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

Thomas Raphael Meinel, Christine Lerch, Urs Fischer, Morin Beyeler, MD, Adnan Mujanovic, MD, Christoph Kurmann, MD, Bernhard Siepen, MD, Adrian Scutelnic, MD, Madlaine Müller, MD, Martina Goedlhin, MD, Nebyat Filate Belachew, MD, Tomas Dobrocky, MD, Jan Grailla, David Seiffge, Simon Jung, MD, Marcel Arnold, MD, Roland Wiest, MD, Raphael Meier, PhD,* and Johannes Kaesmacher, MD*

Neurology® 2022;99:e1009-e1018. doi:10.1212/WNL.0000000000200815

Abstract

Background and Objectives
Very poor outcome despite IV thrombolysis (IVT) and mechanical thrombectomy (MT) occurs in approximately 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 (FRTs) in patients undergoing those therapies.

Methods
Patients from a prospectively collected observational registry of a single academic stroke center treated with MT and/or IVT were included. The data set was split into a training (N = 1,808, 80%) and internal validation (N = 453, 20%) cohort. We used gradient boosted decision tree machine learning models after k-nearest neighbor imputation of 32 variables available at admission to predict FRT defined as modified Rankin scale 5–6 at 3 months. We report feature importance, ability for discrimination, calibration, and decision curve analysis.

Results
A total of 2,261 patients with a median (interquartile range) age of 75 years (64–83 years), 46% female, median NIH Stroke Scale 9 (4–17), 34% IVT alone, 41% MT alone, and 25% bridging were included. Overall, 539 (24%) had FRT, more often in MT alone (34%) as compared with IVT alone (11%). Feature importance identified clinical variables (stroke severity, age, active cancer, prestroke disability), laboratory values (glucose, C-reactive protein, 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 (area under the curve 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. Although it might be particularly useful in low resource healthcare settings, incorporation of further multifaceted variables is necessary to further increase the predictive performance.

*These authors contributed equally to this work as co-senior authors.

From the Department of Neurology (T.M., C.L., M.B., B.S., A.S., M.G., D.S., S.J., M.A.), University Hospital Bern, Inselspital, University of Bern; Department of Neurology and Stroke Center (U.F.), University Hospital Basel and University of Basel; University Institute of Diagnostic and Interventional Neuroradiology (A.M., C.K., N.F.B., T.D., J.G., R.W., R.M., J.K.), Support Center for Advanced Neuroimaging (R.W., R.M., J.K.), Department of Diagnostic, Paediatric and Interventional Radiology (J.K.), University Hospital Bern, Inselspital, University of Bern, Switzerland.

Go to Neurology.org/N for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

The Article Processing Charge was funded by the authors.

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.
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\(^1\) and 1 in 3 patients in real-world settings\(^2\) 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.\(^3\)–\(^5\) Similarly, 1 in 5 patients receiving IV thrombolysis (IVT) has a very poor long-term outcome despite best available treatment with even higher prevalence rates in elderly patients.\(^6\)

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 (FRTs).\(^4\)–\(^7\) Although individual patients’ preferences may vary, the 5-year quality-adjusted life expectancy in stroke survivors reaching mRS 5 is minimal (0.06).\(^8\)–\(^10\) Several clinical and imaging biomarkers associated with FRT have been identified.\(^4\)–\(^11\)–\(^18\) However, the models developed for MT are insufficient to reliably inform patient and proxies and to guide the decision-making in individual patients.\(^19\)–\(^22\) Besides the MR PREDiCT tool,\(^23\) 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 who 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). In addition, because in several countries, access to MT and qualified personnel is limited,\(^24\) but with rising strain and workload,\(^25\) 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 MT to better inform shared decision making and possibly even to avoid futile therapies. In addition, 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 multi-variable prediction model for individual prognosis or diagnosis statement.\(^26\)–\(^27\)

We included consecutive adult patients from the registry treated with IVT and/or MT between January 2015 and October 2020 and excluded patients with missing outcome at 3 months (12%, n = 322). Patients with missing outcome had a slightly worse prognostic profile with a higher stroke severity, more history of transient ischemic attack, longer onset-to-admission times, and more frequent MT (see eTable 1, links.lww.com/WNL/C129 for full comparison). Patients had to receive 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 prestroke 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 patients with AIS 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 presented in Figure 1.\(^29\) In brief, it 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 with Swiss Law.

We randomly split the final data set (n = 2,261) in the training set (80%, n = 1,808) used for development and internal validation (20%, n = 453). The flowchart showing patient inclusion and the study design is shown in eFigure 1 (links.lww.com/WNL/C129).

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) that were not routinely registered when more than 25% of the values were missing. Dropped variables included D-dimer, high-density lipoprotein cholesterol, hemoglobin A1c, triglycerides, troponin, and activated partial thromboplastin time. See eTable 2 (links.lww.com/WNL/C129) 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 semistructured 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 = 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 (3.7.7) scikit learn30 (0.22.1), XGBoost31 (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 data set was split into 10 equally sized subsets.29 Nine of the 10 subsets were used for training and 1 for validation. In the inner loop, hyperparameter optimization was performed based on maximization of F1-score in a 10-fold randomized grid search using the data of the previously formed 9 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).29 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: χ² and Fisher exact tests for categorical variables, the Mann-Whitney U test for non-normally continuous or ordinally scaled variables, and the Welsch f test for independent normally distributed data. We used the pmsampsize Stata Package (Stata Statistical Software: Release 16; StataCorp, College Station, TX) 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 $R^2$ value of 0.25.12 This resulted in a final sample size of 1,230 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 correlation. Furthermore, we report the area under the curve (AUC) of the receiver operating characteristic (ROC) and the average precision score throughout all possible thresholds. For calibration of the model, we report Brier score and expected calibration error grounded on 10 bins. We computed feature importance using Shapley values for the XGBClassifier. Permutation feature importance was defined as a decrease in F1-score by shuffling the values of a single feature randomly. Decision curve analysis33 is reported to quantify the clinical utility of the model on validation data. Because 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. The results are reported as mean and CI based on 10-fold cross-validation (model development). Performance on validation
### Table 1 Baseline Characteristics of all Included Patients

| Feature                  | Unit/definition | Training data (n = 1,808) | Validation data (n = 453) | p Value |
|--------------------------|-----------------|---------------------------|---------------------------|---------|
|                          | No FRT (n = 1,377) | FRT (n = 431)             | No FRT (n = 345)         | FRT (n = 108) |
|                          |                  |                           |                           |         |
| **Clinical variables**   |                 |                           |                           |         |
| Age                      |                 | 72.6 (61.6–81)            | 82.4 (74.5–87.7)          | <0.001  |
|                          |                 | 72.2 (62.4–80.9)          | 83.6 (75.5–88.1)          | <0.001  |
| Sex                      | Male            | 774 (56.2)                | 206 (47.8)                | 0.002   |
|                          |                 | 190 (55.1)                | 55 (50.9)                 | 0.45    |
|                          | Female          | 603 (43.8)                | 225 (52.2)                |         |
|                          |                 | 155 (44.9)                | 53 (49.1)                 |         |
| NIH Stroke Severity Scale|                 | 7 (4–14)                  | 17 (11–22)                | <0.001  |
|                          |                 | 8 (4–15)                  | 17 (11–21)                | <0.001  |
| Active cancer            |                 | 63 (4.6)                  | 65 (15.1)                | <0.001  |
|                          |                 | 15 (4.3)                  | 19 (17.6)                | <0.001  |
| Prestroke disability     | Modified Rankin scale | 0 (0–1)                  | 1 (0–2)                  | <0.001  |
|                          |                 | 0 (0–1)                   | 1 (0–2)                  | <0.001  |
| Prestroke living situation| Home           | 1,303 (94.6)              | 354 (82.1)               | <0.001  |
|                          |                 | 331 (95.9)                | 90 (83.3)                | <0.001  |
|                          | Nursing home    | 57 (4.1)                  | 70 (16.2)                |         |
|                          |                 | 11 (3.2)                  | 13 (12.0)                |         |
|                          | Other medical facility | 12 (0.9)                | 5 (1.2)                  |         |
|                          |                 | 2 (0.6)                   | 5 (4.6)                  |         |
|                          | Rehabilitation hospital | 5 (0.4)                | 2 (0.5)                  |         |
|                          |                 | 1 (0.3)                   | 0 (0.0)                  |         |
| Systolic blood pressure  | mm Hg           | 156 (139–176)            | 161 (135–180)            | 0.40    |
|                          |                 | 160 (142–178)            | 156 (131.5–184)          | 0.61    |
| Diastolic blood pressure | mm Hg           | 84 (71–97)                | 81.27 (69–95)            | 0.059   |
|                          |                 | 83 (74–95)                | 78 (69–93.5)             | 0.028   |
| **Medical history**      |                 |                           |                           |         |
| Atrial fibrillation/flutter |              | 423 (30.7)                | 185 (42.9)               | <0.001  |
|                          |                 | 100 (29.0)                | 49 (45.4)                | 0.002   |
| Coronary heart disease   |                 | 218 (15.8)                | 100 (23.2)               | <0.001  |
|                          |                 | 53 (15.4)                 | 25 (23.1)                | 0.061   |
| Diabetes mellitus        |                 | 215 (15.6)                | 101 (23.4)               | <0.001  |
|                          |                 | 54 (15.7)                 | 24 (22.2)                | 0.11    |
| Hyperlipidemia           |                 | 903 (65.6)                | 252 (58.5)               | 0.007   |
|                          |                 | 236 (68.4)                | 58 (53.7)                | 0.005   |
| Arterial hypertension    |                 | 1,010 (73.3)              | 350 (81.2)               | <0.001  |
|                          |                 | 248 (71.9)                | 86 (79.6)                | 0.11    |
| History of intracranial hemorrhage |      | 23 (1.7)                 | 6 (1.4)                  | 0.69    |
|                          |                 | 3 (0.9)                   | 2 (1.9)                  | 0.39    |
| Peripheral artery disease|                 | 50 (3.6)                 | 35 (8.1)                 | <0.001  |
|                          |                 | 16 (4.6)                  | 6 (5.6)                  | 0.70    |
| Prosthetic heart valves  | None            | 1,329 (96.5)              | 415 (96.3)               | 0.93    |
|                          |                 | 337 (97.7)                | 101 (93.5)               | 0.10    |
|                          | Biological      | 28 (2.0)                 | 10 (2.3)                 | 5 (1.4) |
|                          |                 | 4 (3.7)                   |                           |         |
|                          | Mechanical      | 20 (1.5)                 | 6 (1.4)                  | 3 (0.9) |
|                          |                 | 3 (2.8)                   |                           |         |
| Smoking                  |                 | 324 (23.5)                | 46 (10.7)                | <0.001  |
|                          |                 | 83 (24.1)                 | 12 (11.1)                | 0.004   |
| History of stroke        |                 | 179 (13.0)                | 80 (18.6)                | 0.004   |
|                          |                 | 37 (10.7)                 | 13 (12.0)                | 0.70    |
| History of transient ischemic attack |      | 57 (4.1)                 | 17 (3.9)                 | 0.86    |
|                          |                 | 19 (5.5)                  | 5 (4.6)                  | 0.72    |
| **Neuroimaging**         |                 |                           |                           |         |
| White matter disease     |                 | 0                         | 355 (25.8)               | <0.001  |
|                          |                 | 48 (11.1)                 | 105 (30.4)               | 15 (13.9) | <0.001 |
|                          | 1                | 555 (40.3)                | 127 (29.5)               | 147 (42.6) | 34 (31.5) |
|                          | 2                | 304 (22.1)                | 131 (30.4)               | 60 (17.4) | 34 (31.5) |
|                          | 3                | 163 (11.8)                | 125 (29.0)               | 33 (9.6)   | 25 (23.1) |
| Detectable vessel occlusion |              | 645 (46.8)               | 121 (28.1)               | <0.001  |
|                          |                 | 169 (49.0)                | 24 (22.2)                | <0.001  |
| **Treatment**            |                 |                           |                           |         |
| Intra-arterial treatment |                 | 834 (60.6)                | 364 (84.5)               | <0.001  |
|                          |                 | 205 (59.4)                | 94 (87.0)                | <0.001  |
| IV thrombolysis          |                 | 874 (63.5)                | 184 (42.7)               | <0.001  |
|                          |                 | 234 (67.8)                | 43 (39.8)                | <0.001  |

Continued
Laboratory values

| Feature                      | Unit/definition | Training data (n = 1,808) | Validation data (n = 453) | p Value |
|------------------------------|-----------------|---------------------------|---------------------------|---------|
|                              |                 | No FRT (n = 1,377) | FRT (n = 431) | No FRT (n = 345) | FRT (n = 108) |
| Symptom onset to admission time | min             | 157 (82–321)         | 190 (107–400)          | <0.001  | 135 (74–254)         | 215 (126–539)          | <0.001  |
| Laboratory values            |                 |                          |                          |         |                          |                          |         |
| Serum creatinine             | μmol/L          | 79 (67–92)             | 85 (66–108)            | <0.001  | 80 (67–93)             | 85 (70–104)            | 0.18    |
| C-reactive protein           | mg/L            | 3.93 (3–8.6)           | 8 (3–19)               | <0.001  | 2.5 (1.88–3.18)        | 2.4 (1.9–2.9)          | 0.10    |
| Plasma glucose               | mmol/L          | 6.4 (5.7–7.5)          | 7.1 (5.9–8.5)          | <0.001  | 6.4 (5.8–7.4)          | 7.05 (6–8.65)          | <0.001  |
| Low-density lipoprotein cholesterol | mmol/L           | 2.6 (1.98–3.25) | 2.4 (1.99–2.99) | 0.006   | 2.5 (1.88–3.18)        | 2.4 (1.9–2.9)          | 0.10    |
| Total cholesterol            | mmol/L          | 4.66 (3.9–5.3)         | 4.34 (3.8–5.0)         | <0.001  | 4.62 (3.9–5.3)         | 4.4 (3.9–4.9)          | 0.049   |
| Hemoglobin                   | mg/L            | 137 (128–146)          | 127 (116–137)          | <0.001  | 136 (128–144)          | 127 (116–138)          | <0.001  |
| International normalized ratio|                | 1.02 (1–1.08)         | 1.06 (1.01–1.16)       | <0.001  | 1.02 (1–1.06)          | 1.06 (1.02–1.13)       | <0.001  |
| Platelet count               | g/L             | 222 (190–256)          | 224 (185–264)          | 0.82    | 224 (195–259)          | 225 (187–257)          | 0.45    |

Abbreviation: FRT = futile recanalization therapy. Values are reported as median (interquartile range) or n (%).

Data was reported as mean ±1 SD over 20 runs with different random seeds for algorithm initialization.

Data Availability

Investigators may request access to anonymized individual patient data including analysis-ready data sets, and data set specifications, after publication. Before using the data, proposals need to be approved by an independent review panel at swissethics.ch/basec, and a signed data sharing agreement will then be approved.

Results

Baseline Characteristics

The final cohort included 2,261 patients: median (interquartile range) age 75 years (64–83 years), 46% female, and median NIH Stroke Scale 9 (4–17). In total, 34% of patients 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 with bridging patients (26%) and IVT alone (11%).

Baseline characteristics of patients in training and validation set are presented 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.

Discrimination of ML methods was good. Performance and a complete overview of the results are presented 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) was good, and overall performance was moderate (F1-score 0.63 ± 0.004) in the validation data set (see eFigures 2 and 3, links.lww.com/WNL/C129, for ROC curves of the full model in the derivation and validation cohort).

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 C-reactive protein [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, links.lww.com/WNL/C129, for all features in patients with and without detectable vessel occlusion).

In patients without detectable vessel occlusion, higher international normalized ratio 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.44 No evidence for harm was present throughout all thresholds.

Discussion

The development and validation of a multivariable prediction model for futile thrombolysis and thrombectomy showed the
follow the main findings: (1) FRT overall occurs in 1 in 4 patients and more often in patients receiving MT alone (34%) as compared with 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.26 Given more liberal indications in the real-world as compared with the patients included in the original randomized controlled studies,35,36 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 when compared with established treatments, for example, 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 who 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,14,37-40 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 on 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 is 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 tool23 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.41 However, in high-resource healthcare systems, the algorithm will probably not be used because no patients will be excluded from acute recanalization treatments despite high chances for FRT. In addition, 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 others37,42 to individual patients

### Table 2 Performance of the Model for mRS 5–6 Prediction

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

Abbreviations: AUC = area under the curve; ECE = expected calibration error; mRS = modified Rankin scale; ROC = receiver operating characteristic. Values are reported as mean and 95% CI for 10-fold cross-validation and mean ± SD for validation.
regarding 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-income 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 cutoff will be chosen in an individual scenario.

In addition, 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 volume.

---

**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, and (C) only patients without detectable vessel occlusion. See the Supplement (links.lww.com/WNL/C129) for full model and features. Frame indicates the 10 most important variables. NIHSS = NIH Stroke Scale.
location, penumbra volume, covert brain infarctions, brain atrophy, masseter muscle, and oxygen saturation among 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 data sets, the differences in performance between different analytical ML methods and conventional logistic regression are minimal or nonexistent.

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. In addition, 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, and oxygen saturation among others have not been accounted for. Another limitation is the choice/availability of independent variables, for example, information on dementia/prestroke 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 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 with 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.

**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 and FF18059) and the University of Bern (a digital reference network platform for clinical and experimental neuroscience—deep phenotyping and data integration).

**Disclosure**
U. Fischer has received research grants from Medtronic (SWIFT DIRECT and BEYOND SWIFT) and does consultancy for Medtronic, Stryker and CSL Behring outside the submitted work. J. Gralla has received research grants from
Müller, MD
Madlaine Siepen, MD
Bernhard MD Kurmann, Christoph
Adnan Morin

Department of Neurology

Lerch Christine

Appendix (continued)

| Name            | Location                                                                 | Contribution                                                                 |
|-----------------|--------------------------------------------------------------------------|-----------------------------------------------------------------------------|
| Martina Goeldlin, MD | Department of Neurology, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data |
| Nebiyat Filate Belachew, MD | University Institute of Diagnostic and Interventional Neuroradiology, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data |
| Tomas Dobrocky, MD | University Institute of Diagnostic and Interventional Neuroradiology, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; study concept or design |
| Jan Gralla, MD | University Institute of Diagnostic and Interventional Neuroradiology, University Hospital Bern, Inselspital, University of Bern, Switzerland | 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 | Department of Neurology, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; study concept or design |
| Simon Jung, MD | Department of Neurology, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; study concept or design |
| Marcel Arnold, MD | Department of Neurology, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design |
| Roland Wiest, MD | University Institute of Diagnostic and Interventional Neuroradiology, and Support Center for Advanced Neuroimaging, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; study concept or design |
| Raphael Meier, PhD | University Institute of Diagnostic and Interventional Neuroradiology, Support Center for Advanced Neuroimaging, and Department of Diagnostic, Paediatric and Interventional Radiology, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data |
| Johannes Kaesmacher, MD | University Institute of Diagnostic and Interventional Neuroradiology, Support Center for Advanced Neuroimaging, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data |

Appendix Authors

| Name               | Location                                                                 | Contribution                                                                 |
|--------------------|--------------------------------------------------------------------------|-----------------------------------------------------------------------------|
| Thomas Raphael Meinel | Department of Neurology, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; study concept or design |
| Christine Lerch | Department of Neurology, University Hospital Bern, Inselspital, University of Bern, Switzerland | 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 | Department of Neurology and Stroke Center, University Hospital Basel and University of Basel, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; study concept or design |
| Morin Beyeler, MD | Department of Neurology, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data |
| Adnan Mujanovic, MD | University Institute of Diagnostic and Interventional Neuroradiology, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data |
| Christoph Kurmann, MD | University Institute of Diagnostic and Interventional Neuroradiology, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data |
| Bernhard Siepen, MD | Department of Neurology, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data |
| Adrian Scutelnic, MD | Department of Neurology, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data |
| Madlaine Müller, MD | Department of Neurology, University Hospital Bern, Inselspital, University of Bern, Switzerland | Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data |

Publication History

Received by Neurology January 5, 2022. Accepted in final form April 19, 2022. Submitted and externally peer reviewed. The handling editor was José Merino, MD, MPhil, FAAN.
Multivariable Prediction Model for Futile Recanalization Therapies in Patients With Acute Ischemic Stroke

Thomas Raphael Meinel, Christine Lerch, Urs Fischer, et al.
*Neurology* 2022;99:e1009-e1018 Published Online before print July 8, 2022
DOI 10.1212/WNL.00000000000200815

This information is current as of July 8, 2022

| Updated Information & Services | including high resolution figures, can be found at: [http://n.neurology.org/content/99/10/e1009.full](http://n.neurology.org/content/99/10/e1009.full) |
|-------------------------------|-------------------------------------------------------------------------------------------------------------------------------------|
| References                    | This article cites 48 articles, 16 of which you can access for free at: [http://n.neurology.org/content/99/10/e1009.full#ref-list-1](http://n.neurology.org/content/99/10/e1009.full#ref-list-1) |
| Subspecialty Collections      | This article, along with others on similar topics, appears in the following collection(s): All Cerebrovascular disease/Stroke [http://n.neurology.org/cgi/collection/all_cerebrovascular_disease_stroke](http://n.neurology.org/cgi/collection/all_cerebrovascular_disease_stroke) |
| Permissions & Licensing       | Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at: [http://www.neurology.org/about/about_the_journal#permissions](http://www.neurology.org/about/about_the_journal#permissions) |
| Reprints                      | Information about ordering reprints can be found online: [http://n.neurology.org/subscribers/advertise](http://n.neurology.org/subscribers/advertise) |