

Article

Dynamic Prediction of Mechanical Thrombectomy Outcome for Acute Ischemic Stroke Patients Using Machine Learning

Yixing Hu ^{1,2,†}, Tongtong Yang ^{1,2,†}, Juan Zhang ³, Xixi Wang ⁴, Xiaoli Cui ³, Nihong Chen ⁴, Junshan Zhou ⁴, Fuping Jiang ⁵, Junrong Zhu ^{2,6} and Jianjun Zou ^{2,6,*}

- ¹ School of Basic Medicine and Clinical Pharmacy, China Pharmaceutical University, Nanjing 210009, China; 3320092130@stu.cpu.edu.cn (Y.H.); 3320092195@stu.cpu.edu.cn (T.Y.)
- ² Department of Pharmacy, Nanjing First Hospital, Nanjing Medical University, Nanjing 210029, China; junrong_zhu@aliyun.com
- ³ Department of Neurology, Yuhua Branch of Nanjing First Hospital, Nanjing Yuhua Hospital, Nanjing Medical University, Nanjing 210029, China; zhangjuan_1@126.com (J.Z.); 17720809749@163.com (X.C.)
- ⁴ Department of Neurology, Nanjing First Hospital, Nanjing Medical University, Nanjing 210029, China; w15895820585@163.com (X.W.); y19850867562@163.com (N.C.); zhjsh333@126.com (J.Z.)
- ⁵ Department of Geriatrics, Nanjing First Hospital, Nanjing Medical University, Nanjing 210029, China; fupingjiang@163.com
- ⁶ Department of Pharmacy, Nanjing First Hospital, China Pharmaceutical University, Nanjing 210029, China
- * Correspondence: zoujianjun100@126.com
- † These authors contributed equally to this work.



Citation: Hu, Y.; Yang, T.; Zhang, J.; Wang, X.; Cui, X.; Chen, N.; Zhou, J.; Jiang, F.; Zhu, J.; Zou, J. Dynamic Prediction of Mechanical Thrombectomy Outcome for Acute Ischemic Stroke Patients Using Machine Learning. *Brain Sci.* **2022**, *12*, 938. <https://doi.org/10.3390/brainsci12070938>

Academic Editor:
Gianluca Borghini

Received: 16 June 2022
Accepted: 15 July 2022
Published: 18 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The unfavorable outcome of acute ischemic stroke (AIS) with large vessel occlusion (LVO) is related to clinical factors at multiple time points. However, predictive models used for dynamically predicting unfavorable outcomes using clinically relevant preoperative and postoperative time point variables have not been developed. Our goal was to develop a machine learning (ML) model for the dynamic prediction of unfavorable outcomes. We retrospectively reviewed patients with AIS who underwent a consecutive mechanical thrombectomy (MT) from three centers in China between January 2014 and December 2018. Based on the eXtreme gradient boosting (XGBoost) algorithm, we used clinical characteristics on admission (“Admission” Model) and additional variables regarding intraoperative management and the postoperative National Institute of Health stroke scale (NIHSS) score (“24-Hour” Model, “3-Day” Model and “Discharge” Model). The outcome was an unfavorable outcome at the three-month mark (modified Rankin scale, mRS 3–6: unfavorable). The area under the receiver operating characteristic curve and Brier scores were the main evaluating indexes. The unfavorable outcome at the three-month mark was observed in 156 (62.0%) of 238 patients. These four models had a high accuracy in the range of 75.0% to 87.5% and had a good discrimination with AUC in the range of 0.824 to 0.945 on the testing set. The Brier scores of the four models ranged from 0.122 to 0.083 and showed a good predictive ability on the testing set. This is the first dynamic, preoperative and postoperative predictive model constructed for AIS patients who underwent MT, which is more accurate than the previous prediction model. The preoperative model could be used to predict the clinical outcome before MT and support the decision to perform MT, and the postoperative models would further improve the predictive accuracy of the clinical outcome after MT and timely adjust therapeutic strategies.

Keywords: mechanical thrombectomy; acute ischemic stroke; dynamic prediction; machine learning; unfavorable outcome

1. Introduction

Acute ischemic stroke (AIS) is the main cause of lifelong disability and mortality around the world [1,2]. Therefore, mechanical thrombectomy (MT) is now regarded as a standard of care for managing patients with acute large vessel occlusion stroke [3–6].

However, despite high recanalization rates, almost one-third of stroke patients show unfavorable outcomes [7,8]. Hence, the early dynamic prediction of an unfavorable outcome is important in AIS patients treated with MT, and it can be assessed by the modified Rankin Scale (mRS) score [9].

Previous studies show that many prognostic factors are associated with the unfavorable outcome in AIS patients, such as physiologic factors [10–12] (e.g., age, sex, comorbidity), clinical factors [13,14] (e.g., NIHSS score, glycosylated hemoglobin, creatinine), and neuroimaging prognostic factors [15–17] (e.g., location of the occlusion, Alberta Stroke Program Early CT Score (ASPECTS)). However, these pretreatment prognostic factors still cannot well predict the outcome of AIS patients. This is due to some interventional factors that also affect the prognosis, such as procedure time, surgical technique, recanalization status, and so on [10,14,18].

Models used for dynamically predicting three-month unfavorable outcomes in patients with AIS treated with MT using clinically relevant preoperative and postoperative time point variables have not been developed. Currently, most are built based on preoperative variables, such as THRIVE [10], HIAT [14], GADIS [19], NAC [20], and IER-START [21], and several machine learning (ML) models [22–25], and these scores and models cannot be updated according to the changes in the patient's state and examination results over time. In clinical practice, these scores are intended to inform treatment, and they lack focus on determining functional outcomes after treatment. Additionally, the above scores have been validated externally, and the AUC range was from 0.680 to 0.838 in recent studies [10,20,21], indicating that there is still room to improve accuracy. In addition, these scores, based on the linear regression algorithm, have some limitations in addressing nonlinear problems between the variables in real-world applications. Therefore, developing a high accuracy model for a clinically relevant admission, postoperative 24-h, 3-day, with discharge time points for the dynamical prediction of unfavorable outcomes is desirable.

ML algorithms are in the realm of artificial intelligence. They have been proven to be an excellent approach in modeling multi-factor events in medical fields, such as brain-computer interface (BCI) technology [26–28], deep brain stimulation (DBS) application [29], and outcome prediction [30,31], because ML can contain a mass of data, extract subtle information, and summarize the knowledge gained on new unseen situations efficiently and automatically. At the same time, ML can also deal with the non-linear relationship between data. As can be surmised, applying an ML model at multiple time points after admission has allowed neurologists to better dynamically assess functional neurological outcomes and adjust therapeutic strategies.

To improve personalized stroke care, we developed and validated a dynamic ML model at multiple time points for the dynamic prediction of a three-month unfavorable outcome.

2. Materials and Methods

2.1. Study Population

We analyzed electronic health record data from three Chinese stroke units (Nanjing First Hospital, Changsha Central Hospital, and People's Hospital of Hunan Province) of adult patients (≥ 18 years of age at the time of surgery) between January 2014 and December 2018. Patients included in the analysis went through the following: (i) had an AIS treated with MT; (ii) had their three-month modified Rankin Scale (mRS) scores assessed; (iii) their National Institute of Health stroke scale (NIHSS) scores on admission, postoperative 24-h, 3-day, and discharge were recorded. Patients aged < 18 years or with an intracranial hemorrhage (ICH) on admission or interval from onset to treatment over 24 h were excluded.

In this study, the candidate variables for our model included demographics (age, sex, smoking), medical history (previous cerebral infarction, transient ischemic attack, previous cerebral hemorrhage, diabetes mellitus, hyperlipidemia, hypertension, atrial fibrillation and coronary heart disease), baseline data (diastolic pressure, systolic pressure, NIHSS

score on admission, Trial of ORG 10172 in Acute Stroke Treatment (TOAST) classification, anterior circulation stroke, posterior circulation stroke, and blood biochemical examination), as well as interventional characteristics (interval from groin puncture to recanalization, interval from onset to treatment, endovascular therapy and IV thrombolysis) and post interventional clinical characteristics (symptomatic intracranial hemorrhage (sICH), NIHSS score after 24-h, NIHSS score after 3-day and NIHSS score on discharge).

We estimated that the outcome was neurological–functional three months after the AIS with MT, as represented by an mRS score >2 (i.e., poor prognosis). In addition, the symptomatic intracranial hemorrhage (sICH) was exposed within 24 h after admission and evaluated by the Heidelberg Bleeding Classification [32]. The NIHSS score and three-month mRS were collected by outpatient services and telephone interviews with patients or their families.

2.2. Statistical Analyses

Variables with less than 20% missing values were taken in to generate the complete dataset. As the case may be, missing variables were imputed by mode for categorical variables and mean or median for continuous variables

A statistical analysis was performed using SPSS version 22.0 (IBM Corporation, Armonk, NY, USA), applying the number (percentage) as a measure of categorical variables and the medians (quartile) or means (standard deviation) as measures of continuous variables, as appropriate. In addition, comparisons in the baseline characteristics between favorable outcome and unfavorable outcome groups for continuous variables were performed using the Mann–Whitney U test or Student t test, depending on their normality of distribution, and Pearson’s test or Fisher’s exact test for categorical variables, depending on their sample amount.

2.3. Feature Selection

Feature selection further reduced redundancy and irrelevant variables, which is a critical step in improving a model’s stability. Concerning feature selection, we ranked the model features consistent with their importance (importance type = “gain,” the equal gain of divides that apply to the feature) and chose a group of features by eliminating the least important features. To avoid the false-negative result in the univariable analysis, all features with $p < 0.2$ in the univariable analysis or that were considered to have clinical significance features were taken into the feature selection. The features selected from the model could be divided into four subsets, judging by the time at which they were collected: admission, postoperative 24-h, postoperative 3-day, and discharge.

2.4. Model Development

Predictive ML models for the three-month outcome of AIS patients treated were developed by eXtreme gradient boosting (XGBoost). XGBoost is an ensemble learning classifier that utilizes a Gradient Boosting framework to solve supervised learning problems. XGBoost has been widely used in medical studies to predict or screen prognosis [33–35]. XGBoost also uses a trained prediction model to provide a regularized gradient enhancement and feature importance score, which can be applied to feature selection. For model development, it was performed in Python (version 3.7). We built machine-learning models with four sets of input variables corresponding to time points, and it was named the “Admission” model, “24-Hour” model, “3-Day” model, and “Discharge” model, respectively. The following procedure was carried out in each model: First, patients were randomly split into a train set and a test set according to an 8:2 scale procedure. Then, the train data were used to train and build the model. Meanwhile, we utilized a grid search algorithm and ten-fold cross-validation to optimize model hyperparameters. In a ten-fold cross-validation, the training set was randomly divided into ten equal folds, each fold was used in turn for the validation set, and the remaining nine folds were used for the

training set, which improved the generalization of the model and avoided overfitting. By obtaining the models, the model's performance was measured on the testing set.

2.5. Model Evaluation

We used the area under the receiver operating characteristic curve (AUC) to measure discrimination [36]. The calibration of the models was assessed using the Brier scores method (range: 0–1)—a model with a lower Brier score means better model calibration and discrimination [37]. The AUC comparisons between different models were evaluated with the DeLong test [38]. We further applied the testing set to THRIVE [10] and HIAT [14] scores to evaluate the dynamic model's preponderance in predictive performance.

3. Results

3.1. Study Population

The clinical characteristics of the unfavorable and favorable outcome groups are shown in Table 1. We screened 239 AIS patients treated with MT, patients for age <18 ($n = 1$; 0.4%) were excluded, and our final study sample included 238 patients overall, within the follow-up period, with a median age of 69 (IQR 59–78), and 72 patients (30.3%) were women. The median admission NIHSS score was 15 (IQR 10–20), while the median NIHSS score at postoperative 24-hour, 3-day, and discharge decreased to 13 (IQR 6–21), 12 (IQR 4–23), and 9 (IQR 2–20), respectively. Of the 238 patients, 156 (cumulative rate 65.5%) patients developed unfavorable outcomes in three months. and 61 patients (cumulative rate 25.6%) died (mRS scores = 6). The clinical outcome assessed by mRS scores at three months was 0 in 24 (10%), 1 in 25 (10.5%), 2 in 33 (13.8%), 3 in 34 (14.2%), 4 in 39 (16.3%), 5 in 24 (10%), and (10.5%), 2 in 33 (13.8%), 3 in 34 (14.2%), and 4 in 6 in 59 patients (24.7%). Compared with favorable outcome patients, patients with unfavorable outcomes were seven years older and had lower low-density lipoprotein, lower platelets, and higher creatinine. Patients with an unfavorable outcome also had longer intervals from groin puncture to recanalization and a higher NIHSS score.

Table 1. Clinical, demographic and laboratory data of study population stratified according to three-month favorable or unfavorable outcome after acute ischemic stroke in Chinese patients with mechanical thrombectomy.

Variables	Favorable Outcome (mRS 0–2)	Unfavorable Outcome (mRS 3–6)	<i>p</i> -Value
Patients, <i>n</i>	82	156	
Demographics			
Age, years, median (IQR)	65 (58–74)	72 (60–81)	0.001 *
Sex, <i>n</i> (%)			0.592
Male	59 (72)	107 (68.6)	
Female	23 (28)	49 (31.4)	
Smoking, <i>n</i> (%)	28 (34.1)	49 (31.4)	0.668
Medical history, <i>n</i> (%)			
Transient ischemic attack	0 (0)	3 (1.9)	0.110 *
Previous cerebral infarction	12 (14.6)	29 (18.6)	0.443
Previous cerebral hemorrhage	1 (1.2)	6 (3.8)	0.462
Diabetes mellitus	18 (22)	31 (19.9)	0.706
Hypertension	57 (69.5)	108 (69.5)	0.964
Hyperlipidemia	7 (8.5)	12 (7.7)	0.819
Coronary artery disease	15 (18.3)	44 (28.2)	0.092 *
Atrial fibrillation	25 (30.5)	58 (37.2)	0.303

Table 1. Cont.

Variables	Favorable Outcome (mRS 0–2)	Unfavorable Outcome (mRS 3–6)	p-Value
Baseline data			
NIHSS score on admission, median (IQR)	11 (7–16)	17 (13–22)	<0.0001 *
Systolic pressure, mmHg, median (IQR)	138 (124–155)	142 (129–160)	0.299
Diastolic pressure, mmHg, median (IQR)	83 (74–93)	86 (76–99)	0.154 *
INR, median (IQR)	0.98 (0.93–1.09)	1.02 (0.935–1.12)	0.092 *
HbA1c, mmol/L, median (IQR)	5.80 (5.50–6.53)	5.90 (5.50–6.50)	0.856
TC, mmol/L, median (IQR)	4.32 (3.43–4.98)	4.08 (3.43–4.83)	0.300
TG, mmol/L, median (IQR)	1.12 (0.82–1.68)	1.05 (0.76–1.47)	0.205
LDL, mmol/L, median (IQR)	2.66 (1.98–3.24)	2.23 (1.80–2.83)	0.019 *
FBG, mmol/L, median (IQR)	6.04 (5.08–7.35)	6.48 (5.60–7.99)	0.028 *
PLT, μ mol/L, median (IQR)	193.00 (150.75–235.50)	172.5 (143.00–212.50)	0.017 *
UA, μ mol/L, median (IQR)	284.90 (233.00–357.00)	313.50 (232.32–396.75)	0.125 *
HCY, μ mol/L, median (IQR)	12.46 (10.70–16.76)	13.18 (10.97–16.64)	0.433
Creatinine, μ mol/L, median (IQR)	67.00 (58.37–77.00)	78.00 (61.63–94.00)	0.003 *
Anterior circulation stroke, <i>n</i> (%)	60 (73.2)	126 (80.8)	0.178 *
Posterior circulation stroke, <i>n</i> (%)	22 (26.8)	30 (19.2)	0.178 *
TOAST classification, <i>n</i> (%)			0.119 *
Large artery atherosclerosis	47 (57.3)	69 (44.2)	
Cardioembolism	32 (39.0)	75 (48.1)	
Others	3 (3.7)	12 (7.7)	
Interventional characteristics			
Interval from groin puncture to recanalization, min, median (IQR)	60 (50–85)	81 (59–130)	0.004 *
Interval from onset to treatment, min, median (IQR)	290 (230–411)	280 (206–413)	0.240
Endovascular therapy, <i>n</i> (%)			0.144 *
Tirofiban	29 (35.4)	41 (26.3)	
No tirofiban	53 (64.6)	115 (73.7)	
IV thrombolysis, <i>n</i> (%)			0.806
No thrombolysis	45 (54.9)	83 (53.2)	
Thrombolysis	37 (45.1)	73 (46.8)	
Post interventional characteristics			
sICH, <i>n</i> (%)	0 (0)	18 (11.5)	0.001 *
NIHSS score after 24-hour, median (IQR)	5 (2–10)	17 (12–31)	<0.0001 *
NIHSS score after 3-day, median (IQR)	3 (2–7)	18 (10–34)	<0.0001 *
NIHSS score on discharge, median (IQR)	2 (1–3)	16 (8–34)	<0.0001 *

* Included into the feature selection ($p < 0.2$). mRS, modified Rankin Scale; NIHSS, National Institute of Health stroke scale; TOAST, Trial of ORG 10172 in Acute Stroke Treatment; INR, International normalized ratios; IQR, interquartile range; HbA1c, Glycated hemoglobin; TC, total cholesterol; TG, triglyceride; LDL, Low density lipoprotein; FBG, Fasting blood glucose; PLT, Platelet; UA, Uric Acid; HCY, Homocysteine; IV, Intravenous; sICH, symptomatic intracranial hemorrhage.

3.2. Feature Selection

According to the results (Supplementary Figure S1), we selected the top 10 variables that had a higher importance from the feature selection results as input variables of an “Admission” model. Similarly, the input variables of three other models that were the same as before the NIHSS score and were consistent with the time point were the highest-ranked feature in all models.

3.3. Model Performance

From the identified top-performing XGBoost-selected predictors, the prediction models for the three-month outcome of AIS patients treated with MT were created. Figure 1a,b presents the ROC curves of four models on the training set and testing set, respectively. Tables 2–5 present the confusion matrix of the performance of the “Admission”, “24-Hour”, “3-Day”, and “Discharge” models on the testing dataset. In addition, the Brier scores for unfavorable functional outcome prediction AIS on the testing dataset are shown in Figure 2.

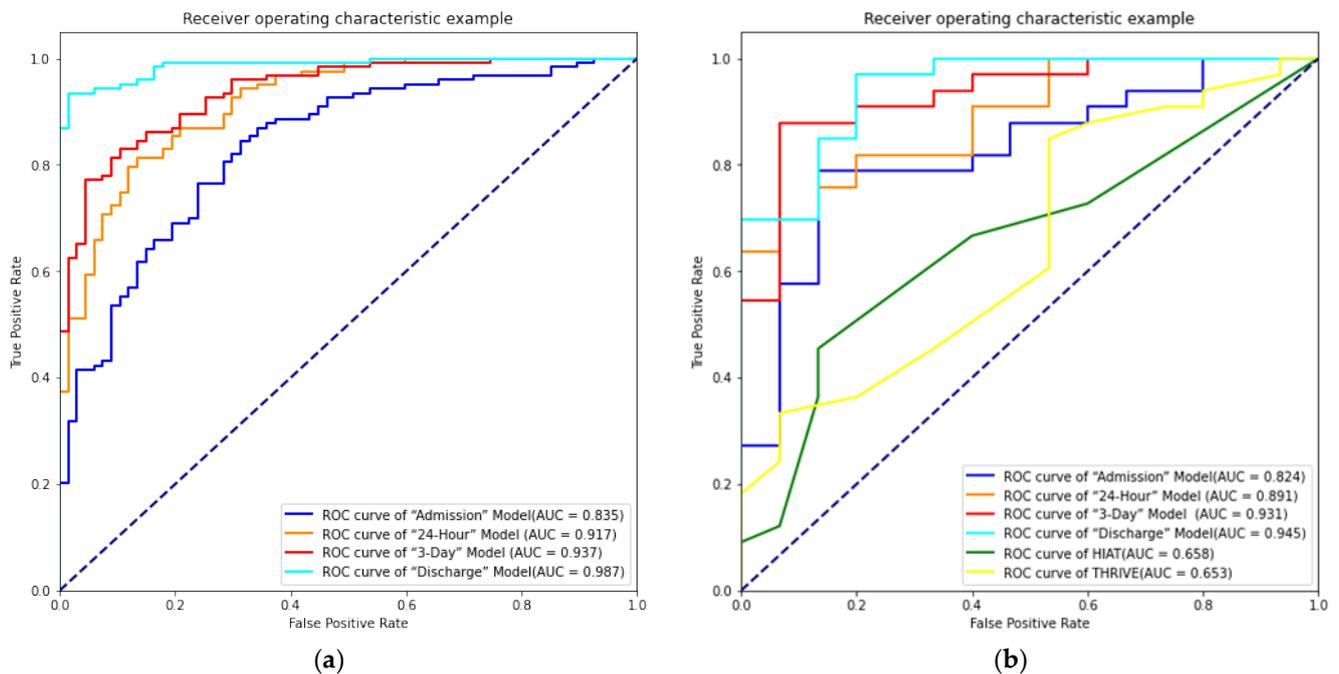


Figure 1. (a) The receiver operating characteristic curve (ROC) of four models on the training set; (b) the receiver operating characteristic curve (ROC) of four models and traditional scores on testing set. AUC, the area under curve; THRIVE, Totalled Health Risks in Vascular Events; HIAT, Houston Intra-arterial Recanalization Therapy.

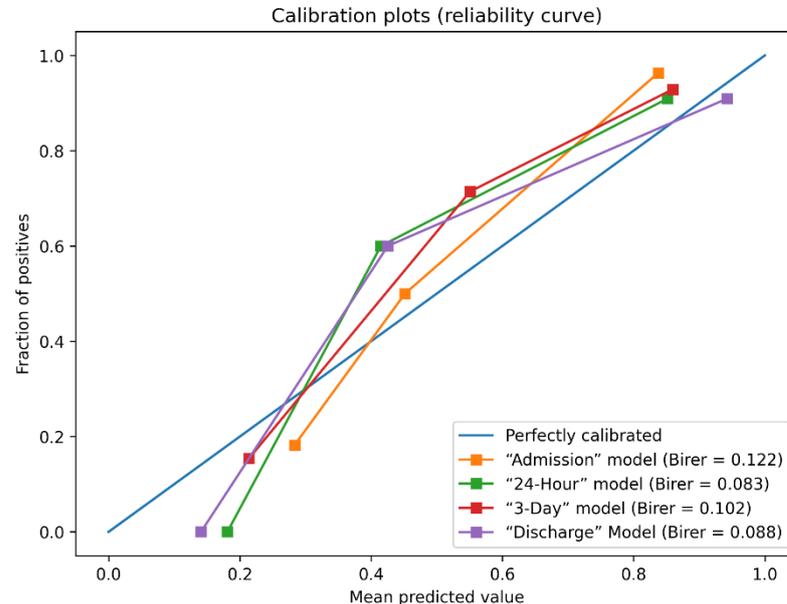


Figure 2. The calibration curves and the Brier score of four models on the testing set.

3.3.1. "Admission" Model

On the training data, the "Admission" model achieved an AUC of 0.835. On the testing set, the model achieved an AUC of 0.824, and the sensitivity and specificity were 0.848 and 0.533. Among the 33 cases with an unfavorable outcome, 28 cases were correctly predicted by the model. The model incorrectly predicted five cases as favorable outcomes, leading to a recall rate of 70%. A total of 75.0% of cases were correctly predicted, contributing to an error rate of 25.0%. The hyper-parameters applied in the "Admission" model are shown in Supplementary Table S1.

3.3.2. “24-Hour” Model

On the training data, the “24-Hour” model achieved an AUC of 0.917. On the testing set, the model achieved an AUC of 0.891, and the sensitivity and specificity were 0.788 and 0.800. Among the 33 cases with an unfavorable outcome, 26 cases were correctly predicted by the model. The model incorrectly predicted seven cases as favorable outcomes, leading to a recall rate of 78.8%. A total of 79.2 % of cases were correctly predicted, corresponding to an error rate of 14.6%. The hyper-parameters applied in the “24-Hour” model are shown in Supplementary Table S1.

3.3.3. The “3-Day” Model

On the training data, the “3-Day” model achieved an AUC of 0.937. On the testing set, the model achieved an AUC of 0.931, and the sensitivity and specificity were 0.758 and 0.933. Among the 33 cases with an unfavorable outcome, 25 cases were correctly predicted by the model. The model incorrectly predicted eight cases as favorable outcomes, leading to a recall rate of 75.8%. A total of 81.2% of cases were correctly predicted, corresponding to an error rate of 18.8%. The hyper-parameters applied in the “3-Day” model are shown in Supplementary Table S1.

3.3.4. “Discharge” Model

On the training data, the “Discharge” model achieved an AUC of 0.987. On the testing set, the model achieved an AUC of 0.945, and the sensitivity and specificity were 0.909 and 0.800. Among the 33 cases with an unfavorable outcome, 30 cases were correctly predicted by the model. The model incorrectly predicted three cases as favorable outcomes, leading to a recall rate of 90.9%. A total of 87.5% of cases were correctly predicted, corresponding to an error rate of 12.5%. The hyper-parameters applied in the “Discharge” model are shown in Supplementary Table S1.

Table 2. Confusion matrix for the “Admission” Model.

Testing Data				Statistical Analysis	
True Predicted	0	1	Total	Accuracy	0.750
0	8	7	15	Precision	0.800
1	5	28	33	Sensitivity	0.848
Total	13	35	48	Specificity	0.533
				AUC	0.824

AUC, the area under curve.

Table 3. Confusion matrix for the “24-H” Model.

Testing Data				Statistical Analysis	
True Predicted	0	1	Total	Accuracy	0.792
0	12	3	15	Precision	0.897
1	7	26	33	Sensitivity	0.788
Total	19	29	48	Specificity	0.800
				AUC	0.891

AUC, the area under curve.

Table 4. Confusion matrix for the “3-Day” Model.

Testing Data				Statistical Analysis	
True Predicted	0	1	Total	Accuracy	0.812
0	14	1	15	Precision	0.962
1	8	25	33	Sensitivity	0.758
Total	22	26	48	Specificity	0.933
				AUC	0.931

AUC, the area under curve.

Table 5. Confusion matrix for the “Discharge” Model.

Testing Data				Statistical Analysis	
True Predicted	0	1	Total	Accuracy	0.875
0	12	3	15	Precision	0.909
1	3	30	33	Sensitivity	0.909
Total	17	31	48	Specificity	0.800
				AUC	0.945

3.4. Model Comparison

Among all prediction models, we found that the “Discharge” model had the highest AUC (0.970; 95%CI 0.901–0.991) and the lowest Brier scores (0.081), and the “Admission” model had the lowest AUC (0.824; 95%CI 0.715–0.899) and the highest Brier scores (0.122). The results of the DeLong test indicated that the “Admission” and “24-Hour” models were statistically different from the AUC of “Discharge” models ($p < 0.05$) (Supplementary Table S2), and the AUC of the four models gradually increased from admission to discharge on the testing set. Meanwhile, the Brier scores decreased from admission to discharge on the testing set. The dynamic model exhibited a good predictive ability on the testing set and improved over time. In addition, the AUCs of the dynamic model were significantly superior to THRIVE [10] and HIAT [14] scores (Figure 1b, Supplementary Table S2).

4. Discussion

To the best of our knowledge, this is the first attempt to construct four XGBoost ML models (i.e., a preoperative, postoperative 24-h, 3-day, and discharge model) to dynamically predict the probability of a three-month unfavorable outcome in Chinese patients with AIS treated with MT. Although traditional prediction scores [10,14,19–21] and ML models [22–25] offer practical guidance and are effective stroke prognosis tools for population-based comparisons, they predict the outcome of AIS patients undergoing MT by a static nature, and their clinical use is limited. Instead, in our study, we combined preoperative and intraoperative data with time-relevant postoperative variables to dynamically predict the three-month outcome of MT for AIS patients. the “Admission” model could be used to predict the clinical outcome before MT and support the decision to perform MT, and the “24-Hour”, “3-Day”, and “Discharge” models would further improve the predictive accuracy of the clinical outcome after MT and adjust therapeutic strategies in time. From admission to discharge, the physiology or pathology will change along with treatment, so it is necessary to consider all possible factors that occurred until a specific time point and provide an outcome prediction in real time that could guide better decisions for treatment [39]. With that in mind, we proposed a method of risk-based dynamic monitoring strategies and targeted the introduction of therapy neurologists when discussing prognosis with AIS patients and their families.

One strength of our study lies in our models showing good discrimination ability compared to previous scoring systems (THRIVE and HIAT scores) by the AUC of models at different time points. In the “Admission” model, it included preoperative factors, and the AUC of the model was 0.824, while in the “24-Hour”, “3-Day”, and “Discharge” models, they included intraoperative and postoperative factors, and the AUC of the model was 0.891–0.945, while THRIVE and HIAT scores with an AUC of 0.653 and 0.658, respectively. The recent single-center cohort study combined individual data of 246 patients from Heidelberg University Hospital with multiple clinical characteristics to develop unfavorable outcome prediction models [23]. The baseline model included clinical and conventional imaging characteristics on admission to reach an AUC of 0.740 to predict unfavorable outcomes in internal validation. Further addition of CT-perfusion, angiographic, and post-interventional characteristics achieved an AUC of 0.856, still below the predictive performance of our model [23]. These four prediction models are beneficial to reducing the

risk of postoperative three-month unfavorable outcome and may have good application prospects. In addition, however, we also found that adding intraoperative data with time-relevant postoperative variables (i.e., NIHSS score in postoperative 24-h, 3-day, and discharge time points) would modestly improve the predictive performance of outcome models, which supported the theory that the admission NIHSS scores are inferior to postoperative 24-h NIHSS, 3-day NIHSS, and discharge NIHSS score in predicting post-stroke functional outcome [40].

There were several additional strong points of this study. First, we derived unfavorable outcome models using state-of-the-art ML methods, which have recently been applied in the development of neurological function outcome prediction models in the domain of neurological disease [22–25]. However, our modeling strategy differed from the conventional single modeling strategy, with multiple models predicting the unfavorable outcome for more precise forecasting. Second, the prediction models were developed using multicenter registration data. A wide range of data sources may reduce non-sampling errors and enhance the reliability of models. Third, it is notable that, in most studies, the importance of the features were first assessed using univariate analysis and based on prior knowledge, then the significant features were engaged in the ML model. In the end, they calculated the variable importance of variables in ML models. This variable selection method is called the filter method and has been widely used. However, it may introduce too many variables to fit models well and may bring the collinearity of variables [23]. Thus, the rank of feature importance in the model was introduced to feature selection, reducing over-fitting and avoiding variables collinearity.

In our study, the variables selected were inexpensive to obtain and commonly available. Therefore, we found it to be similar to former prediction models. The most important features among the four XGBoost models were the NIHSS score, which achieved a dominant position in all variables, and confirmed NIHSS was relatively high in the three-month outcome of AIS patients [40]. Furthermore, a previous study said the NIHSS score after endovascular surgery had a better discrimination in predicting three-month functional outcomes than the pretreatment NIHSS score [41]. Therefore, the “24-Hour”, “3-Day”, and “Discharge” models, which incorporated both preoperative NIHSS score and postoperative NIHSS score could be offered modest improvements in predicting a three-month unfavorable outcome.

It is consistent with previous reports, which incorporated the discharge NIHSS score with good predictive accuracy about the functional outcome. For example, Miguel Monteiro et al. [42] constructed machine learning prognostic models for AIS patients, and the models revealed that incorporating the discharge NIHSS score had the best result in predicting the 90-day functional outcome. Nevertheless, in this paper, all patients were treated using Recombinant Tissue Plasminogen Activator (rtPA), and the literature included many variables, which was not conducive to predicting clinical promotion. Furthermore, our models not only contributed to accurately predicting patients who would and would not benefit from MT, but also they would benefit from improving the postoperative accuracy of AIS patients with MT.

Nevertheless, our study has several limitations. First, this study was based on a retrospective data set, and there may be invisible disturbing factors contributing to potential biases in the results. Thus, prospective studies are necessary and validated in an independent external cohort to reconfirm these results in future work. Second, the dynamic models screened out patients with poor outcomes but did not suggest who would benefit from preventative therapy. It still depends on clinicians to decide whether to use therapeutic interventions. However, as the disease continues, clinical indicators may change. The early and precise prediction of poor outcomes may require more time for neurologists to reform treatment strategies and reduce complications and mortality. Third, the size of our study population and clinical characteristics were limited, while current ML models lack an underlying causal structure: their predictions are entirely based on what humans have done in the past. Hence, the future will expand the sample size for further research.

5. Conclusions

This is the first dynamic predictive model of a three-month unfavorable outcome for Chinese AIS patients with MT that was developed and validated at admission, postoperative 24-h, 3-day, and discharge time points, and these models showed good discrimination and calibration. In terms of AUCs, the range can be up to 0.945, depending on when the prediction was made. Furthermore, compared with previous scoring (THRIVE and HIAT scores), the dynamic risk prediction had a better performance. The preoperative model could be used to predict the clinical outcome before MT and support the decision to perform MT, and the postoperative models would further improve the predictive accuracy of the clinical outcome after MT and adjust therapeutic strategies in timely.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/brainsci12070938/s1>, Table S1: Hyper-parameters of each XGBoost model; Table S2: The *p*-values of pairwise comparisons of AUCs on the testing set for different models with the Delong test; Figure S1: (a) Top 10 features selected using eXtreme gradient boosting (XGBoost) and the corresponding variable importance score in “Admission” model; (b) Top 10 features selected using XGBoost and the corresponding variable importance score in “24-Hour” model; (c) Top 10 features selected using XGBoost and the corresponding variable importance score in “3-Day” model; (d) Top 10 features selected using XGBoost and the corresponding variable importance score in “Discharge” model. X-axis indicates the importance score which is the relative number of a variable that is used to distribute the data, Y-axis indicates the top 10 weighted variables.

Author Contributions: Conceptualization, Y.H. and T.Y.; methodology, Y.H.; software, T.Y.; formal analysis, Y.H.; data curation, J.Z. (Juan Zhang) and X.W.; writing—original draft preparation, Y.H. and T.Y.; writing—review and editing, X.C.; visualization, N.C.; supervision, J.Z. (Junshan Zhou); project administration, F.J., J.Z. (Junrong Zhu) and J.Z. (Jianjun Zou). All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (Grant Number 82173899) and Jiangsu Pharmaceutical Association (Grant Number H202108, A2021024).

Institutional Review Board Statement: The study was conducted in accordance with the Nanjing First Hospital, the First Affiliated Hospital (People’s Hospital of Hunan Province), and Changsha Central Hospital, approved by the Ethics Committee of Nanjing First Hospital and the document number of the ethics committee’s approval: ChiCTR-OCH-14004382.

Informed Consent Statement: Patient consent was waived because this is a retrospective study.

Data Availability Statement: The datasets we used in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wu, S.; Wu, B.; Liu, M.; Chen, Z.; Wang, W.; Anderson, C.S.; Sandercock, P.; Wang, Y.; Huang, Y.; Cui, L.; et al. Stroke in China: Advances and challenges in epidemiology, prevention, and management. *Lancet Neurol.* **2019**, *18*, 394–405. [[CrossRef](#)]
2. Benjamin, E.J.; Virani, S.S.; Callaway, C.W.; Chamberlain, A.M.; Chang, A.R.; Cheng, S.; Chiuve, S.E.; Cushman, M.; Delling, F.N.; Deo, R.; et al. Heart Disease and Stroke Statistics-2018 Update: A Report from the American Heart Association. *Circulation* **2018**, *137*, e67–e492. [[CrossRef](#)] [[PubMed](#)]
3. Saver, J.L.; Goyal, M.; van der Lugt, A.; Menon, B.K.; Majoie, C.B.L.M.; Dippel, D.W.; Campbell, B.C.; Nogueira, R.G.; Demchuk, A.M.; Tomasello, A.; et al. Time to Treatment with Endovascular Thrombectomy and Outcomes from Ischemic Stroke: A Meta-analysis. *JAMA* **2016**, *316*, 1279–1288. [[CrossRef](#)] [[PubMed](#)]
4. Schonewille, W.J. BEST evidence on mechanical thrombectomy for patients with vertebrobasilar occlusion. *Lancet Neurol.* **2020**, *19*, 102–103. [[CrossRef](#)]
5. Powers, W.J.; Rabinstein, A.A.; Ackerson, T.; Adeoye, O.M.; Bambakidis, N.C.; Becker, K.; Biller, J.; Brown, M.; Demaerschalk, B.M.; Hoh, B.; et al. 2018 Guidelines for the Early Management of Patients With Acute Ischemic Stroke: A Guideline for Healthcare Professionals from the American Heart Association/American Stroke Association. *Stroke* **2018**, *49*, e46–e110. [[CrossRef](#)]
6. Jovin, T.G.; Chamorro, A.; Cobo, E.; De Miquel, M.A.; Molina, C.A.; Rovira, A.; Román, L.S.; Serena, J.; Abilleira, S.; Ribo, M.; et al. Thrombectomy within 8 hours after symptom onset in ischemic stroke. *N. Engl. J. Med.* **2015**, *372*, 2296–2306. [[CrossRef](#)]

7. Singer, O.C.; Berkefeld, J.; Nolte, C.; Bohner, G.; Haring, H.-P.; Trenkler, J.; Gröschel, K.; Müller-Forell, W.; Niederkorn, K.; Deutschmann, H.; et al. Mechanical recanalization in basilar artery occlusion: The ENDOSTROKE study. *Ann. Neurol.* **2015**, *77*, 415–424. [[CrossRef](#)]
8. Saini, V.; Guada, L.; Yavagal, D.R. Global Epidemiology of Stroke and Access to Acute Ischemic Stroke Interventions. *Neurology* **2021**, *97* (Suppl. 2), S6–S16. [[CrossRef](#)]
9. van Swieten, J.C.; Koudstaal, P.J.; Visser, M.C.; Schouten, H.J.; van Gijn, J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* **1988**, *19*, 604–607. [[CrossRef](#)]
10. Flint, A.C.; Cullen, S.P.; Faigeles, B.S.; Rao, V.A. Predicting long-term outcome after endovascular stroke treatment: The totaled health risks in vascular events score. *AJNR Am. J. Neuroradiol.* **2010**, *31*, 1192–1196. [[CrossRef](#)]
11. Almekhlafi, M.; Davalos, A.; Bonafe, A.; Chapot, R.; Gralla, J.; Pereira, V.; Goyal, M. Impact of age and baseline NIHSS scores on clinical outcomes in the mechanical thrombectomy using solitaire FR in acute ischemic stroke study. *AJNR Am. J. Neuroradiol.* **2014**, *35*, 1337–1340. [[CrossRef](#)] [[PubMed](#)]
12. Hametner, C.; Kellert, L.; Ringleb, P.A. Impact of sex in stroke thrombolysis: A coarsened exact matching study. *BMC Neurol.* **2015**, *15*, 10. [[CrossRef](#)] [[PubMed](#)]
13. Linfante, I.; Starosciak, A.K.; Walker, G.R.; Dabus, G.; Castonguay, A.C.; Gupta, R.; Sun, C.-H.J.; Martin, C.; Holloway, W.E.; Mueller-Kronast, N.; et al. Predictors of poor outcome despite recanalization: A multiple regression analysis of the NASA registry. *J. Neurointerv. Surg.* **2016**, *8*, 224–229. [[CrossRef](#)]
14. Halleivi, H.; Barreto, A.D.; Liebeskind, D.S.; Morales, M.M.; Martin-Schild, S.B.; Abraham, A.T.; Gadia, J.; Saver, J.L.; Grotta, J.C.; Savitz, S.I.; et al. Identifying patients at high risk for poor outcome after intra-arterial therapy for acute ischemic stroke. *Stroke* **2009**, *40*, 1780–1785. [[CrossRef](#)]
15. Liggins, J.T.P.; Yoo, A.J.; Mishra, N.; Wheeler, H.M.; Straka, M.; Leslie-Mazwi, T.M.; Chaudhry, Z.A.; Kemp, S.; Mlynash, M.; Bammer, R.; et al. A score based on age and DWI volume predicts poor outcome following endovascular treatment for acute ischemic stroke. *Int. J. Stroke* **2015**, *10*, 705–709. [[CrossRef](#)] [[PubMed](#)]
16. Lansberg, M.G.; Christensen, S.; Kemp, S.; Mlynash, M.; Mishra, N.; Federau, C.; Tsai, J.P.; Kim, S.; Nogueira, R.G.; Jovin, T.; et al. Computed tomographic perfusion to Predict Response to Recanalization in ischemic stroke. *Ann. Neurol.* **2017**, *81*, 849–856. [[CrossRef](#)] [[PubMed](#)]
17. Elijovich, L.; Goyal, N.; Mainali, S.; Hoit, D.; Arthur, A.S.; Whitehead, M.; Choudhri, A.F. CTA collateral score predicts infarct volume and clinical outcome after endovascular therapy for acute ischemic stroke: A retrospective chart review. *J. Neurointerv. Surg.* **2016**, *8*, 559–562. [[CrossRef](#)]
18. Rangaraju, S.; Liggins, J.T.; Aghaebrahim, A.; Streib, C.; Sun, C.; Gupta, R.; Nogueira, R.; Frankel, M.; Mlynash, M.; Lansberg, M.; et al. Pittsburgh outcomes after stroke thrombectomy score predicts outcomes after endovascular therapy for anterior circulation large vessel occlusions. *Stroke* **2014**, *45*, 2298–2304. [[CrossRef](#)]
19. O'Connor, K.P.; Hathidara, M.Y.; Danala, G.; Xu, C.; McCoy, T.M.; Sidorov, E.V.; Zheng, B.; Bohnstedt, B.N.; Ray, B. Corrigendum to: Predicting Clinical Outcome After Mechanical Thrombectomy: The GADIS (Gender, Age, Diabetes Mellitus History, Infarct Volume, and Current Smoker) Score [World Neurosurgery 134 (2020) e1130-e1142]. *World Neurosurg.* **2020**, *138*, 606. [[CrossRef](#)]
20. Li, X.; Zou, Y.; Hu, J.; Li, X.M.; Huang, C.P.; Shan, Y.J.; Nyame, L.; Zhao, Z.; Sun, C.; Ibrahim, M.; et al. A NAC nomogram to predict the probability of three-month unfavorable outcome in Chinese acute ischemic stroke patients treated with mechanical thrombectomy. *Int. J. Neurosci.* **2021**, *131*, 163–169. [[CrossRef](#)]
21. Cappellari, M.; Mangiafico, S.; Saia, V.; Pracucci, G.; Nappini, S.; Nencini, P.; Konda, D.; Sallustio, F.; Vallone, S.; Zini, A.; et al. IER-START nomogram for prediction of three-month unfavorable outcome after thrombectomy for stroke. *Int. J. Stroke* **2020**, *15*, 412–420. [[CrossRef](#)] [[PubMed](#)]
22. Chiu, I.M.; Zeng, W.H.; Cheng, C.Y.; Chen, S.H.; Lin, C.R. Using a Multiclass Machine Learning Model to Predict the Outcome of Acute Ischemic Stroke Requiring Reperfusion Therapy. *Diagnostics* **2021**, *11*, 80. [[CrossRef](#)] [[PubMed](#)]
23. Nishi, H.; Oishi, N.; Ishii, A.; Ono, I.; Ogura, T.; Sunohara, T.; Chihara, H.; Fukumitsu, R.; Okawa, M.; Yamana, N.; et al. Predicting Clinical Outcomes of Large Vessel Occlusion before Mechanical Thrombectomy Using Machine Learning. *Stroke* **2019**, *50*, 2379–2388. [[CrossRef](#)] [[PubMed](#)]
24. Van Os, H.J.A.; Ramos, L.A.; Hilbert, A.; van Leeuwen, M.; Van Walderveen, M.A.A.; Kruijff, N.D.; Dippel, D.W.J.; Steyerberg, E.W.; Van Der Schaaf, I.C.; Lingsma, H.F.; et al. Predicting Outcome of Endovascular Treatment for Acute Ischemic Stroke: Potential Value of Machine Learning Algorithms. *Front. Neurol.* **2018**, *9*, 784. [[CrossRef](#)]
25. Asadi, H.; Dowling, R.; Yan, B.; Mitchell, P. Machine learning for outcome prediction of acute ischemic stroke post intra-arterial therapy. *PLoS ONE* **2014**, *9*, e88225. [[CrossRef](#)]
26. Roy, A.M. A multi-scale fusion CNN model based on adaptive transfer learning for multi-class MI-classification in BCI system. *BioRxiv* **2022**. [[CrossRef](#)]
27. Sharma, R.; Kim, M.; Gupta, A. Motor imagery classification in brain-machine interface with machine learning algorithms: Classical approach to multi-layer perceptron model. *Biomed. Signal. Process.* **2022**, *71*, 103101. [[CrossRef](#)]
28. Kapgate, D. Efficient Quadcopter Flight Control Using Hybrid SSVEP + P300 Visual Brain Computer Interface. *Int. J. Hum-Comput Int.* **2022**, *38*, 42–52. [[CrossRef](#)]
29. Watts, J.; Khojandi, A.; Shylo, O.; Ramdhani, R.A. Machine Learning's Application in Deep Brain Stimulation for Parkinson's Disease: A Review. *Brain Sci.* **2020**, *10*, 809. [[CrossRef](#)]

30. Thorsen-Meyer, H.C.; Nielsen, A.B.; Nielsen, A.P.; Kaas-Hansen, B.S.; Toft, P.; Schierbeck, J.; Strøm, T.; Chmura, P.J.; Heimann, M.; Dybdahl, L.; et al. Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: A retrospective study of high-frequency data in electronic patient records. *Lancet Digit Health* **2020**, *2*, e179–e191. [[CrossRef](#)]
31. Khera, R.; Haimovich, J.; Hurley, N.C.; McNamara, R.; Spertus, J.A.; Desai, N.; Rumsfeld, J.S.; Masoudi, F.A.; Huang, C.; Normand, S.-L.; et al. Use of Machine Learning Models to Predict Death after Acute Myocardial Infarction. *JAMA Cardiol.* **2021**, *6*, 633–641. [[CrossRef](#)] [[PubMed](#)]
32. von Kummer, R.; Broderick, J.P.; Campbell, B.C.V.; Demchuk, A.; Goyal, M.; Hill, M.D.; Treurniet, K.M.; Majoie, C.B.L.M.; Marquering, H.A.; Mazya, M.V.; et al. The Heidelberg Bleeding Classification: Classification of Bleeding Events after Ischemic Stroke and Reperfusion Therapy. *Stroke* **2015**, *46*, 2981–2986. [[CrossRef](#)] [[PubMed](#)]
33. Xu, Y.; Han, D.; Huang, T.; Zhang, X.; Lu, H.; Shen, S.; Lyu, J.; Wang, H. Predicting ICU Mortality in Rheumatic Heart Disease: Comparison of XGBoost and Logistic Regression. *Front. Cardiovasc. Med.* **2022**, *9*, 847206. [[CrossRef](#)] [[PubMed](#)]
34. Nam, S.M.; Peterson, T.A.; Seo, K.Y.; Han, H.W.; Kang, J.I. Discovery of Depression-Associated Factors from a Nationwide Population-Based Survey: Epidemiological Study Using Machine Learning and Network Analysis. *J. Med. Internet Res.* **2021**, *23*, e27344. [[CrossRef](#)]
35. Hou, N.; Li, M.; He, L.; Xie, B.; Wang, L.; Zhang, R.; Yu, Y.; Sun, X.; Pan, Z.; Wang, K. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: A machine learning approach using XGboost. *J. Transl. Med.* **2020**, *18*, 462. [[CrossRef](#)]
36. Hanley, J.A.; McNeil, B.J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **1982**, *143*, 29–36. [[CrossRef](#)]
37. Selvarajah, S.; Kaur, G.; Haniff, J.; Cheong, K.C.; Hiong, T.G.; van der Graaf, Y.; Bots, M.L. Comparison of the Framingham Risk Score, SCORE and WHO/ISH cardiovascular risk prediction models in an Asian population. *Int. J. Cardiol.* **2014**, *176*, 211–218. [[CrossRef](#)]
38. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* **1988**, *44*, 837–845. [[CrossRef](#)]
39. Maldaner, N.; Zeintlberger, A.M.; Sosnova, M.; Goldberg, J.; Fung, C.; Bervini, D.; May, A.; Bijlenga, P.; Schaller, K.; Roethlisberger, M.; et al. Development of a Complication- and Treatment-Aware Prediction Model for Favorable Functional Outcome in Aneurysmal Subarachnoid Hemorrhage Based on Machine Learning. *Neurosurgery* **2021**, *88*, E150–E157. [[CrossRef](#)]
40. Reznik, M.E.; Yaghi, S.; Jayaraman, M.V.; McTaggart, R.A.; Hemendinger, M.; Mac Grory, B.C.; Burton, T.M.; Cutting, S.M.; Siket, M.S.; Madsen, T.E.; et al. Baseline NIH Stroke Scale is an inferior predictor of functional outcome in the era of acute stroke intervention. *Int. J. Stroke* **2018**, *13*, 806–810. [[CrossRef](#)]
41. Jeong, H.-G.; Kim, B.J.; Choi, J.C.; Hong, K.-S.; Yang, M.H.; Jung, C.; Han, M.-K.; Bae, H.-J. Posttreatment National Institutes of Health Stroke Scale Is Superior to the Initial Score or Thrombolysis in Cerebral Ischemia for 3-Month Outcome. *Stroke* **2018**, *49*, 938–944. [[CrossRef](#)] [[PubMed](#)]
42. Monteiro, M.; Fonseca, A.C.; Freitas, A.T.; e Melo, T.P.; Francisco, A.P.; Ferro, J.M.; Oliveira, A.L. Using Machine Learning to Improve the Prediction of Functional Outcome in Ischemic Stroke Patients. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2018**, *15*, 1953–1959. [[CrossRef](#)] [[PubMed](#)]