outcomes, enlightening the importance of the mTICI score as a critical variable in our model.[26,27] This helps contextualize why the mTICI score would logically serve as an informative predictor. The history of malignancy and diabetes mellitus were other important features revealed by SHAP analysis. It has been shown that systemic malignancy is linked to an increased risk of ischemic stroke, and patients with a history of cancer are more likely to have recurrent strokes and die from cardiovascular disease.[28,29] Furthermore, diabetes has long been recognized as a risk factor for stroke mortality.[30] These provide further context as to why a history of malignancy and diabetes would logically be an informative predictor.

Our study has several limitations that should be considered when interpreting the results. The retrospective design and modest cohort size (n=164) inherently limit the generalizability of our findings. While we rigorously evaluated model performance using a 5-repeat, 5-fold stratified cross-validation methodology, this approach does not replace the value of external validation in entirely separate populations. The absence of an independent test set further limits our ability to fully assess the model's generalizability. Future studies should aim for larger, multi-center, prospective cohorts with independent test sets for more robust validation. Data was drawn from two high-volume comprehensive stroke centers, which may not represent the full spectrum of clinical settings where DMVOs are treated. Outcomes may depend substantially on facility volume, operator experience, and other institutional factors. The evaluation of imaging data by a single neuroradiologist and the involvement of only four neurointerventionalists introduce potential bias, particularly given the importance of thrombectomy technique as a predictive factor. We acknowledge limitations in our input variables. The NIHSS may not fully capture stroke symptoms, especially in posterior circulation strokes. The mTICI score, designed for anterior circulation large vessel occlusions, may not be optimal for all vessels involved in DMVOs. The mRS, our outcome measure, has known variability within each category. The use of kNN for imputing missing data has limitations in dealing with non-random missingness, potentially introducing bias. Additionally, a substantial number of patients were excluded due to missing 90-day mRS scores, which could affect the model's generalizability. Lastly, while certain variables demonstrate prognostic relationships with outcomes, these associations should not be inferred as causal without further analysis. The model predicts outcomes based on patterns and correlations without determining causality. Model outputs should be viewed as prognostic forecasts rather than endorsements of causal mechanisms or treatment effectiveness.

Future studies should aim to address these limitations through larger, prospective, multi-center designs with more diverse patient populations and clinical settings. Incorporating more granular details, including but not limited to discharge disposition, discharge mRS, use of other functional neurorehabilitation scales, and readmissions within 30 days, may provide even more predictive models in the future. No clinical decisions or interventions should be actively guided by this model without further validation, as its current function is exclusively prognostic rather than prescriptive.

CONCLUSIONS

Identifying outcome risks and engaging in shared decision-making is crucial for patient care in DMVOs. Our machine learning approach, utilizing the TabPFN algorithm, achieved good discrimination and calibration in predicting 90-day functional outcomes in DMVO patients. By translating predictive modeling into accessible and interpretable risk estimates, our methodology exemplifies how precision medicine tools can empower granular prognosis for stroke variants like DMVOs. External validation with larger multicenter datasets is needed to confirm generalizability before integrating such tools into routine practice.

ACKNOWLEDGMENTS

None.

REFERENCES

1. Kobeissi H, Bilgin C, Ghozy S, et al. A review of acute ischemic stroke caused by distal, medium vessel occlusions. *Interv Neuroradiol* https://doi.org/10.1177/15910199231197616.
2. Saver JL, Chapot R, Agid R, et al. Thrombectomy for Distal, Medium Vessel Occlusions: A Consensus Statement on Present Knowledge and Promising Directions. *Stroke* 2020;51:2872–84.
3. Sousa JA, Sondermann A, Bernardo-Castro S, et al. CTA and CTP for Detecting Distal Medium Vessel Occlusions: A Systematic Review and Meta-analysis. *AJNR Am J Neuroradiol* 2023;45:51–6.
4. Sousa JA, Sondermann A, Bernardo-Castro S, et al. Diagnostic accuracy of CT angiography and CT perfusion imaging for detecting distal medium vessel occlusions: Protocol for a systematic review and meta-analysis. *PLoS One* 2023;18:e0284116.
5. Bilgin C, Bolsegui ML, Ghozy S, et al. Common design and data elements reported in active mechanical thrombectomy trials focusing on distal medium vessel occlusions and minor strokes: a systematic review. *J Neurointerv Surg* https://doi.org/10.1136/jnis-2023-021073.
6. Nedelcu S, Gulati A, Henninger N. Endovascular therapy versus medical management for mild strokes due to medium and distal vessel occlusions. *Interv Neuroradiol* https://doi.org/10.1177/15910199231216510.
7. Matsumoto K, Nohara Y, Soejima H, et al. Stroke Prognostic Scores and Data-Driven Prediction of Clinical Outcomes After Acute Ischemic Stroke.

Stroke 2020;51:1477–83.
8. Drozdowska BA, Singh S, Quinn TJ. Thinking About the Future: A Review of Prognostic Scales Used in Acute Stroke. *Front Neurol* 2019;10:274.
9. Mainali S, Darsie ME, Smetana KS. Machine Learning in Action: Stroke Diagnosis and Outcome Prediction. *Front Neurol* 2021;12:734345.
10. Ozkara BB, Karabacak M, Kotha A, et al. Development of machine learning models for predicting outcome in patients with distal medium vessel occlusions: a retrospective study. *Quant Imaging Med Surg* 2023;13:5815–30.
11. Weisscher N, Vermeulen M, Roos YB, et al. What should be defined as good outcome in stroke trials; a modified Rankin score of 0-1 or 0-2? *J Neurol* 2008;255:867–74.
12. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* https://doi.org/10.1136/bmj-2023-078378.
13. Ozkara BB, Karabacak M, Hamam O, et al. Prediction of Functional Outcome in Stroke Patients with Proximal Middle Cerebral Artery Occlusions Using Machine Learning Models. *J Clin Med* 2023;12:839.
14. Beretta L, Santaniello A. Nearest neighbor imputation algorithms: a critical evaluation. *BMC Medical Informatics and Decision Making* 2016;16:74.
15. Tibshirani R. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society Series B (Methodological)* 1996;58:267–88.
16. Hollmann N, Müller S, Eggensperger K, et al. TabPFN: A Transformer That Solves Small Tabular Classification Problems in a Second. https://doi.org/10.48550/ARXIV.2207.01848.
17. Saber H, Saver JL. Distributional Validity and Prognostic Power of the National Institutes of Health Stroke Scale in US Administrative Claims Data. *JAMA Neurol* 2020;77:606–12.
18. Saver JL. Optimal Endpoints for Acute Stroke Therapy Trials: Best Ways to Measure Treatment Effects of Drugs and Devices. *Stroke* 2011;42:2356–62.
19. Ntaios G, Faouzi M, Ferrari J, et al. An integer-based score to predict functional outcome in acute ischemic stroke: the ASTRAL score. *Neurology* 2012;78:1916–22.
20. de Havenon A, Castonguay A, Nogueira R, et al. Prestroke Disability and Outcome After Thrombectomy for Emergent Anterior Circulation Large Vessel Occlusion Stroke. *Neurology* 2021;97:e1914–9.
21. Kwok CS, Clark A, Ford GA, et al. Association between prestroke disability and inpatient mortality and length of acute hospital stay after acute stroke. *J Am Geriatr Soc* 2012;60:726–32.
22. Leker RR, Gavriliuc P, Yaghmour NE, et al. Increased Risk for Unfavorable Outcome in Patients with Pre-Existing Disability Undergoing Endovascular Therapy. *J Stroke Cerebrovasc Dis* 2018;27:92–6.
23. Salwi S, Cutting S, Salgado AD, et al. Mechanical Thrombectomy in Patients With Ischemic Stroke With Prestroke Disability. *Stroke* 2020;51:1539–45.
24. Zaidat OO, Yoo AJ, Khatri P, et al. Recommendations on Angiographic Revascularization Grading Standards for Acute Ischemic Stroke. *Stroke* 2013;44:2650–63.
25. Chamorro Á, Blasco J, López A, et al. Complete reperfusion is required for maximal benefits of mechanical thrombectomy in stroke patients. *Sci Rep* 2017;7:11636.
26. Hulscher F, Farouki Y, Mine B, et al. Predictors of Good Clinical Outcome after Thrombectomy for Distal Medium Vessel Occlusions. *World Neurosurg* 2022;160:e566–72.
27. Abdelrady M, Derraz I, Dargazanli C, et al. Complete recanalization predicts favorable outcome in patients with distal M2-M3 middle cerebral artery occlusions following endovascular thrombectomy. *J Neuroradiol* 2023;50:230–6.
28. Lau K-K, Wong Y-K, Teo K-C, et al. Stroke Patients with a Past History of Cancer Are at Increased Risk of Recurrent Stroke and Cardiovascular Mortality. *PLoS One* 2014;9:e88283.
29. Yoo J, Nam HS, Kim YD, et al. Short-Term Outcome of Ischemic Stroke Patients With Systemic Malignancy. *Stroke* 2019;50:507–11.
30. Tuomilehto J, Rastenyte D, Jousilahti P, et al. Diabetes mellitus as a risk factor for death from stroke. Prospective study of the middle-aged Finnish population. *Stroke* 1996;27:210–5.

SUPPLEMENTAL FILES

Variable		n (%) or median (Q1-Q3)	Missing Data (%)
Age		71 (62.8-80)	-
Sex	**Female**	91 (55.5%)	-
	Male	73 (44.5%)	
Race	**White**	87 (53.1%)	-
	Asian	3 (1.8%)	
	Black	72 (43.9%)	
	Other	2 (1.2%)	
Current or Former Smoker	**No**	84 (51.2%)	-
	Yes	80 (48.8%)	
Diabetes	**No**	125 (76.2%)	-
	Yes	39 (23.8%)	
Hypertension	**No**	26 (15.9%)	-
	Yes	138 (84.2%)	
Heart Disease	**No**	84 (51.2%)	-

	Yes	80 (48.8%)	
History of Malignancy	**No**	134 (81.7%)	-
	Yes	30 (18.3%)	
Antiplatelet Use	**No**	109 (66.5%)	-
	Yes	55 (33.5%)	
DVT or PE	**No**	141 (86%)	-
	Yes	23 (14%)	
Admission BMI		27.4 (23.1-32.1)	3 (1.8%)
Admission Hemoglobin		13 (11.9-14.3)	-
Admission NIHSS Score	**0**	5 (3.1%)	1 (0.6%)
	1-4	37 (22.7%)	
	5-15	74 (45.4%)	
	16-20	29 (17.8%)	
	21-42	18 (11%)	
Premorbid mRS Score	**0**	94 (57.7%)	1 (0.6%)
	1	33 (20.3%)	
	2	13 (8%)	
	3	16 (9.8%)	
	4	4 (2.5%)	
	5	3 (1.8%)	
Occlusion Site	**MCA**	142 (18.6%)	-
	PCA	17 (10.4%)	
	ACA	5 (3.1%)	
Occlusion Laterality	**Left**	99 (60.4%)	-
	Right	65 (39.6%)	
MCA Dot Sign	**No**	108 (67.5%)	4 (2.4%)
	Yes	52 (32.5%)	
Intravenous Thrombolysis	**No**	108 (65.9%)	-
	Yes	56 (34.2%)	
ST - Type of Thrombectomy	**ST not attempted**	80 (51%)	7 (4.3%)
	Direct aspiration	38 (24.2%)	
	Stent retriever	12 (7.6%)	
	Combined	27 (17.2%)	
ST - mTICI	**ST not attempted**	80 (50%)	4 (2.4%)
	0	8 (5%)	
	1	3 (1.9%)	
	2a	4 (2.5%)	
	2b	15 (9.4%)	
	2c	7 (4.4%)	
	3	43 (26.9%)	
90-Day mRS Score	**0**	31 (18.9%)	-
	1	37 (22.6%)	
	2	22 (13.4%)	
	3	17 (10.4%)	
	4	17 (10.4%)	

	5	4 (2.4%)
	6	36 (22%)

Online Supplemental Data: Patient characteristics (n: number, Q1: quartile 1, Q2: quartile 3, BMI: body mass index, NIHSS: National Institutes of Health Stroke Scale, mRS: modified Rankin scale, MCA: middle cerebral artery, PCA: posterior cerebral artery, ACA: anterior cerebral artery, ST: stroke thrombectomy, mTICI: modified thrombolysis in cerebral infarction).

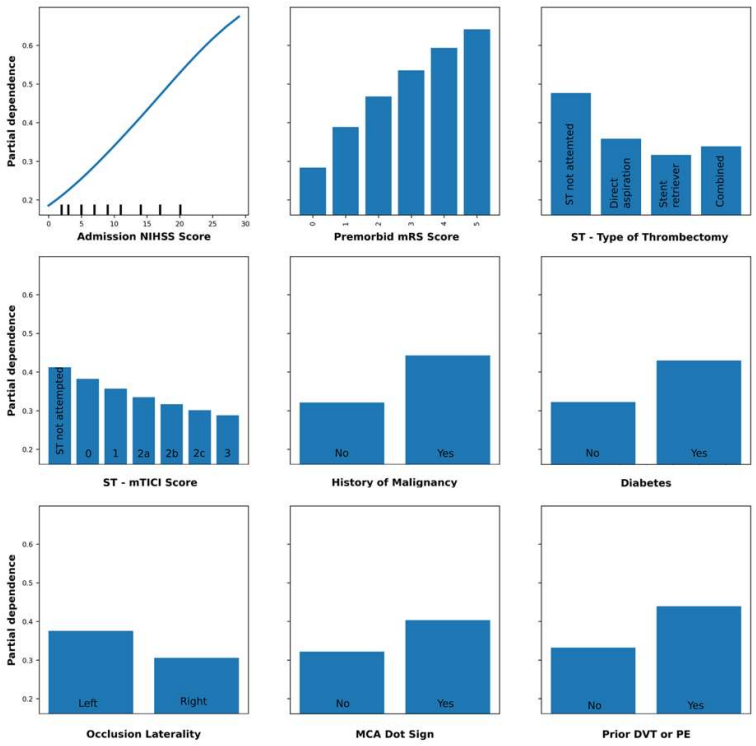
Variable		n (%) or median (Q1-Q3)	Missing Data (%)
Age		71 (62.8-80)	-
Sex	**Female**	91 (55.5%)	-
	Male	73 (44.5%)	
Race	**White**	87 (53.1%)	-
	Asian	3 (1.8%)	
	Black	72 (43.9%)	
	Other	2 (1.2%)	
Current or Former Smoker	**No**	84 (51.2%)	-
	Yes	80 (48.8%)	
Diabetes	**No**	125 (76.2%)	-
	Yes	39 (23.8%)	
Dyslipidemia	**No**	64 (39%)	-
	Yes	100 (61%)	
Hypertension	**No**	26 (15.9%)	-
	Yes	138 (84.2%)	
Heart Disease	**No**	84 (51.2%)	-
	Yes	80 (48.8%)	
Atrial Fibrillation	**No**	105 (64%)	-
	Yes	59 (36%)	
Peripheral Vascular Disease	**No**	152 (92.7%)	-
	Yes	12 (7.3%)	
Prior DVT or PE	**No**	141 (86%)	-
	Yes	23 (14%)	
Prior Stroke or TIA	**No**	136 (82.9%)	-
	Yes	28 (17.1%)	
Chronic Kidney Disease	**No**	137 (83.5%)	-
	Yes	27 (16.5%)	
History of Malignancy	**No**	134 (81.7%)	-
	Yes	30 (18.3%)	
Antiplatelet Use	**No**	109 (66.5%)	-
	Yes	55 (33.5%)	
Admission BMI		27.4 (23.1-32.1)	3 (1.8%)
Admission SBP		150 (133-174)	-
Admission DBP		83 (73.5-97.5)	1 (0.6%)
Admission HR		83.5 (71.8-98)	-

Admission RR		18 (16-19.8)	2 (1.2%)
Admission Saturation		98 (96-100)	-
Admission Carbon Dioxide		24 (22-26)	3 (1.8%)
Admission Sodium		139 (137-141)	-
Admission Potassium		4.1 (3.7-4.4)	2 (1.2%)
Admission Chloride		103 (100-107)	2 (1.2%)
Admission Calcium		9.1 (8.7-9.5)	-
Admission Magnesium		1.9 (1.7-2.1)	37 (22.6%)
Admission Phosphorous		3.2 (2.9-3.9)	46 (28.0%)
Admission Glucose		117.5 (101.8-137.3)	-
Admission BUN		17 (14-23)	-
Admission Creatinine		1 (0.9-1.2)	-
Admission Albumin		3.9 (3.5-4.3)	7 (4.3%)
Admission Total Protein		6.9 (6.4-7.3)	7 (4.3%)
Admission Total Bilirubin		0.5 (0.4-0.7)	7 (4.3%)
Admission ALP		80 (65-102)	7 (4.3%)
Admission ALT		18 (14-27)	7 (4.3%)
Admission AST		21.5 (17-30)	18 (11.0%)
Admission Hematocrit		40.5 (37-43.6)	-
Admission Hemoglobin		13 (11.9-14.3)	-
Admission WBC Count		8.1 (6.5-10.2)	-
Admission Platelet Count		233.5 (186-292.8)	-
Admission INR		1.1 (1-1.1)	17 (10.4%)
Admission PTT		25.4 (23.6-27)	26 (15.9%)
Admission NIHSS Score	**0**	5 (3.1%)	1 (0.6%)
	1-4	37 (22.7%)	
	5-15	74 (45.4%)	
	16-20	29 (17.8%)	
	21-42	18 (11%)	
Premorbid mRS Score	**0**	94 (57.7%)	1 (0.6%)
	1	33 (20.3%)	
	2	13 (8%)	
	3	16 (9.8%)	
	4	4 (2.5%)	
	5	3 (1.8%)	
Stroke Etiology	**Cardioembolism**	80 (48.8%)	-
	Large artery atherosclerosis	18 (11%)	
	Small-vessel occlusion	2 (1.2%)	
	Stroke of other determined etiology	5 (3.1%)	
	Stroke of undetermined etiology	59 (36%)	
Occlusion Site	**MCA**	142 (86.6%)	-

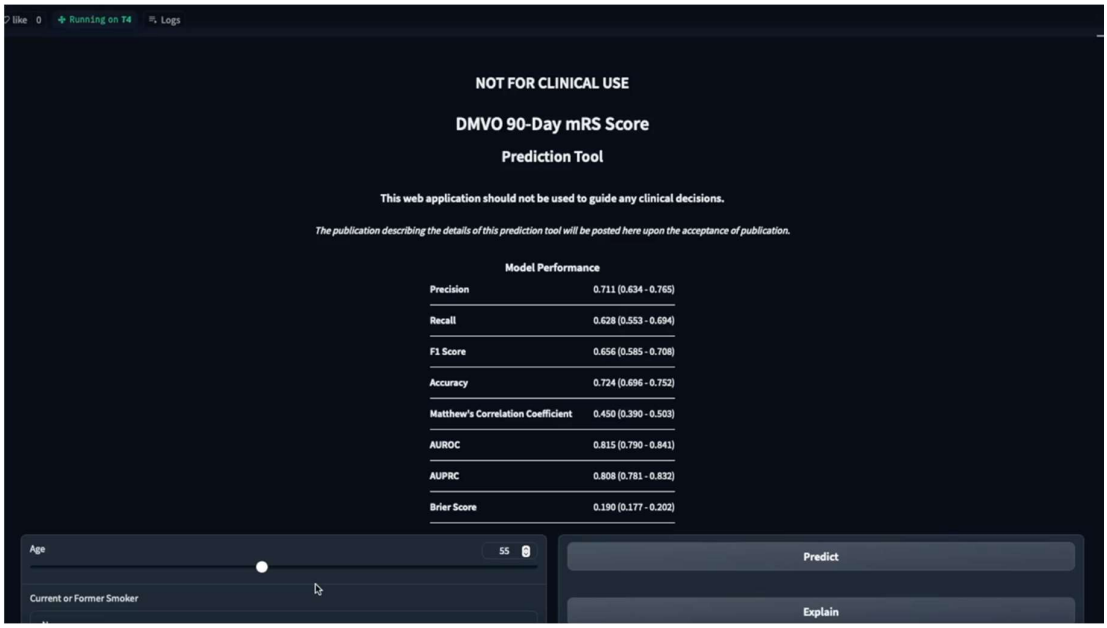
	PCA	17 (10.4%)	
	ACA	5 (3.1%)	
Occlusion Laterality	**Left**	99 (60.4%)	-
	Right	65 (39.6%)	
NCCT-ASPECTS	**10**	90 (56.3%)	4 (2.4%)
	9	30 (18.8%)	
	≤ 8	40 (25%)	
MCA Dot Sign	**No**	108 (67.5%)	4 (2.4%)
	Yes	52 (32.5%)	
Intravenous Thrombolysis	**No**	108 (65.9%)	-
	Yes	56 (34.2%)	
ST - Type of Thrombectomy	**ST not attempted**	80 (51%)	7 (4.3%)
	Direct aspiration	38 (24.2%)	
	Stent retriever	12 (7.6%)	
	Combined	27 (17.2%)	
ST – mTICI Score	**ST not attempted**	80 (50%)	4 (2.4%)
	0	8 (5%)	
	1	3 (1.9%)	
	2a	4 (2.5%)	
	2b	15 (9.4%)	
	2c	7 (4.4%)	
	3	43 (26.9%)	
ST - Number of Passes	**ST not attempted**	80 (49.7%)	3 (1.8%)
	1	52 (31.1%)	
	2	13 (8.1%)	
	3	11 (6.8%)	
	≥ 4	5 (3.1%)	
Symptom Onset to Door		71 (47-201)	93 (56.7%)
Last Known Well to CT		290 (109-782)	7 (4.3%)
Last Known Well to Door		262.5 (75.8-755)	8 (4.9%)
Door to CT		28 (17.8-45.3)	4 (2.4%)
Door to Groin Puncture		848 (152-854)	1 (0.6%)
Door to Needle Time		180 (93.8-180)	-
Door to Recanalization		2444 (367.3-3035)	2 (1.2%)
First Pass to Recanalization		92 (10.3-92)	26 (15.9%)
Groin Puncture to First Pass Time		123 (27-123)	10 (6.1%)
Groin Puncture to Recanalization		128 (37-128)	15 (9.1%)
90-Day mRS Score	**0**	31 (18.9%)	-

	1	37 (22.6%)
	2	22 (13.4%)
	3	17 (10.4%)
	4	17 (10.4%)
	5	4 (2.4%)
	6	36 (22%)

Supplemental Table 1: Detailed patient characteristics (n: number, Q1: quartile 1, Q2: quartile 3, DVT: deep vein thrombosis, PE: pulmonary embolism, TIA: transient ischemic attack, BMI: body mass index, SBP: systolic blood pressure, DBP: diastolic blood pressure, HR: heart rate, RR: respiratory rate, BUN: blood urea nitrogen, ALP: alkaline phosphatase, ALT: alanine transaminase, AST: aspartate aminotransferase, WBC: white blood cell, INR: international normalized ratio, PTT: partial thromboplastin time, NIHSS: National Institutes of Health Stroke Scale, mRS: modified Rankin scale, MCA: middle cerebral artery, PCA: posterior cerebral artery, ACA: anterior cerebral artery, NCCT-ASPECTS: non-contrast computed tomography Alberta Stroke Program Early Computed Tomography Score, CT: computed tomography, ST: mechanical thrombectomy, mTICI: modified thrombolysis in cerebral infarction).



SUP FIG 1. The partial dependency plot (PDP) for the 9 most important features of the model (NIHSS: National Institutes of Health Stroke Scale, mRS: modified Rankin scale, ST: stroke thrombectomy, mTICI: modified thrombolysis in cerebral infarction, MCA: middle cerebral artery, DVT: deep vein thrombosis, PE: pulmonary embolism).



SUP VIDEO 1. Demonstration of the web application.