Multivariable Prediction Model for Futile Recanalization Therapies in Patients With Acute Ischemic Stroke

Author(s):
Thomas Meinel; Christine Lerch; Urs Fischer; Morin Beyeler, MD; Adnan Mujanovic, MD; Christoph Kurmann, MD; Bernhard Siepen, MD; Adrian Scutelnic, MD; Madlaine Muller, MD; Martina Goeldlin, MD; Nebiyat Filate Belachew, MD; Tomas Dobrocky, MD; Jan Gralla, MD; David Seiffge; Simon Jung, MD; Marcel Arnold, MD; Roland Wiest, MD; Raphael Meier, PhD; Johannes Kaesmacher, MD

Corresponding Author:
Thomas Meinel, thomas.meinel@students.unibe.ch

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND), which permits downloading and sharing the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Neurology® Published Ahead of Print articles have been peer reviewed and accepted for publication. This manuscript will be published in its final form after copyediting, page composition, and review of proofs. Errors that could affect the content may be corrected during these processes.
Affiliation Information for All Authors:

1. Department of Neurology, University Hospital Bern, Inselspital, University of Bern, Bern, Switzerland;
2. Department of Neurology and Stroke Center, University Hospital Basel and University of Basel, Switzerland;
3. University Institute of Diagnostic and Interventional Neuroradiology, University Hospital Bern, Inselspital, University of Bern, Bern, Switzerland;
4. Support Center for Advanced Neuroimaging, University Hospital Bern, Inselspital, University of Bern, Bern, Switzerland;
5. Department of Diagnostic, Paediatric and Interventional Radiology, University Hospital Bern, Inselspital, University of Bern, Bern, Switzerland

Equal Author Contribution:
co-senior authors: Raphael Meier and Johannes Kaesmacher

Contributions:
Thomas Meinel: Drafting/revision of the manuscript for content, including medical writing for content; Study concept or design; Analysis or interpretation of data
Christine Lerch: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design
Urs Fischer: Drafting/revision of the manuscript for content, including medical writing for content; Study concept or design
Morin Beyeler: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data
Adnan Mujanovic: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data
Christoph Kurmann: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data
Bernhard Siepen: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data
Adrian Scutelnic: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data
Madlaine Müller: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data
Martina Goeldlin: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data
Nebiyat Filate Belachew: Drafting/revision of the manuscript for content, including medical writing for content; Analysis or interpretation of data
Tomas Dobrocky: Drafting/revision of the manuscript for content, including medical writing for content; Study concept or design
Jan Gralla: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design
David Seiffge: Drafting/revision of the manuscript for content, including medical writing for content; Study concept or
Figure Count:
3

Table Count:
2

Search Terms:

Acknowledgment:

Study Funding:
This study was supported by funding received from Swiss National Science Foundation (grant no. 320030L_170060 STRAY-CATS), the Swiss Heart Foundation (grant no. FF17033 & FF18059) and the University of Bern (A digital reference network platform for clinical and experimental neuroscience – deep phenotyping and data integration).
Abstract

**Background and Objectives:** Very poor outcome despite intravenous thrombolysis (IVT) and mechanical thrombectomy (MT) occurs in about 1 of 4 patients with ischemic stroke and is associated with a high logistic and economic burden. We aimed to develop and validate a multivariable prognostic model to identify futile recanalization therapies (FRT) in patients undergoing those therapies.

**Materials and Methods:** Patients from a prospectively collected observational registry of a single academic stroke center treated with MT and/or IVT were included. The dataset was split into a training (N=1808, 80%) and internal validation (N=453, 20%) cohort. We used gradient boosted decision tree machine-learning models after k-NN imputation of 32 variables available at admission to predict FRT defined as modified Rankin-Scale (mRS) 5-6 at 3 months. We report feature importance, ability for discrimination, calibration and decision curve analysis.

**Results:** 2261 patients with a median (IQR) age 75 years (64-83), 46% female, median NIHSS 9 (4-17), 34% IVT alone, 41% MT alone, 25% bridging were included. Overall 539 (24%) had FRT, more often in MT alone (34%) as compared to IVT alone (11%). Feature importance identified clinical variables (stroke severity, age, active cancer, prestroke disability), laboratory values (glucose, CRP, creatinine), imaging biomarkers (white matter hyperintensities) and onset-to-admission time as the
most important predictors. The final model was discriminatory for predicting 3-month FRT (AUC 0.87, 95% CI 0.87-0.88) and had good calibration (Brier 0.12, 0.11-0.12). Overall performance was moderate (F1-score 0.63 ± 0.004) and decision curve analyses suggested higher mean net benefit at lower thresholds of treatment (up to 0.8).

Conclusions: This FRT prediction model can help inform shared decision making and identify the most relevant features in the emergency setting. While it might be particularly useful in low resource healthcare settings, incorporation of further multifaceted variables is necessary to further increase the predictive performance.
Introduction

Even with the outstanding efficacy of mechanical thrombectomy (MT) for treatment of large-vessel acute ischemic stroke (AIS), 1 in 5 patients in the randomized controlled trials and 1 in 3 patients in real-world settings have very poor long-term outcome (modified Rankin Scale mRS 5-6 at 90 days) despite a technically successful intervention. The term futile intervention was coined to elucidate this issue. Similarly, 1 in 5 patients receiving intravenous thrombolysis (IVT) has very poor long-term outcome despite best available treatment with even higher prevalence rates in elderly patients. Given the need for informed shared-decision making and the high societal and health economics burden of both treatments, there is a need to proof that MT remains (cost-)effective in patients at high risk for futile recanalization therapies (FRT). Although individual patients’ preferences may vary, the five-year quality-adjusted-life-expectancy in stroke survivors reaching mRS 5 is minimal (0.06). Several clinical and imaging biomarkers associated with FRT have been identified. However, the models developed for MT are insufficient to reliably inform patient and proxies and to guide the decision-making in individual patients. Besides the MR PREDICT tool, which can be used for patients undergoing bridging IVT, no model for IVT-only patients not fulfilling endovascular trial criteria has been developed in the context of contemporary MT. A reliable prediction algorithm to identify patients that will have FRT with variables available at baseline is missing. Such an algorithm is needed to realistically inform patient and proxies about the potential risk/benefits of treatment and facilitate patient-oriented informed decision (e.g. withholding maximal therapy if advanced directives state that the patient does not wish to live with severe dependency). Additionally, since in several countries, access to MT and qualified personnel is limited, but with rising strain and workload, reliable FRT prediction could accommodate the rising demand, and increasing healthcare costs.

We hypothesized that the combination of several clinical, laboratory, imaging, and workflow variables would enable accurate prediction of FRT after both IVT and/or MT in order to better inform shared-decision making and possibly even to avoid futile therapies. Additionally, we wanted to perform internal validation of the developed model, analyze the clinical utility of the final model, and analyze potential differences in predictors of FRT between patients with and without detectable vessel occlusion.
Methods

This study adheres to the Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) statement.26,27

We included consecutive adult patients from the registry treated with IVT and/or MT between 01/2015-10/2020 and excluded patients with missing outcome at three months (12%, n=322). Patients with missing outcome had a slightly worse prognostic profile with a higher stroke severity, more history of TIA, longer onset-to-admission times and more frequent MT (see eTable 1 in the Supplement for full comparison). Patients had to receive either IVT alone, MT alone or a combination (bridging approach). At our center, we have a liberal approach of IVT and/or MT indications performing treatment also in borderline indications such as low stroke severity, distal occlusions, extended time window and pre-stroke impairment - see our guidelines for detailed indications/contraindications.28

The purpose of the developed model is the prospective prediction of very poor 3-month functional outcome in AIS patients who are potential candidates for IVT and/or MT. The intended use of the machine learning (ML) algorithm in the acute stroke workflow has been published and is provided in figure 1.29 In brief, is intended to inform shared-decision making with the patient and next of kin based on the probability of FRT, after the decision for IVT and/or MT treatment has been made, but before treatment has been actually started.

Standard Protocol Approvals, Registrations, and Patient Consents

Approval by local ethics committee was granted (Kantonale Ethikkommission, Bern, Switzerland: ID 231/2014 and ID 2020-01696) waiving consent of participants in accordance to Swiss Law.

[Figure 1 intended workflow]

We randomly split the final dataset (n=2261) in the training set (80%, n=1808) used for development and internal validation (20%, n=453). The flow chart showing patient inclusion and the study design is shown in eFigure 1 in the Supplement.
Clinical, laboratory, imaging and workflow variables were collected prospectively by dedicated research staff and missing values completed by using the clinical information system. Initially, 32 baseline admission variables were considered, but we dropped input variables (features) which were not routinely registered when more than 25% of the values were missing. Dropped variables included D-Dimer, HDL cholesterol, HbA1c, triglycerides, troponin and activated partial thromboplastin time. See eTable 2 in the Supplement for a full list of features considered and their respective definitions. The choice of independent variables was based on existing literature on pathophysiologically plausible associations with FRT that were available and documented in a sufficiently high data quality at our institution.

FRT was defined as mRS 5-6 at 3 months dichotomizing the outcome as a binary target variable, and was assessed by certified physicians during routine clinical visits or certified study nurses by semi-structured telephone interviews. Death was assessed through linkage with the national mortality registry. Assessors were not blinded, but unaware of this project at the time of assessment. Classes of FRT were imbalanced (24% mRS ≥ 5, 76% mRS <5, ratio 3.2:1). We imputed missing data with k-Nearest Neighbor (k-NN, k=15 neighbors) imputation on normalized data of the training and validation separately. We normalized features to [0,1] interval with min-max normalization.

We implemented Gradient Boosting (XGBoost classifier), by means of Python’s (3.7.7) scikit learn\(^30\) (0.22.1), xgboost\(^31\) (1.2.0, for XGBoost classifier) based on a previous project showing that this algorithm had good overall performance and high robustness.\(^29\)

We used a nested, stratified 10-fold cross-validation strategy for model development (XGBoost classifier).\(^29\) In the outer cross-validation loop, the training dataset was split into 10 equally sized subsets.\(^29\) Nine out of the ten subsets were used for training and one for validation. In the inner loop, hyperparameter optimization was performed based on maximization of f1-score in a 10-fold randomized gridsearch using the data of the previously formed nine folds.\(^29\)

For validation, we trained ML algorithms using the setting of the nested cross validation’s inner loop on the complete training data (10-fold randomized grid search) and applied the resulting ML model to
separate validation data (n=453). We repeated this process 20 times using a different random seed for algorithm initialization in each run.

Statistical analysis

We assessed univariate associations of clinical variables with FRT in the overall cohort, as well as training and validation data separately using standard descriptive statistics: Chi squared and Fisher’s exact tests for categorical variables, Whitney-Mann-U-Test for non-normally continuous or ordinally scaled variables, and Welsch’s t-test for independent normally distributed data. We used the pmsampsize Stata Package (StataCorp. 2019. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC) to calculate the minimum sample size required using 40 candidate predictors based on an assumed outcome prevalence of 25% and a lower bound for the new model's R-squared value of 0.25. This resulted in a final sample size of 1230 patients.

We report model discrimination, calibration and clinical utility. For discrimination of the models at the threshold of p=0.5, we report precision, recall, F1-score, accuracy, balanced accuracy, specificity, and Matthew’s correlation. Furthermore, we report the Area Under-the Curve (AUC) of the Receiver operating characteristic (ROC) and the Average Precision (AP) score throughout all possible thresholds. For calibration of the model, we report Brier score and Expected Calibration Error (ECE) grounded on 10 bins. We computed feature importance using Shapley Values Feature Importance feature importance for the XGBClassifier. Permutation feature importance was defined as decrease in F1-score by shuffling the values of a single feature randomly. Decision curve analysis is reported to quantify the clinical utility of the model on validation data. Since the a-priori probability threshold will vary according to healthcare resources, we did not define a fix a-priori threshold for risk of FRT to assess the net benefit. Results are reported as mean and confidence interval based on 10-fold cross-validation (model development). Performance on validation data was reported as mean ± 1 standard deviation over 20 runs with different random seeds for algorithm initialization.

Data Availability
Investigators may request access to anonymized individual patient data including analysis-ready datasets, and dataset specifications, after publication. Prior to use of the data, proposals need to be approved by an independent review panel at https://swissethics.ch/basec and a signed data sharing agreement will then be approved.
Results

Baseline characteristics

The final cohort included 2261 patients: median (IQR) age 75 years (64-83), 46% female, median NIHSS 9 (4-17). 34% received IVT alone, 41% MT alone and 25% both acute recanalization treatments (bridging approach). FRT occurred overall in 24% of patients and more often in patients receiving MT alone (34%) as compared to bridging patients (26%) and IVT alone (11%).

Baseline characteristics of patients in training and validation set are shown in Table 1. Taken together, older age, higher stroke severity, active cancer, and a higher cardiovascular risk profile were associated with FRT on univariate analysis.

[Table 1]

Discrimination of ML methods was good. Performance and a complete overview of the results are provided in Table 2. Overall, discrimination for predicting FRT (AUC 0.87 95% CI 0.87-0.88) and calibration (Brier 0.12, 0.11-0.12) were good and overall performance moderate (F1 score 0.63 ± 0.004) in the validation dataset (see eFigure 2 and 3 in the Supplement for ROC curves of the full model in the derivation and validation cohort).

[Table 2]

Shapley Values Feature Importance revealed that the most important features included clinical variables (higher stroke severity, older age, active cancer, prestroke disability), laboratory values (higher glucose, higher CRP, less dyslipidemia), imaging biomarkers (more white matter hyperintensities) and longer onset-to-admission time. The feature importances are shown in Figure 2 (see eFigures 4 and 5 in the Supplement for all features in patients with and without detectable vessel occlusion).

[Figure 2 feature importance]

In patients without detectable vessel occlusion, higher INR, and atrial fibrillation seemed to be more important, whereas white matter hyperintensity severity and active cancer seemed to be less important.
The decision curve for XGBClassifier is shown in Figure 3. The mean net benefit for p=0.8 was minimal (0.02 ± 0.01), and relevant net benefit was only present in lower a-priori thresholds. The net benefit – in this context – is weighing the profit obtained by classifying an individual with the outcome and the loss caused by falsely classifying an individual without the outcome. No evidence for harm was present throughout all thresholds.

[Figure 3 Decision Curve]
Discussion

The development and validation of a multivariable prediction model for futile thrombolysis and thrombectomy showed the following main findings:

(1) FRT overall occurs in 1 in 4 patients and more often in patients receiving MT alone (34%) as compared to IVT alone (11%). (2) The most relevant predictors of FRT included clinical variables (higher stroke severity, older age, active cancer, prestroke disability), laboratory values (higher glucose, higher CRP, higher creatinine), imaging biomarkers (more white matter hyperintensities) and longer onset-to-admission time. (3) The combination of several clinical, laboratory, neuroimaging and workflow variables at baseline showed good discrimination for prediction of FRT. (4) Our model will help to inform shared-decision making, but its usability to withhold treatment is uncertain and depends on healthcare resources. (5) Potential differences for prediction of FRT between patients with and without detectable vessel occlusion were identified.

Despite the success of IVT and MT in improving stroke outcome, rates of FRT remain considerable for both treatments. Given more liberal indications in the real-world as compared to the patients included in the original randomized controlled studies, for both treatments there is a gradual shift from selecting patients to deselecting patients. However, this development is likely to cause an increase in FRT. Although the costs of IVT and EVT are much lower as compared to established treatments e.g. for palliative cancer, the treatments pose enormous logistic, economic and ethical challenges in acute stroke treatment.

With this analysis, we have developed and internally validated a multivariable prediction algorithm to discriminate between patients which are likely or unlikely to have a very poor clinical outcome despite best available treatment with IVT and/or MT.

The lower rate of FRT in IVT patients might be related to the fact that patients meeting IVT criteria are early presenters, without extensive hemorrhagic changes. Some features included in our model have been partially described before, but several are novel and their relative importance can now be estimated (Figure 2). This allows clinicians to rapidly clarify the most relevant features in the
emergency setting. All variables considered in our model are available upon admission or can be obtained within a few minutes until the decision to perform IVT/MT has to be taken.

The intended use of our model are ischemic stroke patients before treatment has been started, but the decision for treatment with IVT and/or MT has been made. The derivation cohort is from a high-volume academic stroke center with low restrictions to perform those treatments. For clear-cut indications for MT, such as patients presenting early with severe stroke, the updated MR CLEAN predict tool might be more useful, but it cannot be applied to borderline indications or to inform decision on IVT-only patients not fulfilling endovascular trial criteria.

Our model has good discrimination and reasonable calibration, both in the derivation as well as the validation cohort. The model output can be used to inform and discuss with patients and/or next of kin the high risk of poor outcome and help to set realistic expectations. However, in high-resource healthcare systems, the algorithm will probably not be used since no patients will be excluded from acute recanalization treatments despite high chances for FRT. Additionally, false positives (classifying patients as FRT, despite the fact that they might gain functional independence) should be weighted more than false negative classifications. In other words, the positive predictive value of any FRT algorithm has to be as great as possible to evade skipping an evidence-based treatment from appropriate patients which may benefit from IVT/MT. Hence, we caution colleagues to apply the appealing AUC reported here and by others to individual patients with regards to the clinical utility in excluding patients from reperfusion therapies. However, there was no evidence of harm as shown by the decision curve analysis. Randomized controlled trials need to clarify whether the cost-effectiveness of MT is preserved in patients with high risk of poor outcome.

Nevertheless, the situation might be very different in low and mid income countries where access to IVT/MT is limited by shortage of MT devices, IVT medication, staff or other hurdles. Fittingly, our models exhibited net benefit in the lower probability thresholds. This means, that they might be most useful in settings of very limited healthcare resources. Cultural perception and individual preferences might also influence which a priori cut-off will be chosen in an individual scenario.
Additionally, this rather simple algorithm provides evidence that prediction of FRT is possible when different information sources (clinical, laboratory, imaging, time metrics) are combined. Further and more sophisticated information sources are promising additions for refinement of this algorithm. Those include ischemic core volume, ischemic core location, penumbra volume, covert brain infarctions, brain atrophy, masseter muscle, oxygen saturation, amongst others have not been accounted for. Reliable identification of FRT ahead of treatment is of utmost importance and this model and proposed refinements can serve as a starting point to reach the intended use. However, a recent study found that a large part of the variance in outcome after MT is explained by variables that are only known after the treatment decision has been made challenging the possibility to predict FRT in the emergency setting before knowing the outcome of the intervention – at least for MT patients.

Regarding the model performance, previous studies have shown that for most tabulated datasets, the differences in performance between different analytical machine learning methods and conventional logistic regression are minimal or non-existent. Strengths of this analysis include its large sample size with good quality data of predictors easily obtained in an emergency setting. This study has the limitations of a single center, retrospective registry potentially limiting its generalizability to other settings. Most importantly, no medical comparison group was available. Hence inference on a potential residual clinical benefit even in a patient with high probability for FRT is not possible. Additionally, 12% (322 patients) of the cohort had missing 3-month outcome and those had a slightly worse prognostic profile. Moreover, several advanced imaging parameters such as ischemic core volume, ischemic core location, penumbra volume and mismatch profile, covert brain infarctions, brain atrophy, masseter muscle volume, oxygen saturation amongst others have not been accounted for. Another limitation is the choice/availability of independent variables, since for example information on dementia/pre stroke cognitive impairment was unavailable. Information on medical history was obtained prospectively in the registry and includes variables obtained during hospital stay, so it is uncertain, how many of the predictors incorporated could be obtained within a short time period in an emergency setting. Hence,
our results need to be replicated by other groups and verified by upcoming randomized controlled trials on this issue in several subgroups with high-risk for FRT, such as low ASPECTS (NCT03805308, NCT03811769).

In conclusion, FRT occurs in 1 in 4 patients and more often in patients receiving MT alone (34%) as compared to IVT alone (11%). We identified clinical variables (higher stroke severity, older age, active cancer, prestroke disability), laboratory values (higher glucose, higher CRP, higher creatinine), imaging biomarkers (white matter hyperintensities) and time from onset-to-admission as the most relevant predictors of FRT. The prediction algorithm will help to inform shared-decision making and to set realistic expectations. Although the clinical benefit and usability of this algorithm for withholding treatments in settings with high healthcare resources is to be established in future studies, the development of a reliable algorithm for prediction of FRT seems to be within reach and should incorporate more advanced admission imaging features.
References


32. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ. 2020;368. doi:10.1136/bmj.m441


## Table 1. Baseline characteristics of all included patients.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Unit/ Definition</th>
<th>Training data (n=1808)</th>
<th>Validation data (n=453)</th>
<th>p-value</th>
<th>Training data (n=1808)</th>
<th>Validation data (n=453)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td>72.6 (61.6-81)</td>
<td>82.4 (74.5-87.7)</td>
<td>&lt;0.001</td>
<td>72.2 (62.4-80.9)</td>
<td>83.6 (75.5-88.1)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Female Sex</td>
<td></td>
<td>774 (56.2%)</td>
<td>206 (47.8%)</td>
<td>0.002</td>
<td>190 (55.1%)</td>
<td>55 (50.9%)</td>
<td>0.45</td>
</tr>
<tr>
<td>National Institute of Health Stroke Severity Scale</td>
<td></td>
<td>603 (43.8%)</td>
<td>225 (52.2%)</td>
<td></td>
<td>155 (44.9%)</td>
<td>53 (49.1%)</td>
<td></td>
</tr>
<tr>
<td>Active Cancer</td>
<td></td>
<td>7 (4-14)</td>
<td>17 (11-22)</td>
<td>&lt;0.001</td>
<td>8 (4-15)</td>
<td>17 (11-21)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Prestroke disability</td>
<td>Modified Rankin Scale</td>
<td>63 (4.6%)</td>
<td>65 (15.1%)</td>
<td>&lt;0.001</td>
<td>15 (4.3%)</td>
<td>19 (17.6%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Prestroke living situation</td>
<td></td>
<td>1,303 (94.6%)</td>
<td>354 (82.1%)</td>
<td>&lt;0.001</td>
<td>331 (95.9%)</td>
<td>90 (83.3%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>mmHg</td>
<td>156 (139-176)</td>
<td>161 (135-180)</td>
<td>0.40</td>
<td>160 (142-178)</td>
<td>156 (131.5-184)</td>
<td>0.61</td>
</tr>
<tr>
<td>Diastolic blood pressure</td>
<td>mmHg</td>
<td>84 (71-97)</td>
<td>81.27 (69-95)</td>
<td>0.059</td>
<td>83 (74-95)</td>
<td>78 (69-93.5)</td>
<td>0.028</td>
</tr>
<tr>
<td><strong>Medical History</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atrial Fibrillation / Flutter</td>
<td></td>
<td>423 (30.7%)</td>
<td>185 (42.9%)</td>
<td>&lt;0.001</td>
<td>100 (29.0%)</td>
<td>49 (45.4%)</td>
<td>0.002</td>
</tr>
<tr>
<td>Coronary heart disease</td>
<td></td>
<td>218 (15.8%)</td>
<td>100 (23.2%)</td>
<td>&lt;0.001</td>
<td>53 (15.4%)</td>
<td>25 (23.1%)</td>
<td>0.061</td>
</tr>
<tr>
<td>Diabetes Mellitus</td>
<td></td>
<td>215 (15.6%)</td>
<td>101 (23.4%)</td>
<td>&lt;0.001</td>
<td>54 (15.7%)</td>
<td>24 (22.2%)</td>
<td>0.11</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td></td>
<td>903 (65.6%)</td>
<td>252 (58.5%)</td>
<td>0.007</td>
<td>236 (68.4%)</td>
<td>58 (53.7%)</td>
<td>0.005</td>
</tr>
<tr>
<td>Arterial Hypertension</td>
<td></td>
<td>1,010 (73.3%)</td>
<td>350 (81.2%)</td>
<td>&lt;0.001</td>
<td>248 (71.9%)</td>
<td>86 (79.6%)</td>
<td>0.11</td>
</tr>
<tr>
<td>History of intracranial hemorrhage</td>
<td></td>
<td>23 (1.7%)</td>
<td>6 (1.4%)</td>
<td>0.69</td>
<td>3 (0.9%)</td>
<td>2 (1.9%)</td>
<td>0.39</td>
</tr>
<tr>
<td>Peripheral artery disease</td>
<td></td>
<td>50 (3.6%)</td>
<td>35 (8.1%)</td>
<td>&lt;0.001</td>
<td>16 (4.6%)</td>
<td>6 (3.6%)</td>
<td>0.70</td>
</tr>
<tr>
<td>Prosthetic heart valves</td>
<td></td>
<td>None</td>
<td></td>
<td></td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biological</td>
<td></td>
<td>1,329 (96.5%)</td>
<td>415 (96.3%)</td>
<td>0.93</td>
<td>337 (97.7%)</td>
<td>101 (93.5%)</td>
<td>0.10</td>
</tr>
<tr>
<td>Mechanical</td>
<td></td>
<td>28 (2.0%)</td>
<td>10 (2.3%)</td>
<td></td>
<td>5 (1.4%)</td>
<td>4 (3.7%)</td>
<td></td>
</tr>
<tr>
<td>Biological</td>
<td></td>
<td>20 (1.5%)</td>
<td>6 (1.4%)</td>
<td></td>
<td>3 (0.9%)</td>
<td>3 (2.8%)</td>
<td></td>
</tr>
<tr>
<td>Smoking</td>
<td>324 (23.5%)</td>
<td>46 (10.7%)</td>
<td>&lt;0.001</td>
<td>83 (24.1%)</td>
<td>12 (11.1%)</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>------------</td>
<td>--------</td>
<td>-------------</td>
<td>------------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>History of stroke</td>
<td>179 (13.0%)</td>
<td>80 (18.6%)</td>
<td>0.004</td>
<td>37 (10.7%)</td>
<td>13 (12.0%)</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>History of transient ischemic attack</td>
<td>57 (4.1%)</td>
<td>17 (3.9%)</td>
<td>0.86</td>
<td>19 (5.5%)</td>
<td>5 (4.6%)</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td><strong>Neuroimaging</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White matter disease</td>
<td>0</td>
<td>355 (25.8%)</td>
<td>48 (11.1%)</td>
<td>&lt;0.001</td>
<td>105 (30.4%)</td>
<td>15 (13.9%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>1</td>
<td>555 (40.3%)</td>
<td>127 (29.5%)</td>
<td>147 (42.6%)</td>
<td>34 (31.5%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>304 (22.1%)</td>
<td>131 (30.4%)</td>
<td>60 (17.4%)</td>
<td>34 (31.5%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>163 (11.8%)</td>
<td>125 (29.0%)</td>
<td>33 (9.6%)</td>
<td>25 (23.1%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detectable vessel occlusion</td>
<td>645 (46.8%)</td>
<td>121 (28.1%)</td>
<td>&lt;0.001</td>
<td>169 (49.0%)</td>
<td>24 (22.2%)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intraarterial treatment</td>
<td>834 (60.6%)</td>
<td>364 (84.5%)</td>
<td>&lt;0.001</td>
<td>205 (59.4%)</td>
<td>94 (87.0%)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Intravenous thrombolysis</td>
<td>874 (63.5%)</td>
<td>184 (42.7%)</td>
<td>&lt;0.001</td>
<td>234 (67.8%)</td>
<td>43 (39.8%)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Symptom onset to admission time minutes</td>
<td>157 (82-321)</td>
<td>190 (107-400)</td>
<td>&lt;0.001</td>
<td>135 (74-254)</td>
<td>215 (126-539)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td><strong>Laboratory Values</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum creatinine umol/L</td>
<td>79 (67-92)</td>
<td>85 (66-108)</td>
<td>&lt;0.001</td>
<td>80 (67-93)</td>
<td>85 (70-104)</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>C reactive protein (CRP) mg/L</td>
<td>3.93 (3-8.6)</td>
<td>8 (3-19)</td>
<td>&lt;0.001</td>
<td>2.5 (1.88-3.18)</td>
<td>2.4 (1.9-2.9)</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Plasma glucose mmol/L</td>
<td>6.4 (5.7-7.5)</td>
<td>7.1 (5.9-8.5)</td>
<td>&lt;0.001</td>
<td>6.4 (5.8-7.4)</td>
<td>7.05 (6-8.65)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Low-density lipoprotein cholesterol mmol/L</td>
<td>2.6 (1.98-3.25)</td>
<td>2.4 (1.99-2.99)</td>
<td>0.006</td>
<td>2.5 (1.88-3.18)</td>
<td>2.4 (1.9-2.9)</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Total cholesterol mmol/L</td>
<td>4.66 (3.9-5.3)</td>
<td>4.34 (3-5.0)</td>
<td>&lt;0.001</td>
<td>4.62 (3.9-5.3)</td>
<td>4.4 (3.9-4.9)</td>
<td>0.049</td>
<td></td>
</tr>
<tr>
<td>Haemoglobin mg/L</td>
<td>137 (128-146)</td>
<td>127 (116-137)</td>
<td>&lt;0.001</td>
<td>136 (128-144)</td>
<td>127 (116-138)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>International normalized ratio (INR)</td>
<td>1.02 (1-1.08)</td>
<td>1.06 (1.01-1.16)</td>
<td>&lt;0.001</td>
<td>1.02 (1-1.06)</td>
<td>1.06 (1.02-1.13)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Platelet count G/L</td>
<td>222 (190-256)</td>
<td>224 (185-264)</td>
<td>0.82</td>
<td>224 (195-259)</td>
<td>225 (187-257)</td>
<td>0.45</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Performance of the model for mRS 5-6 prediction. Values are reported as mean and 95% CI for 10-fold cross validation and mean +/- SD for validation.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>10-fold cross-validation</th>
<th>Validation data</th>
<th>Subgroup with detectable vessel occlusion</th>
<th>Subgroup without detectable vessel occlusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>0.84 (0.80-0.87)</td>
<td>0.84 (0.82-0.85)</td>
<td>0.78 (0.76-0.80)</td>
<td>0.88 (0.86-0.90)</td>
</tr>
<tr>
<td>Balanced Accuracy</td>
<td>0.73 (0.67-0.80)</td>
<td>0.78 (0.75-0.80)</td>
<td>0.72 (0.70-0.75)</td>
<td>0.72 (0.64-0.79)</td>
</tr>
<tr>
<td>Precision</td>
<td>0.71 (0.59-0.83)</td>
<td>0.65 (0.62-0.68)</td>
<td>0.66 (0.60-0.72)</td>
<td>0.64 (0.55-0.73)</td>
</tr>
<tr>
<td>Recall</td>
<td>0.53 (0.39-0.67)</td>
<td>0.67 (0.62-0.72)</td>
<td>0.58 (0.51-0.64)</td>
<td>0.48 (0.32-0.64)</td>
</tr>
<tr>
<td>F1-Score</td>
<td>0.60 (0.50-0.71)</td>
<td>0.66 (0.63-0.69)</td>
<td>0.61 (0.58-0.65)</td>
<td>0.54 (0.44-0.64)</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.93 (0.88-0.98)</td>
<td>0.89 (0.88-0.90)</td>
<td>0.87 (0.82-0.91)</td>
<td>0.95 (0.92-0.98)</td>
</tr>
<tr>
<td>Brier score</td>
<td>0.12 (0.10-0.14)</td>
<td>0.12 (0.11-0.12)</td>
<td>0.15 (0.15-0.15)</td>
<td>0.10 (0.09-0.11)</td>
</tr>
<tr>
<td>ECE</td>
<td>0.07 (0.04-0.09)</td>
<td>0.06 (0.05-0.07)</td>
<td>0.07 (0.05-0.09)</td>
<td>0.11 (0.08-0.13)</td>
</tr>
<tr>
<td>ROC-AUC</td>
<td>0.86 (0.82-0.91)</td>
<td>0.87 (0.87-0.88)</td>
<td>0.83 (0.82-0.83)</td>
<td>0.84 (0.82-0.87)</td>
</tr>
</tbody>
</table>
Figure Legends

Figure 1. Intended use of the model in a stroke workflow.
IVT: intravenous thrombolysis; MT: mechanical thrombectomy; mRS: modified Rankin Scale; ML: machine learning. The ML model outputs a probability (risk score) for mRS 5-6 based on variables available ahead of recanalization therapies. This information is provided to the treating physicians after selection of patient for recanalization and could serve as a marker for futile recanalization. Workflow derived from Meier et al. with permission. 29

Figure 2. Shapley Values Feature Importance
computed for the XGBoost (mean values of 10 feature permutations and 20 random initializations). A) all patients B) only patients with detectable vessel occlusion C) only patients without detectable vessel occlusion. See the Supplement for full model and features. Frame indicates the ten most important variables.
Figure 3. Decision curve analysis for XGBClassifier.
The blue line indicates the mean over 20 random initializations. The rug plot shows the samples used for computation of the decision curve. ‘No risk’ denotes the trivial strategy of always predicting mRS 0-4, and ‘All Risk’ denotes the strategy of always predicting mRS 5-6.
Multivariable Prediction Model for Futile Recanalization Therapies in Patients With Acute Ischemic Stroke

Thomas Meinel, Christine Lerch, Urs Fischer, et al.

Neurology published online July 8, 2022
DOI 10.1212/WNL.0000000000200815

This information is current as of July 8, 2022