Finding predictors of stroke survival. Evidence from the Perm region.

Kulikova Sofya[1*], Polyakova Irina[1,2†], Kuzmicheva Elizaveta[1†], Kulesh Alexey[3,4], Chazov Stanislav[3,4]

[1]Center for Cognitive Neuroscience, HSE University, blvd. Gagarina, Perm, 614060, Russia.
[2]The International Laboratory of Intangible-driven Economy, HSE University, blvd. Gagarina, Perm, 614060, Russia.
[3]Perm City Clinical Hospital №4, Kim St, Perm, 614107, Russia.
[4]Ministry of Health of Russia, Acad. E.A. Vagner Perm State Medical University, Petropavlovskaya St, Perm, 614990, Russia.

*Corresponding author(s). E-mail(s): SPKulikova@hse.ru;
†These authors contributed equally to this work.

Abstract

Predicting the outcome of the ischemic stroke is a complicated task since its mortality and disability rates depend on multiple factors, including age, sex, stroke type and severity, various comorbidities. In addition, survival rate also varies across the countries, reflecting such factors as genetics and differences in the organization of the national health care systems. This study aims at finding predictors of a 1-year survival after ischaemic stroke in a sample of 254 patients from the Perm region of Russia. Each patient was described with 75 parameters, registered following routine clinical examinations and including information about the stroke subtype and severity, lesion size and location, stroke-associated deficits, comorbidities, smoking, results of the blood tests and cardiac characteristics, measure using echocardiography. The final prediction model was based on a logistic regression and included the 6 most informative characteristics: age, NIHSS score at discharge, hemoglobin level, involvement of the anterior choroidal artery, recurrent stroke within the following year and the cardioembolic origin. The proposed model has reached the accuracy level of 84% in a 10-fold cross-validation. Further multicenter studies are needes to test how well the revealed predictors could be extrapolated to other clinical settings and samples.

Keywords: ischaemic stroke, prediction models, outcome, MRI, NIHSS, blood tests

1 Introduction

Stroke is a severe neurological disorder, characterized by a high-level of mortality and chronic disabilities among survivals [1]. Taking into account the tremendous economical and social impact of stroke, finding predictors of stroke outcome is an important task of clinical decision-making. However, the etiology of stroke is highly variable, including a significant portion of cryptogenic cases with undetermined causes. This makes predicting stroke outcome a complicated task. Recent models predicting unfavourable stroke outcome include a 30-day mortality risk prediction model [2], a nomogram for predicting death within 6 months of stroke [3], the PLAN score [4], the IScore [5], the Ischaemic Stroke Survival Score [6], various scores and models for predicting survival up to 1 year [7–9], including a dynamic prediction model suggested by Huang et al. [10], and a 10-year mortality model proposed by Szlachetka et al. [11]. Despite the differences in the proposed models, there is a common evidence that certain factors may be considered as promising predictors of stroke survival, including age [4, 10, 11], sex [5, 10, 11], stroke severity [2, 5, 10], stroke type [4, 7, 10] and various comorbidities [4, 5, 7], such as heart-related diseases, diabetes, renal disfunctions, etc. However, the exact combination of these factors in the proposed prediction models varies across the studies (a detailed comparison of these studies is given in Tab. 1). This fact may reflect not only the differences in the used prediction models but also the different mortality levels from stroke across the countries [12], which may stem from multiple factors such as race [13], lifestyle [14], genetics [15, 16] and differences in the organization of the national health care systems w.r.t. stroke treatment [17]. Taking into consideration these factors and the fact that most of the predicting models were designed based on Western populations, the goal of the present work was to identify factors that could be used to predict mortality rate 1 year after stroke in a Russian population from the Ural region.

2 Materials and methods

2.1 Clinical sample

This study represents analysis of 254 patients (mean age - 65.8±10.2 years, min - 28, max - 90; 140 males / 114 females) with an acute ischemic stroke, who were followed for at least one year after their initial hospitalization to the Neurological Department of the Perm City Clinical Hospital №4. The inclusion criteria were the following:

1. An acute ischemic stroke at the time of admission, confirmed by a clinical specialist based on both CT and MRI scans;
2. a complete clinical, laboratory, and instrumental examination according to the current health care guidelines and protocols;
3. a given written informed consent.

During the following year 54 patients died, which corresponds to a 21% mortality rate. Following the routine clinical examinations during the hospitalization, a comprehensive dataset of 75 clinical parameters was collected for each individual. All

Table 1 Previously reported factors predicting stroke survival

First author	Wang	Sha	O'Donnell	Saposnik	Williams	Anderson	Solberg	Wang	Huang	Szlachetka
pub. year	(2022)	(2021)	(2012)	(2011)	(2000)	(1994)	(2007)	(2003)	(2020)	(2022)
sample (N)	488497	210	4943	12262	453	492	800	440	4315	10366
sample origin	UK	China	Canada	Canada	US	Australia	Norway	Australia	UK	UK
recording period	2013-2019	2013-2019	2003-2008	2003-2008	1993-1999	1989-1990	1993-1998	1995-1997	1995-2016	2003-2016
prediction period	1 months	6 months	1 year	1 year	1 year	1 year	1 year	1 year	1 year	10 years
sex	+	ns.	ns.	++	ns.	ns.	ns.	ns.	++	++
age	++	+	++	++	++	ns.	++	ns.	++	++
stroke type	++[1]	ns.[1]	ns.	++	NA	++	++[1]	++[6]	++	++
stroke severity	++	NA	NA	++	++	++	++	NA	++	++
consciousness level	++[2]	NA	++	NA	NA	++	NA	++	++	NA
dependence	NA	++[5]	++	++	++	++	NA	NA	++	++
neurological deficit	NA	NA	++	NA	++	++	++	++[7]	++	++
heart diseases	+[3]	+	++[3]	++	NA	++	++	++	NA	++
hypertension	ns.	+	ns.	NA	NA	ns.	+	ns.	NA	++
white blood cells	NA	+	ns.	NA	NA	NA	NA	ns.	NA	++
PLR[4]	NA	++	NA	NA	NA	NA	NA	NA	NA	NA
serum albumin	NA	++	NA	NA	NA	NA	NA	NA	NA	NA
serum D-dimer	NA	+	NA	NA	NA	NA	NA	NA	NA	NA
smoking	NA	+	ns.	++	ns.	ns.	NA	NA	NA	NA
nutritional status	NA	+	NA	NA	NA	NA	NA	NA	NA	NA
diabetes	ns.	ns.	ns.	+	NA	ns.	++	++	NA	ns.
glucose level	NA	NA	ns.	++	NA	NA	+	ns.	NA	NA
hypertermia	NA	NA	ns.	NA	NA	NA	ns.	++	NA	NA

'ns.' is indicated if the factor was included in the study but considered as non-significant for predicting stroke outcome; 'NA' is indicated if the factor was not considered or reported in the study; '+' stands if the factor has shown some correlation with the stroke survival but was not included in the final prediction model; '++' is indicated when the factor was considered important for predicting stroke outcome.

[1]haemorrhagic or ischemic

[2]measured with NIHSS

[3]atrial fibrilation

[4]platelet/lymphocyte ratio

[5]measured by Barthel Index

[6]unilateral or bilateral

[7]dysphagia and/or urinary incontinence

clinical examinations were conducted in accordance with national health care guidelines and protocols. The recorded parameters included both numerical (Tab. 2) and binary (Tab. 3) variables, that could be categorized into distinct groups for better interpretation.

The binary variables were used to describe: 1) stroke etiology - atherothrombotic, cardioemdolic, embolic stroke of undetermined source (ESUS) or other; 2) location of stroke lesion(s) within particular brain regions or partical artery basins; 3) various comorbidities (ex. ischemic heart disease, cancer, etc.); 4) stroke-associated deficits (ex. hemiparesis, aphasia, etc.) and 5) smoking.

The numerical variables include age, weight, height and the related to them Body Mass Index (BMI) and Body Surface Area (BSA), hours before admission to hospital, different characteristics of stroke severity (ex.The National Institutes of Health Stroke Scale score (NIHSS) and The Modified Rankin Scale score(mRS)), lesion size, results of the clinical blood tests (ex. red blood cells, lymphocytes, thromobocytes, etc.) and cardiac characteristics (ex. ejection fraction, end systolic volume, end diastolic volume, etc.), measured using echocardiography (EchoCG).

2.2 Methods

All data analysis was conducted using WEKA v.3.6.13 Software [18]. The feature selection procedure was performed on the whole dataset with the *CfsSubsetEval* function. This function implements the approach proposed by Hall and Smith [19], which evaluates the value of a subset of attributes by considering the individual predictive values of each feature along with the degree of redundancy between them. As the result, this approach yields a subset of parameters that are highly correlated with the outcome but have little or no correlation with each other.

The suggested prediction model for a 1-year stroke survival was a linear binary classification model based on a logistic regression. In this model, a linear combination of features is taken as an argument of the logistic function:

$$logit(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

The coefficients in the model were estimated by the maximum likelihood method. The accuracy of the prediction model was tested using a 10-fold cross-validation procedure.

3 Results

The results of the feature selection procedure on the whole dataset yielded 10 potential predictors of the stroke survival: age, stroke type, NIHSS scores at admission and at discharge, mRS score at discharge and after 90 days, hemoglobin level, recurrent stroke during the follow-up period, involvement of the anterior choroidal artery, and ejection fraction, measured with EchoCG. Out of these parameters 6 were retained in the final prediction model, which was built using logistic regression with 10-fold validation:

$$Survival = logit(1.01 - 0.02 \cdot Age - 0.06 \cdot NIHSS - 0.32 \cdot Cardioembolic - 0.02 \cdot Stroke_2$$

$$+ 0,74 \cdot AntChArt + 0.01 \cdot Hg), \quad (2)$$

where $NIHSS$ refers to NIHSS score at discharge, $Cardioembolic$ is equal to 1 whenever the stroke was confirmed to be of a cardioembolic origin, $Stroke_2$ is set to 1 if a recurrent stroke happened during 1 year after the first event, $AntChArt$ indicates involvement of the anterior choroidal artery, Hg reflects blood hemoglobin level. The accuracy of the suggested model reached 84.6%. More details on the model performance are provided in Tab. 4.

4 Discussion

In this work we presented an attempt to create a prediction model for a 1-year survival after acute ischemic stroke among patients residing in the Perm region of Russia. Since existing studies report multiple different factors as potential predictors of survival after stroke, we have considered all available clinical data to select the best parameters that could be correlated with the 1-year outcome. These parameters included data on stroke type and location, stroke-associated deficits and their severity, existing comorbidities, as well as data from routine blood and heart investigations. Out of 75 available characteristics, only 6 were retained in the final prediction model, namely: age, NIHSS score at discharge, hemoglobin level, recurrent stroke withing the following year, involvement of the anterior choidal artery and the cardioembolic ethiology of stroke. While increasing age has been negatively associated with stroke survival in almost every existing prediction model, other factors deserve more discussion.

Stroke severity, measure either by NIHSS [2], mRS [11] or CNS (Canadian Neurological Scale) [8] scores, are commonly considered as indicators of poor prognosis and are widely used in various prediction models. However, it remains an open question which scale may better describe the outcome, since in most of the studies only one scale is used and it is not possible to compare their prediction abilities on the same sample. In the present study we considered both NIHSS and mRS scales. Although both of them were selected during the feature selection procedure, only NIHSS score at discharge was retained in the final model. When NIHSS score was replaced in the final prediction model with the mRS score at discharge, the prediction accuracy remained high (83%) but was slightly smaller than for the NIHSS score. Adding both scores to the model did not lead to any significant improvement in the model performance, since NIHSS and mRS scores were highly correlated with each other. Thus, we conclude that both of these scales could be used interchangeably, however NIHSS score should be preferred at least on our sample.

Stroke type is another important factor that may influence survival rate. In most of the studies only 2 types of stroke (ischemic or hemorrhagic) are considered, with the hemorrhagic type being a significant predictor for a poor prognosis. Indeed, a 30-day mortality rate from hemorrhagic stroke is twice as large as of the ischemic stroke [20]. However, unlike such studies, in this work we have focused only on the ischemic stroke, which itself presents a heterogeneous disease [21]. Subtypes of ischemic stroke differ in their management strategies, severity of the associated deficits, recurrence rates and possibly, mortality rates. In line with the previous study of Stead et al. [22], the only subtype, included in the final prediction model, was cardioembolic, which has been previously reported to show the worst prognosis amongst ischemic stroke subtypes [23, 24]. Similar to Stead et al. [22], cardioembolic subtype in our model was considered as a separate predictor, independent of age, gender, and NIHSS score. Of interest, cryptogenic stroke, which was not included in the prediction model, was associated with a less severe outcome than cardioembolic stroke. Taking into account the diversity of the pathological mechanisms, underlying different stroke subtypes, it is possible that building prediction models separately for each stroke subtype will result in better predictions and bring under attention more clinical factors that would be significant for a particular type of stroke [25]. However, this kind of investigation was

left beyond the scope of the present work because the sample size was not sufficiently large to consider each stroke subtype separately.

Recurrent strokes have been previously associated with long-term mortality rates in patients with ischemic stroke [26–28]. In our prediction model a recurrent stroke within a 1-year period was also considered as a separate predictor of a poor outcome, highlighting the importance of prevention of recurrent vascular events. The prevention of recurrent stroke should be managed in accordance to its pathogenetic subtype [29]. However, up to 30% of all ischemic stroke cases are crytogenic [30], i.e. of an unknown origin. That is why developing new algorithms for identifying stroke subtypes presents an important direction for future research, aimed at improving a long-term outcome after ischemic stroke.

Higher hemoglobin level in our model was considered as a favourable predictor, contributing to survival after stroke. Considering it impact on oxygen delivery and blood viscosity, it is not surprising that higher level of hemoglobin has a positive influence on the stroke outcome. Indeed, recent review and meta-analysis suggest that anemia in ischemic stroke could be linked to increased mortality and disability [31, 32]. However, it is possible that the association between the levels of hemoglobin and stroke mortality is not linear [33], and both too low or too high values would increase mortality and lead poorer functional outcome.

Finaly, the involvement of the anterior choroidal artery, which was considered by our model as a favourable factor, may reflect the fact that such type of stoke is likely to be related to microangiopathy rather than to atherosclerosis or other "more serious" causes. And indeed, previous studies point to the fact that infarcts in the anterior choroidal artery basins have more favourable outcome compared to other cases of supratentorial ischaemic stroke [34].

The relative contribution of the above mentioned factors to stroke survival, as well as their subset, may potentially vary across different clinical centers, regions and with the change of the prediction period. Thus, the presented results should be extrapolated with reasonable caution and future multi-center studies are needed to confirm the generalizability of the proposed model.

Acknowledgments. The publication was prepared within the framework of the Academic Fund Program at HSE University (grant № 23-00-026 "Development of automatic approaches to determine the etiology of cryptogenic stroke for the purpose of preventing secondary acute cerebrovascular accidents").

References

[1] Donkor, E.S.: Stroke in the 21st century: A snapshot of the burden, epidemiology, and quality of life. Stroke Res. Treat. (2018) https://doi.org/10.1155/2018/3238165

[2] Wang, W., Rudd, A.G., Wang, Y., Curcin, V., Wolfe, C.D., Peek, N., Bray, B.: Risk prediction of 30-day mortality after stroke using machine learning: a nationwide registry-based cohort study. BMC Neurology **22**(195) (2022) https://doi.org/10.1186/s12883-022-02722-1

[3] Sha, L., Xu, T., Ge, X., Shi, L., Zhang, J., Guo, H.: Predictors of death within 6 months of stroke onset: A model with barthel index, platelet/lymphocyte ratio and serum albumin. Nurs Open. **8**(3), 1380–1392 (2021) https://doi.org/10.1002/nop2.754

[4] O'Donnell, M.J., Fang, J., D'Uva, C., Saposnik, G., Gould, L., McGrath, E., Kapral, M.K., Canadian Stroke Network, f.t.: The PLAN Score: A Bedside Prediction Rule for Death and Severe Disability Following Acute Ischemic Stroke. Archives of Internal Medicine **172**(20), 1548–1556 (2012) https://doi.org/10.1001/2013.jamainternmed.30

[5] Saposnik, G., Kapral, M., Liu, Y., Hall, R., O'Donnell, M., Raptis, S., Tu, J., Mamdani, M., Austin, P., Canadian Stroke Network, f.t., Group, S.O.R.C.S.W.: IScore: a risk score to predict death early after hospitalization for an acute ischemic stroke. Circulation **123**, 739–749 (2011) https://doi.org/10.1161/CIRCULATIONAHA.110.983353

[6] Williams, G.R., Jiang, J.G.: Development of an ischemic stroke survival score. Stroke **31**(10), 2414–2420 (2000) https://doi.org/10.1161/01.STR.31.10.2414

[7] Anderson, C.S., Jamrozik, K.D., Broadhurst, R.J., Stewart-Wynne, E.G.: Predicting survival for 1 year among different subtypes of stroke. results from the perth community stroke study. Stroke **25**(10), 1935–1944 (1994) https://doi.org/10.1161/01.STR.25.10.1935

[8] Solberg, O.G., Dahl, M., Mowinckel, P., Stavem, K.: Derivation and validation of a simple risk score for predicting 1-year mortality in stroke. Journal of Neurology **254**(10), 1376–1383 (2007) https://doi.org/10.1007/s00415-007-0555-2

[9] Wang, Y., Lim, L.L., Heller, R.F., Fisher, J., Levi, C.R.: A prediction model of 1-year mortality for acute ischemic stroke patients. . Archives of physical medicine and rehabilitation **84**(7), 1006–1011 (2003) https://doi.org/10.1016/s0003-9993(03)00032-7

[10] Huang, Y., Douiri, A., Fahey, M.: A dynamic model for predicting survival up to 1 year after ischemic stroke. Journal of Stroke and Cerebrovascular Diseases **29**(10), 105–133 (2020) https://doi.org/10.1016/j.jstrokecerebrovasdis.2020.105133

[11] Szlachetka, W., Pana, T., Mamas, M., Bettencourt-Silva, J., Metcalf, A., Potter, J., McLernon, D., Myint, P.: Predicting 10-year stroke mortality: development and validation of a nomogram. Acta Neurol Belg. **122**(3), 685–693 (2022) https://doi.org/10.1007/s13760-021-01752-9

[12] Yusuf, S., Reddy, S., Ounpuu, S., Anand, S.: Global burden of cardiovascular diseases. Circulation **104**(22), 2746–2753 (2001) https://doi.org/10.1161/hc4601.099487

[13] Chiu, M., Austin, P.C., Manuel, D.G., Tu, J.V.: Comparison of cardiovascular risk profiles among ethnic groups using population health surveys between 1996 and 2007. CMAJ **182**(8), 301–310 (2010) https://doi.org/10.1503/cmaj.091676

[14] Tu, J.V., Chu, A., Rezai, M.R., Guo, H., Maclagan, L.C., Austin, P.C., Booth, G.L., Manuel, D.G., Chiu, M., Ko, D.T., Lee, D.S., Shah, B.R., Donovan, L.R., Sohail, Q.Z., Alter, D.A.: Incidence of major cardiovascular events in immigrants to ontario, canada. Circulation **132**(16), 1549–1559 (2015) https://doi.org/10.1161/CIRCULATIONAHA.115.015345 https://www.ahajournals.org/doi/pdf/10.1161/CIRCULATIONAHA.115.015345

[15] Wang, L., Zhao, C., Xia, Q.-X., Qiao, S.-J.: Association between 12p13 snp rs11833579 and ischemic stroke in asian population: An updated meta-analysis. Journal of the neurological sciences **345**(1-2), 198–201 (2014) https://doi.org/10.1016/j.jns.2014.07.047

[16] Cheong, M.-Y., Bang, O.-S., Cha, M.-H., Park, Y.-K., Kim, S.-H., Kim, Y.J.: Association of the adiponectin gene variations with risk of ischemic stroke in a korean population. Yonsei Med J. **52**(1), 20–25 (2011) https://doi.org/10.3349/ymj.2011.52.1.20

[17] Sarfo, F.S., Akpa, O.M., Ovbiagele, B., Akpalu, A., Wahab, K., Obiako, R., Komolafe, M., Owolabi, L., Ogbole, G., Fakunle, A., Okekunle, A.P., Asowata, O.J., Calys-Tagoe, B., Uvere, E.O., Sanni, T., Olowookere, S., Ibinaiye, P., Akinyemi, J.O., Arulogun, O., Jenkins, C., (2023), .: Patient-level and system-level determinants of stroke fatality across 16 large hospitals in ghana and nigeria: a prospective cohort study. The Lancet. Global health **11**(4), 575–585 (2023) https://doi.org/10.1016/S2214-109X(23)00038-4

[18] Frank, E., Hall, M.A., Holmes, G., Kirkby, R., Pfahringer, B., Witten, I.H.: In: Maimon, O., Rokach, L. (eds.) Weka: A machine learning workbench for data mining., pp. 1305–1314. Springer, Berlin (2005). http://researchcommons.waikato.ac.nz/handle/10289/1497

[19] Hall, M.A., Smith, L.A.: Feature Subset Selection: A Correlation Based Filter Approach. In: 1997 International Conference on Neural Information Processing and Intelligent Information Systems. Springer, Berlin (1997). https://researchcommons.waikato.ac.nz/handle/10289/1515

[20] Rothwell, P.M., Coull, A.J., Giles, M.F., Howard, S.C., Silver, L.E., Bull, L.M., Gutnikov, S.A., Edwards, P., Mant, D., Sackley, C.M., Farmer, A., Sandercock, P.A., Dennis, M.S., Warlow, C.P., Bamford, J.M., Anslow, P., Study, .O.V.: Change in stroke incidence, mortality, case-fatality, severity, and risk factors in oxfordshire, uk from 1981 to 2004 (oxford vascular study). Lancet **363**(9425), 1925–1933 (2004) https://doi.org/10.1016/S0140-6736(04)16405-2

[21] Adams, H.P., Bendixen, B.H., Kappelle, L.J., Biller, J., Love, B.B., Gordon, D.L., Marsh, E.E.: Classification of subtype of acute ischemic stroke. definitions for use in a multicenter clinical trial. toast. trial of org 10172 in acute stroke treatment. Stroke **24**(1), 35–41 (1993) https://doi.org/10.1161/01.STR.24.1.35

[22] Stead, L.G., Gilmore, R.M., Bellolio, F., Jain, A., Rabinstein, A.A., Decker, W.W., Agarwal, D., Brown, R.D.: Cardioembolic but not other stroke subtypes predict mortality independent of stroke severity at presentation. Stroke Research and Treatment (2011) https://doi.org/10.4061/2011/281496

[23] de Jong, G., van Raak, L., Kessels, F., Lodder, J.: Stroke subtype and mortality: a follow-up study in 998 patients with a first cerebral infarct. Journal of Clinical Epidemiology **56**(3), 262–268 (2003) https://doi.org/10.1016/S0895-4356(02)00572-3

[24] Sumer, M.M., Erturk, O.: Ischemic stroke subtypes: risk factors, functional outcome and recurrence. Neurol Sci **22**, 449–454 (2002) https://doi.org/10.1007/s100720200004

[25] Guo, Y., Zhang, M., Su, Y., Liu, J., Fu, H., Wang, Q., Chen, Y.: Analysis of risk factors for poor short-term outcomes in acute cardioembolic stroke patients without reperfusion therapy. Neuropsychiatric disease and treatment **17**, 3431–3437 (2011) https://doi.org/10.2147/NDT.S335274

[26] Broman, J., Fandler-Höfler, S., Sarnowski, B., Elmegiri, M., Gattringer, T., Holbe, C., Linden, J., Malinowski, R., Martola, J., Pinter, D., Ropele, S., Schminke, U., Tatlisumak, T., Enzinger, C., Putaala, J., Aarnio, K.: Long-term risk of recurrent vascular events and mortality in young stroke patients: Insights from a multicenter study. European Journal of Neurology **30**(9), 2675–2683 (2023) https://doi.org/10.1111/ene.15850 https://onlinelibrary.wiley.com/doi/pdf/10.1111/ene.15850

[27] Aarnio, K., Haapaniemi, E., Melkas, S., Kaste, M., Tatlisumak, T., Putaala, J.: Long-term mortality after first-ever and recurrent stroke in young adults. Stroke **45**(9), 2670–2676 (2014) https://doi.org/10.1161/STROKEAHA.114.005648 https://www.ahajournals.org/doi/pdf/10.1161/STROKEAHA.114.005648

[28] Khanevski, A.N., Bjerkreim, A.T., Novotny, V., Næss, H., Thomassen, L., Logallo, N., Kvistad, C.E., group: Recurrent ischemic stroke: Incidence, predictors, and impact on mortality. Acta Neurologica Scandinavica **140**(1), 3–8 (2019) https://doi.org/10.1111/ane.13093 https://onlinelibrary.wiley.com/doi/pdf/10.1111/ane.13093

[29] Kulesh, A.A., Demin, D.A., Vinogradov, O.I.: Pathogenetic mechanisms of ischemic stroke: from verification to secondary prevention. Consilium Medicum **23**(11), 792–799 (2021)

[30] Yaghi, S., Bernstein, R.A., Passman, R., Okin, P.M., Furie,

K.L.: Cryptogenic stroke. Circulation Research **120**(3), 527–540 (2017) https://doi.org/10.1161/CIRCRESAHA.116.308447 https://www.ahajournals.org/doi/pdf/10.1161/CIRCRESAHA.116.308447

[31] Barlas, R.S., Honney, K., Loke, Y.K., McCall, S.J., Bettencourt-Silva, J.H., Clark, A.B., Bowles, K.M., Metcalf, A.K., Mamas, M.A., Potter, J.F., Myint, P.K.: Impact of hemoglobin levels and anemia on mortality in acute stroke: Analysis of uk regional registry data, systematic review, and meta‐analysis. Journal of the American Heart Association **5**(8), 003019 (2016) https://doi.org/10.1161/JAHA.115.003019 https://www.ahajournals.org/doi/pdf/10.1161/JAHA.115.003019

[32] Desai, A., Oh, D., Rao, E.M., Sahoo, S., Mahajan, U.V., Labak, C.M., Mauria, R., Shah, V.S., Nguyen, Q., Herring, E.Z., Elder, T., Stout, A., Shammassian, B.H.: Impact of anemia on acute ischemic stroke outcomes: A systematic review of the literature. PLOS ONE **18**(1), 1–11 (2023) https://doi.org/10.1371/journal.pone.0280025

[33] Zhang, R., Xu, Q., Wang, A., Jiang, Y., Meng, X., Zhou, M., Wang, Y., Liu, G.: Hemoglobin concentration and clinical outcomes after acute ischemic stroke or transient ischemic attack. Journal of the American Heart Association **10**(23), 022547 (2021) https://doi.org/10.1161/JAHA.121.022547 https://www.ahajournals.org/doi/pdf/10.1161/JAHA.121.022547

[34] Hupperts, R.M.M., Lodder, J., Heuts-van Raak, E.P.M., Kessels, F.: Infarcts in the anterior choroidal artery territory: Anatomical distribution, clinical syndromes, presumed pathogenesis and early outcome. Brain **117**(4), 825–834 (1994) https://doi.org/10.1093/brain/117.4.825 https://academic.oup.com/brain/article-pdf/117/4/825/803291/117-4-825.pdf

Table 2 Summary of the numerical variables across the sample

Parameters	Mean	Sd
Age	65.83	10.23
Weight (kg)	78.80	15.73
Height (cm)	167.21	9.66
BSA	1.88	0.20
BMI	27.95	5.05
Hours before admission	25.23	41.93
Stroke severity		
NIHSS at admission	6.78	5.28
NIHSS at discharge	3.56	4.50
mRS at discharge	1.91	1.37
mRS after 90 days	1.61	1.37
Lesion size (max diameter), mm	34.17	25.68
Blood Tests		
Red Blood Cells, x$10^{12}/L$	4.66	0.57
Hematocrit, %	40.87	5.01
Hemoglobin, g/L	138.78	18.03
Lymphocytes, %	25.03	9.74
Thrombocytes, x$10^{9}/L$	220.95	71.34
LDL (Low-Density Lipoprotein)	3.11	1.07
HDL (High-Density Lipoprotein)	1.24	0.33
TC (Total Cholesterol)	5.06	1.26
TG (Triglycerides)	1.59	0.89
GFR (Glomerular Filtration Rate)	73.87	19.42
Cardiac parameters		
Ejection fraction (EF), %	57.69	7.41
End-Diastolic Diameter (EDD)	47.17	5.34
Ventricular Septum size	12.78	1.80
Posterior wall size	11.71	1.44
End-Diastolic Volume (EDV)	101.83	26.06
End-Systolic Volume (ESV)	43.96	18.02
Heart stroke volume (SV)	57.87	13.41
SV index	30.81	6.99
LA (Left Atrial) train	20.66	5.42
LA volume	71.21	29.48
LA volume index	38.27	16.24
LA diameter	3.86	0.54
LA 2nd size	3.93	0.51
LA 3rd size	5.19	0.70
LADi (Left Atrial Diameter index)	2.08	0.39
Post-Stroke SBP (Systolic Blood Pressure)	148.98	19.59

Table 3 Summary of the binary parameters across the sample

Parameters	Male (N)	Female (N)	Total (N)
Stroke type			
Embolic stroke of undetermined source (ESUS)	70	59	129
Atherothrombotic	36	14	50
Cardioembolic	19	31	50
Other non-ESUS	15	10	25
Stroke localization			
Cortical	63	56	119
Cortico-subcortical	44	40	84
Lacunar	11	2	13
Deep lesion	32	18	50
Insula lesion	17	25	42
Corpus callosum lesion	2	6	8
Brainstem	8	4	12
Anterior cerebral artery	6	6	12
Middle cerebral artery	91	83	174
Posterior cerebral artery	18	9	27
Anterior choroidal artery	12	14	26
Posterior inferior cerebellar artery	22	12	34
Multiple lesions within the same basin	43	24	67
Multiple lesions in different basins	21	9	30
Comorbidities			
Cancer	12	11	23
Obesity	31	47	78
Arterial hypertension	136	110	246
Diabetes mellitus	15	36	51
Acute myocardial infarction	3	3	6
Myocardial infarction (>1 month ago)	21	9	30
Atrial fibrillation	7	7	14
Ischemic heart disease	40	25	65
Previous stroke	38	27	65
Neurological deficits (after the event)			
Dysphagia	2	2	4
Hemianopsia	14	8	22
Hemihypersthesia	52	34	86
Hemiparesis	98	77	175
Ataxia	14	14	28
Oculomotor disturbances	10	4	14
Neglect	12	20	32
Motor aphasia	18	17	35
Sensory-motor aphasia	22	29	51
Acute vestibular syndrome	12	8	20
Smoking	65	9	74

Table 4 Detailed model performance by class

Class	TP rate	FP rate	Precision	Recall	F-measure	ROC area
Died	0.426	0.04	0.742	0.426	0.541	0.817
Survived	0.96	0.574	0.86	0.96	0.907	0.802
Weighted Avg.	0.846	0.46	0.836	0.846	0.829	0.805