



# Early Identification of High-Risk TIA or Minor Stroke Using Artificial Neural Network

*Ka Lung Chan*<sup>1</sup>, *Xinyi Leng*<sup>1,2\*</sup>, *Wei Zhang*<sup>3</sup>, *Weinan Dong*<sup>3</sup>, *Quanli Qiu*<sup>3</sup>, *Jie Yang*<sup>4</sup>, *Yannie Soo*<sup>1</sup>, *Ka Sing Wong*<sup>1</sup>, *Thomas W. Leung*<sup>1</sup> and *Jia Liu*<sup>3\*</sup>

<sup>1</sup> Department of Medicine and Therapeutics, Prince of Wales Hospital, The Chinese University of Hong Kong, Hong Kong, China, <sup>2</sup> Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China, <sup>3</sup> Shenzhen Institutes of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen, China, <sup>4</sup> Department of Neurology, The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou, China

## OPEN ACCESS

### Edited by:

Fabien Scalzo,  
University of California, Los Angeles,  
United States

### Reviewed by:

Ruogu Fang,  
University of Florida, United States  
Maurizio Acampa,  
Azienda Ospedaliera Universitaria  
Senese, Italy

### \*Correspondence:

Xinyi Leng  
xinyi\_leng@cuhk.edu.hk  
Jia Liu  
jia.liu@siat.ac.cn

### Specialty section:

This article was submitted to  
Stroke,  
a section of the journal  
Frontiers in Neurology

**Received:** 01 April 2018

**Accepted:** 08 February 2019

**Published:** 01 March 2019

### Citation:

Chan KL, Leng X, Zhang W, Dong W,  
Qiu Q, Yang J, Soo Y, Wong KS,  
Leung TW and Liu J (2019) Early  
Identification of High-Risk TIA or Minor  
Stroke Using Artificial Neural Network.  
*Front. Neurol.* 10:171.  
doi: 10.3389/fneur.2019.00171

**Background and Purpose:** The risk of recurrent stroke following a transient ischemic attack (TIA) or minor stroke is high, despite of a significant reduction in the past decade. In this study, we investigated the feasibility of using artificial neural network (ANN) for risk stratification of TIA or minor stroke patients.

**Methods:** Consecutive patients with acute TIA or minor ischemic stroke presenting at a tertiary hospital during a 2-year period were recruited. We collected demographics, clinical and imaging data at baseline. The primary outcome was recurrent ischemic stroke within 1 year. We developed ANN models to predict the primary outcome. We randomly down-sampled patients without a primary outcome to 1:1 match with those with a primary outcome to mitigate data imbalance. We used a 5-fold cross-validation approach to train and test the ANN models to avoid overfitting. We employed 19 independent variables at baseline as the input neurons in the ANN models, using a learning algorithm based on backpropagation to minimize the loss function. We obtained the sensitivity, specificity, accuracy and the c statistic of each ANN model from the 5 rounds of cross-validation and compared that of support vector machine (SVM) and Naïve Bayes classifier in risk stratification of the patients.

**Results:** A total of 451 acute TIA or minor stroke patients were enrolled. Forty (8.9%) patients had a recurrent ischemic stroke within 1 year. Another 40 patients were randomly selected from those with no recurrent stroke, so that data from 80 patients in total were used for 5 rounds of training and testing of ANN models. The median sensitivity, specificity, accuracy and c statistic of the ANN models to predict recurrent stroke at 1 year was 75%, 75%, 75%, and 0.77, respectively. ANN model outperformed SVM and Naïve Bayes classifier in our dataset for predicting relapse after TIA or minor stroke.

**Conclusion:** This pilot study indicated that ANN may yield a novel and effective method in risk stratification of TIA and minor stroke. Further studies are warranted for verification and improvement of the current ANN model.

**Keywords:** transient ischemic attack, minor stroke, artificial neural network, risk stratification, prognosis

## INTRODUCTION

The prevalence of transient ischemic attack (TIA) is estimated to be 103.3 per 100,000 in the Chinese population (1). Although TIA may be regarded as a “benign” cerebrovascular event, subsequent stroke could be disabling. Studies conducted over 15 years ago reported that 12–20% of TIA or minor stroke patients would have a recurrent stroke within 3 months (2). The risk of stroke recurrence in such patients has been declining over the past decade partly due to advances in the stroke service system—5.1% of TIA or minor stroke patients in hospitals with dedicated TIA/minor stroke service systems had a recurrent stroke within 1 year in the recently published TIAregistry.org Project, but there have been concerns that the recurrent rate is probably higher in routine clinical practice (3). In the Clopidogrel in High-Risk Patients with Acute Nondisabling Cerebrovascular Events (CHANCE) trial, 10.0% of 5,170 minor stroke patients or TIA patients with an ABCD<sup>2</sup> score  $\geq 4$  recruited from 114 hospitals of different levels in China had a recurrent stroke at 3 months, despite of early-initiated dual or mono antiplatelet treatment (4).

A few risk scores have been developed to identify high-risk TIA or minor stroke patients, for instance, the ABCD<sup>2</sup> score, which has been commonly used in research and in clinical practice. Patients with an ABCD<sup>2</sup> score  $\geq 4$  are generally considered as high-risk patients (5, 6). However, recent studies indicated that ABCD<sup>2</sup> score may not reliably differentiate TIA or minor stroke with mimics, or those at high or low risk of recurrent stroke (7). Moreover, those with a ABCD<sup>2</sup> score  $< 4$  and  $\geq 4$  could have similar risks of recurrent stroke at 3 months (8). Other factors have been considered to supplement the ABCD<sup>2</sup> score, for instance, presence of new infarct(s) and carotid arterial stenosis and dual TIA, to form the ABCD<sup>3</sup>-I score (9). These new scoring systems have been well validated in other populations, reporting the c-statistics of 0.60–0.64 in predicting the recurrent stroke within 3 month following TIA. However, new scores such as ABCD<sup>3</sup>-I score have not been recommended for risk stratification in such patients by the guidelines by far (10, 11).

In the current study, we aimed to use a novel method to predict the risk of stroke recurrence in TIA or minor stroke patients, which was the artificial neural network (ANN) technique. It is a commonly used machine learning algorithm to form a diagnostic or risk prediction model, which typically consists of three layers of neurons, an input layer of independent variables, a hidden layer with no real-world meaning but allowing nonlinear interactions among the input variables, and an output layer for the probability of an outcome (Figure 1). Advantages of ANN over conventional statistical methods in forming a risk prediction model lie in that it detects interactions between the input variables that commonly exist in clinical studies, and that it takes into account the weights of input variables in their correlations with the outcome (12, 13).

Previous studies have used ANN to diagnose acute myocardial infarction and stroke, and to predict mortality in ischemic stroke, intracerebral hemorrhage, traumatic brain injury, etc., which demonstrates improved accuracy against conventional methods in most circumstances (14–18). However, to our knowledge,

ANN had not been applied in predicting the risk of recurrence following a TIA or minor stroke. Therefore, in this pilot study, we developed and tested ANN models for risk stratification of TIA or minor stroke. In addition, other algorithms for machine learning such as Support Vector Machine (SVM) and Naïve Bayes classifiers (NBC) have also been utilized in medical research (19, 20). For instance, SVM has been extensively applied in diagnosis of a disease or classification of groups with certain features based on imaging data (20). The NBC algorithm was able to diagnose carpal tunnel syndrome with the highest detection rate among four machine learning methods (21). Thus, in the current cohort, we also compared the performance of ANN with SVM and NBC in risk stratification of TIA and minor stroke patients.

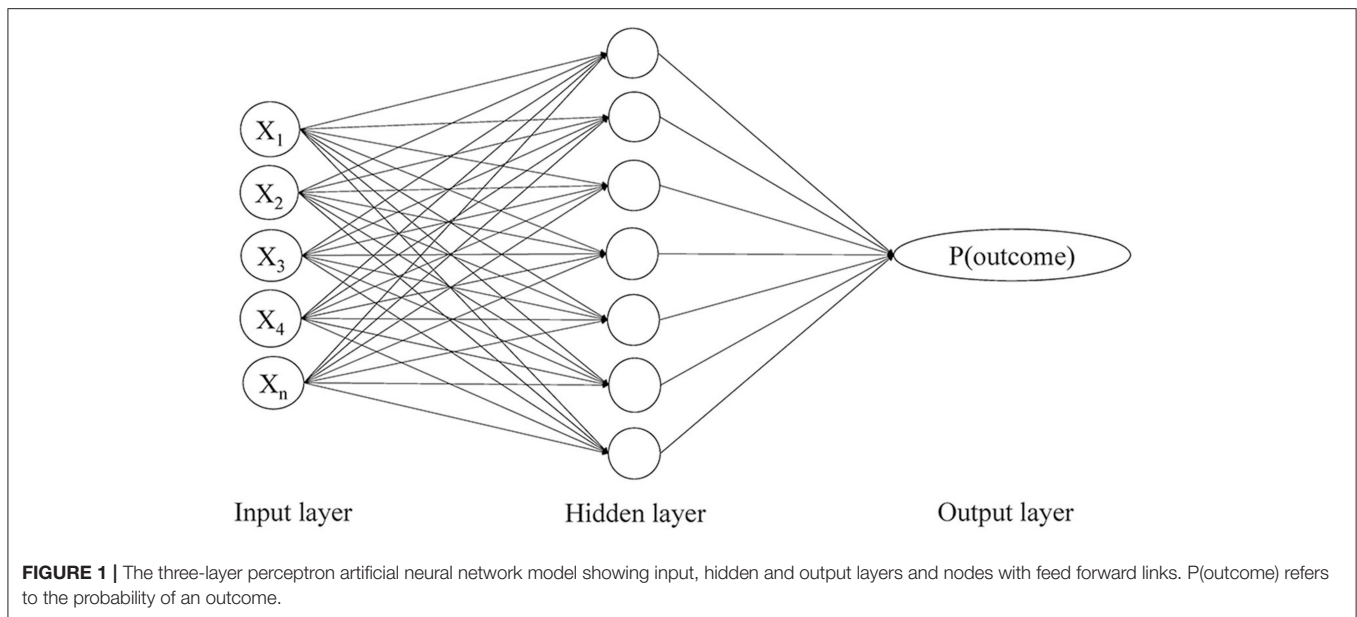
## MATERIALS AND METHODS

### Study Design and Subjects

Consecutive patients with acute TIA or minor ischemic stroke presenting at a tertiary hospital between January 2004 and December 2005 were recruited. TIA was defined as a transient episode of neurological dysfunction caused by focal brain ischemia, which completely resolved within 24 h. A minor ischemic stroke was diagnosed as sudden onset of neurological deficits caused by brain ischemia lasting longer than 24 h, with admission NIHSS score of 0–3. Diagnosis of TIA and minor ischemic stroke was made by the neurologists in charge. Stroke or TIA mimics such as toxic metabolic syndrome, seizure, migraine, demyelinating disorders, drug ingestion were excluded (22). We collected patients' characteristics at baseline as detailed below, including well-established factors readily available in clinical practice that might be associated with stroke recurrence. TIA and minor stroke patients were regularly followed up at the outpatient clinic, when recurrent cerebral ischemic events and other events were recorded. The primary outcome was defined as recurrent ischemic stroke within 1 year, as confirmed with CT or MR imaging, or diagnosed by the neurologist in charge. We developed and tested ANN models based on patients' characteristics to predict the primary outcome. We also conducted conventional statistical analyses for independent predictors for the primary outcome. The study was approved by the Joint Chinese University of Hong Kong–New Territories East Cluster Clinical Research Ethics Committee (The Joint CUHK-NTEC CREC).

### Data Collection

We collected demographic characteristics (sex, age) and vascular risk factors (smoking, hypertension, diabetes mellitus, dyslipidemia, prior TIA or ischemic stroke, atrial fibrillation, ischemic heart disease) at baseline. We also collected certain clinical characteristics including systolic and diastolic blood pressure at admission, National Institutes of Health Stroke Scale (NIHSS) score at admission, premorbid modified Rankin Scale (mRS) score, symptom duration (for transient deficits) and symptom type (unilateral weakness and slurring speech). We also gathered medications prescribed at discharge, such



as antiplatelets, antihypertensive, anticoagulants, antidiabetics and statins.

We also collected neuroimaging features including new infarct(s) on brain CT or magnetic resonance imaging (MRI), and findings by cerebrovascular workup including presence of extra-/intra-cranial arterial stenosis. Extracranial arterial stenosis was defined as at least 50% narrowing of the internal carotid artery or vertebral artery lumen on carotid duplex ultrasound, CT or MR angiography, or digital subtraction angiography, by the NASCET method (23). Intracranial artery stenosis was defined as at least 50% narrowing of middle cerebral artery, anterior cerebral artery, posterior cerebral artery, intracranial segment of internal carotid artery and vertebrobasilar arteries in transcranial Doppler, cerebral CT or MR angiography, or digital subtraction angiography, using WASID method (24). We defined large artery stenosis as either extracranial arterial stenosis or intracranial arterial stenosis.

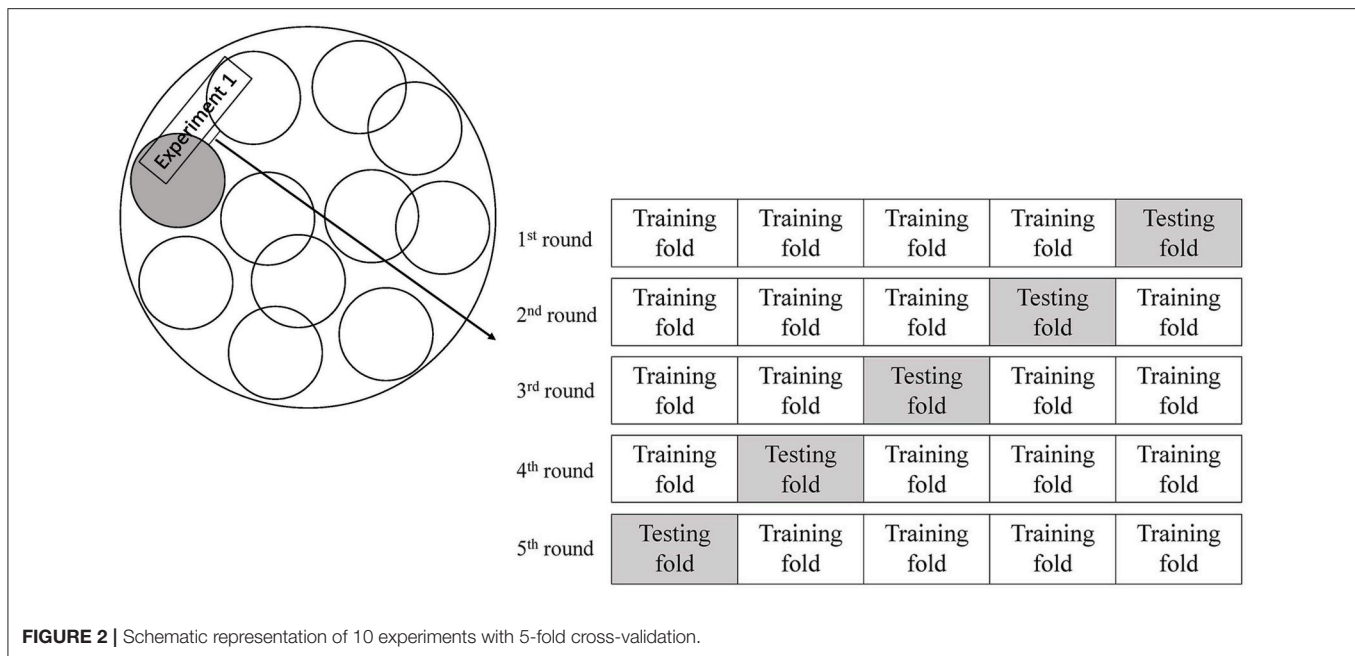
## Training and Testing in ANN

We employed a three-layer multilayer perceptron (MLP) model in this study, a most common type of ANN. It was composed of an input layer of 19 independent variables (**Supplemental Table S1**), a hidden layer with a certain number of neurons that were adjusted through training, and an output layer representing the probability of the primary outcome (**Figure 1**). A backpropagation algorithm was used to minimize the loss function by iteratively updating the weights between the neurons and thus maximize the predictive power of the ANN model for the primary outcome. Loss function represented the inconsistency between the predictive and actual values. Within each iteration of the backpropagation algorithm, the partial derivatives of the loss function with respect to each weight were propagated backward from the output layer and passed through

the hidden layer, which eventually adjusted all the weights back to the input neurons.

The study cohort was imbalanced in the numbers of patients with or without a primary outcome. This may cause biased prediction toward the no-recurrence group. Therefore, we randomly selected the same number of patients without a primary outcome as patients with a primary outcome for the training and testing (1:1 matched). In view of the relatively small sample size, we used a 5-fold cross-validation approach to train and test the ANN models (**Figure 2**), to avoid overfitting of the models (25, 26). The dataset was randomly partitioned into 5 folds, and we performed 5 rounds of training and testing of the ANN models. In each round of the experiments, 4 folds were the training subsets and the remaining subset was retained to test the ANN model. Each of the 5 subsets was only used once as the testing set in the 5-fold cross-validation process. We implemented the training and testing procedures for the ANN models in Matlab. We trained and tested ANN models with 4, 6, 8, 10, and 12 hidden neurons, respectively, and finalized the number of hidden neurons when the ANN model reached a minimal loss function in each round of the cross-validation experiments. We repeated such cross-validation procedures for 10 times (10 Experiments, **Figure 2**), each time with a new group of randomly selected patients without a primary outcome 1:1 matched with patients with a primary outcome.

For continuous independent variables, data normalization was required to speed up the gradient, thus to find the optimal solution. Shapiro-Wilk test was performed to determine whether a continuous independent variable was normally distributed. If not normally distributed, we scaled the data to a range of 0–1 before entering an ANN model. To be efficient in achieving the minimization of the loss function, we employed an adaptive moment estimation (Adam) rate rather than the constant learning rate. In addition, we assigned a weight between



–1 and +1 for the independent variables to speed up the learning process and escape local minima. We obtained the sensitivity, specificity, accuracy, and the *c* statistic of the ANN models developed in each of the 5 rounds of training and testing, in each of the 10 cross-validation experiments; we then computed the overall medians (interquartile range, IQR) of sensitivity, specificity, accuracy and *c* statistic of the models in all of the 50 rounds of training and testing.

### Training and Testing in SVM and NBC

For training the SVM models, we tried a few standard kernel functions, such as sigmoid, Gaussian, polynomial and linear kernel, and adjusted different values of the parameter *C* in the range of 0.1–10. We finally determined the kernel function and the value of the parameter *C* by the highest accuracy in predicting the test set.

Additionally, using the Naïve Bayesian equation to calculate the posterior probability for each class, we first computed the prior probability of each class  $P(c)$ , of each predictor  $P(x)$ , and the likelihood which is the probability of predictor given class  $P(x|c)$ . The outcome of prediction was the class with the highest posterior probability, using the Maximum A Posteriori (MAP) estimation.

In accordance with the ANN, we also applied the 5-fold cross-validation approach for both SVM and NBC and repeated it for 10 times with randomly selected patients without a primary outcome to 1:1 match the patients with a primary outcome. The overall medians (IQR) of sensitivity, specificity, accuracy and *c* statistic were also calculated for SVM models and NBC. We used Kruskal–Wallis test to compare the overall medians of sensitivity, specificity, accuracy and *c* statistic among the three machine learning algorithms and *post-hoc* comparisons between any two of these algorithms.

### Other Statistical Analyses

In addition, we also conducted conventional statistical analyses for predictors for the primary outcome in the study cohort. Continuous variables were presented as medians (interquartile range [IQR]), whilst categorical variables were presented as numbers (percentage). For univariate comparisons between patients with and without the primary outcome, continuous variables were analyzed with independent *t*-tests or Mann–Whitney *U*-test, whilst categorical variables were analyzed with  $\chi^2$ -test or Fisher's exact test. To identify factors independently associated with the primary outcome, variables with  $p < 0.1$  in univariate analyses were entered in a multivariate logistic regression model for further analysis. Odds ratio (OR) and 95% confidence interval (CI) were calculated.  $P < 0.05$  were considered statistically significant. All the conventional statistical analyses were conducted in IBM SPSS Statistics version 22.0 (SPSS Inc., Chicago, IL, United States).

## RESULTS

In total, 451 patients were recruited; 201 (44.6%) patients had a TIA and the remaining had a minor stroke as an index ischemic event. The median NIHSS was 1 (IQR 0–2). Forty (8.9%) patients had the primary outcome of recurrent ischemic stroke within 1 year. Twelve patients died within 1 year, three of which developed recurrent ischemic stroke before death; and the remaining nine patients died from other reasons. More of the patients with a primary outcome had a history of TIA (15 vs. 5.4%,  $p = 0.038$ ), and extra- and/or intra-cranial large artery stenosis (62.5 vs. 35.8%,  $p = 0.001$ ), compared with patients without a primary outcome event (**Table 1**). Other baseline characteristics or medications prescribed at discharge were not significantly different between those with and without

**TABLE 1** | Characteristics of patients with or without recurrent ischemic stroke within 1 year.

Variables	Recurrence (n = 40)	No recurrence (n = 411)	P-value
Age, years	71 (60–74)	65 (55–73)	0.116
Male	20 (50.0)	231 (56.2)	0.451
History of hypertension	29 (72.5)	256 (62.3)	0.201
History of diabetes mellitus	13 (32.5)	123 (29.9)	0.735
History of dyslipidemia	21 (52.5)	252 (61.3)	0.276
History of AF	7 (17.5)	37 (8.9)	0.084
History of ischemic stroke	10 (25.0)	68 (16.5)	0.177
History of TIA	6 (15.0)	22 (5.4)	0.038
History of ischemic heart disease	4 (10.0)	39 (9.5)	0.916
Smoker	8 (20.0)	142 (34.5)	0.062
Unilateral weakness	25 (62.5)	226 (55.0)	0.213
Slurring speech	15 (37.5)	107 (26.0)	0.361
Symptom duration			0.355
≤10 min	3 (7.5)	63 (15.3)	
11–60 min	4 (10)	50 (12.2)	
>60 min	33 (82.5)	298 (72.5)	
Systolic blood pressure, mmHg	166 (146–182)	160 (143–180)	0.453
Diastolic blood pressure, mmHg	80 (71–97)	83 (74–94)	0.949
NIHSS at admission	1 (0.25–2)	1 (0–2)	0.893
Premorbid mRS	0 (0–0)	0 (0–0)	0.125
Large artery stenosis	25 (62.5)	147 (35.8)	0.001
New infarct	11 (27.5)	102 (24.8)	0.709
Antiplatelet(s)	29 (72.5%)	320 (77.9%)	0.916
Anticoagulant	0 (0.0%)	9 (2.2%)	1.000
Antihypertensives	19 (47.5%)	209 (50.9%)	0.851
Antidiabetics	6 (15.0%)	92 (22.4%)	0.409
Statins	16 (40.0%)	182 (44.3%)	0.984

AF indicates atrial fibrillation; TIA indicates transient ischemic attack; NIHSS indicates National Institute of Health Stroke Scale; mRS indicates modified Rankin Scale.

a primary outcome in the study cohort (Table 1). History of TIA, large artery stenosis, atrial fibrillation and smoking were further analyzed in multivariate logistic regression to predict the primary outcome. Only presence of large artery stenosis (OR: 2.87; 95% CI: 1.45–5.67;  $p = 0.002$ ) was significantly associated with recurrent ischemic stroke within 1 year following a TIA or minor stroke in multivariate analysis.

In each of the 10 experiments of developing and testing the ANN models with 5-fold cross-validation, 40 patients were randomly selected from those without a primary outcome to 1:1 match with the 40 patients with a primary outcome (Figure 2). The number of neurons in the hidden layer was finalized as 10, after testing 4, 6, 8, 10, and 12 hidden neurons in the models. The median sensitivity, specificity and accuracy of the ANN models was 75% (63.3–83.3%), 75% (62.5–83.3%), and 75% (68.8–76.6%), respectively. The median c statistic was 0.77 (0.68–0.84) (Table 2).

After testing several kernel functions and values of parameter C, we found that the SVM model with the linear kernel

and parameter C equaling to 1 was optimal among all the others. The median sensitivity, specificity and accuracy of the SVM models was 62.5% (50–62.5%), 75% (50–87.5%), and 62.5% (56.3–68.8%).

Moreover, we also calculated the posterior probability for each class based on the Naïve Bayesian equation, and selected the outcome with highest probability. The median sensitivity, specificity and accuracy of the NBC was 62.5% (50–75%), 75% (62.5–75%), and 62.5% (56.3–68.8%). The performance of ANN models in identifying patients with recurrent ischemic stroke was better than that of SVM and NBC algorithm (Table 2).

## DISCUSSION

This pilot study demonstrated the feasibility of using ANN to predict the risk of recurrent stroke within 1 year after a TIA or minor stroke, based on parameters that are readily available in clinical practice. With a relatively small sample size and a smaller number of the primary outcome event, conventional univariate and multivariate analyses only identified the presence of cervico-cerebral large artery stenosis as an independent predictor for stroke recurrence within 1 year. However, the ANN models developed based on this study cohort showed moderate-to-good accuracy in predicting the primary outcome in comparison with SVM model and Naïve Bayes classifier, which suggested ANN as an alternative or even more effective approach for risk stratification of TIA or minor stroke.

Despite of the small sample size, presence of extra- and/or intra-cranial large artery stenosis was identified as an independent risk factor of recurrent stroke in the current study. This was consistent with relevant findings in the TIAregistry.org project and other previous studies. For instance, in the TIAregistry.org project, major brain imaging findings, including  $\geq 1$  acute ischemic lesion and  $\geq 1$  intra- and/or extra-cranial stenosis  $>50\%$ , were associated with increased risk of stroke recurrence at 3 months or 1 year after a TIA or minor stroke (3). Particularly in subjects recruited from Asia in the TIAregistry.org project, presence of intracranial stenosis tended to increase the 1-year stroke risk, independent of other confounding factors ( $p = 0.09$ ) (27). Therefore, the current study reinforced intracranial stenosis as a strong risk factor for stroke recurrence in Asians.

ANN models developed in the current study showed moderate-to-good accuracy in predicting the primary outcome, while we used a 5-fold cross-validation approach to avoid overfitting of the models. Previous studies also found ANN models accurate and effective in differentiating cerebral ischemia from stroke mimics, and in predicting mortality in patients with intracerebral hemorrhage, etc. (14, 17). As mentioned above, ANN possessed advantages over conventional statistical methods in forming a risk prediction or diagnostic model. ANN could detect complex nonlinear relationships between independent and dependent variables and assign weights to the independent variables in their associations with the outcome, thus enhancing the model fit as compared with logistic regression methods (13).

**TABLE 2** | Predictive performance of ANN, SVM and NBC models.

Statistics*	ANN	SVM	P-value (ANN vs. SVM)	Naïve Bayes	P-value (ANN vs. NBC)	P-value (overall comparison)
Sensitivity	75% (63.3–83.3%)	62.5% (50–62.5%)	<0.001	62.5% (50–75%)	0.001	<0.001
Specificity	75% (62.5–83.3%)	75% (50–87.5%)	0.081	75% (62.5–75%)	0.121	0.151
Accuracy	75% (68.8–76.6%)	62.5% (56.3–68.8%)	<0.001	62.5% (56.3–68.8%)	<0.001	<0.001
C statistic	0.77 (0.68–0.84)	0.63 (0.56–0.69)	<0.001	0.63 (0.56–0.69)	<0.001	<0.001

\*presented with medians (IQR).

ANN indicates artificial neural network; SVM indicates support vector machine; NBC indicates Naïve Bayes classifier.

ANN could also take into account possibly complex interactions between the independent variables, (28) which commonly exist in clinical scenarios, e.g., interactions between age and presence of the vascular risk factors. Moreover, in the era of precision medicine, simple dichotomization of a factor as a continuous variable in nature (e.g., age and blood pressure) in conventional scoring systems may not accurately reflect the effects of these variables in determining the risk of stroke recurrence, while the ANN approach could accommodate variables as they are in the risk prediction models.

Our results showed that the ANN outperformed the SVM and NBC. For the SVM, we tested the standard kernel functions and the best accuracy was achieved with merely 62.5% by a linear kernel. This suggests that the standard nonlinear kernel functions we tested might not be appropriate for projecting the data into a space where they can be classified by an SVM. Finding a better kernel function for this particular problem is however not intuitive and is not in the scope of this study. In contrast, the major advantage of the proposed ANN is that projecting the data into a space for classification is driven by the data. Thus, there is no need to pre-define a kernel function. For the NBC, the assumption of independence between the input variables might not be well satisfied for this study. This can negatively affect the accuracy of the NBC. Though this can possibly be improved by carefully selecting the input variables, we however did not try this procedure in order to show the ANN is an end-to-end approach that does not require data pre-selection.

The present study had several limitations. It was a retrospective single-center study with data collected years ago, but this pilot study indicated potential application of ANN in risk stratification of TIA and minor stroke patients. In addition, the study cohort was inevitably imbalanced in view of the numbers of patients with and without the primary outcome. We mitigated the imbalance between the two groups by randomly down-sampling the no-recurrence cases. However, useful information may be discarded by such resampling, and the cases randomly selected may not represent accurately the rest of the patients. We are currently collecting recent data with a larger sample size to further validate and improve the current models. Last but not least, in the current ANN models, we only employed clinical factors that are readily available in clinical practice and imaging features that could be reliably identified with routine imaging exams, while subsequent relevant studies could accommodate more clinical and imaging factors that

might influence the risk of recurrence in TIA and minor stroke patients. Automatic image analysis and image feature extraction by methods such as convolutional neural networks would help in establishing more intelligent models for risk stratification of such patients.

## CONCLUSION

Under the modern stroke service system, timely attention and management for TIA and minor stroke patients are becoming more readily available, which has significantly reduced the risk of stroke relapse in these patients. However, certain subgroups of patients are still at a high risk of subsequent disabling stroke, who may not be accurately identified with conventional risk predicting scores. Therefore, a more accurate and intelligent risk prediction strategy is needed. The ANN approach has advantages over conventional statistical methods or risk prediction scores that it could account for relationships between the independent variables, reflect complex relationships between continuous and categorical independent variables and the outcome, and quantify the weights of independent variables regarding their impact upon the outcome. The current pilot study indicated that ANN may yield a novel and effective method in risk stratification of TIA or minor stroke patients. Further studies are warranted for verification and improvement of such ANN models.

## AUTHOR CONTRIBUTIONS

XL and JL made substantial contributions to the conception and design of the study. JY, YS, KW, and TL made substantial contributions to the acquisition of data. WZ, KC, WD, and QQ made substantial contributions to the analysis of data. KC, XL, WZ, and WD contributed to the interpretations of data. KC drafted the first version of the manuscript and XL made valuable revisions. All the authors revised the draft for intellectual content, gave their final approval of the final version for publication, and agreed to be accountable for all aspects of the work.

## FUNDING

The Young Elite Scientist Sponsorship Program 2017–2019 (Reference No. 2017QNRC001), the China Association

for Science and Technology; the National Natural Science Foundation of China / Research Grants Council Joint Research Scheme (Reference No. 81661168015); Young Scientists Fund (Reference No. 81601000), National Natural Science Foundation of China.

## REFERENCES

- Jiang B, Sun H, Ru X, Sun D, Chen Z, Liu H, et al. Prevalence, incidence, prognosis, early stroke risk, and stroke-related prognostic factors of definite or probable transient ischemic attacks in China, 2013. *Front Neurol.* (2017) 8:309. doi: 10.3389/fneur.2017.00309
- Coull AJ, Lovett JK, Rothwell PM, Oxford Vascular S. Population based study of early risk of stroke after transient ischaemic attack or minor stroke: implications for public education and organisation of services. *BMJ.* (2004) 328:326. doi: 10.1136/bmj.37991.635266.44
- Amarenco P, Lavallee PC, Labreuche J, Albers GW, Bornstein NM, Canhao P, et al. One-year risk of stroke after transient ischemic attack or minor stroke. *N Engl J Med.* (2016) 374:1533–42. doi: 10.1056/NEJMoa1412981
- Wang Y, Wang Y, Zhao X, Liu L, Wang D, Wang C, et al. Clopidogrel with aspirin in acute minor stroke or transient ischemic attack. *N Engl J Med.* (2013) 369:11–9. doi: 10.1056/NEJMoa1215340
- Johnston SC, Rothwell PM, Nguyen-Huynh MN, Giles MF, Elkins JS, Bernstein AL, et al. Validation and refinement of scores to predict very early stroke risk after transient ischaemic attack. *Lancet.* (2007) 369:283–92. doi: 10.1016/S0140-6736(07)60150-0
- Lemmens R, Smet S, Thijs VN. Clinical scores for predicting recurrence after transient ischemic attack or stroke: how good are they? *Stroke.* (2013) 44:1198–203. doi: 10.1161/STROKEAHA.111.000141
- Wardlaw JM, Brazzelli M, Chappell FM, Miranda H, Shuler K, Sandercock PA, et al. ABCD2 score and secondary stroke prevention Meta-analysis and effect per 1,000 patients triaged. *Neurology.* (2015) 85:373–80. doi: 10.1212/WNL.0000000000001780
- Amarenco P, Labreuche J, Lavallee PC. Patients with transient ischemic attack with ABCD2 <4 can have similar 90-day stroke risk as patients with transient ischemic attack with ABCD2 ≥4. *Stroke.* (2012) 43:863–5. doi: 10.1161/STROKEAHA.111.636506
- Merwick Á, Albers GW, Amarenco P, Arsava EM, Ay H, Calvet D, et al. Addition of brain and carotid imaging to the ABCD2 score to identify patients at early risk of stroke after transient ischaemic attack: a multicentre observational study. *Lancet Neurol.* (2010) 9:1060–9. doi: 10.1016/S1474-4422(10)70240-4
- Kiyohara T, Kamouchi M, Kumai Y, Ninomiya T, Hata J, Yoshimura S, et al. ABCD3 and ABCD3-I scores are superior to ABCD2 score in the prediction of short- and long-term risks of stroke after transient ischemic attack. *Stroke.* (2014) 45:418–25. doi: 10.1161/STROKEAHA.113.003077
- Mayer L, Ferrari J, Krebs S, Boehme C, Toell T, Matosevic B, et al. ABCD3-I score and the risk of early or 3-month stroke recurrence in tissue- and time-based definitions of TIA and minor stroke. *J Neurol.* (2018) 265:530–4. doi: 10.1007/s00415-017-8720-8
- Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol.* (2017) 2:230–43. doi: 10.1136/svn-2017-000101
- Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemiol.* (1996) 49:1225–31. doi: 10.1016/S0895-4356(96)00002-9
- Abedi V, Goyal N, Tsvigoulis G, Hosseinichimeh N, Hontecillas R, Bassaganya-Riera J, et al. Novel screening tool for stroke using artificial neural network. *Stroke.* (2017) 48:1678–81. doi: 10.1161/STROKEAHA.117.017033
- Baxt WG. Use of an artificial neural network for the diagnosis of myocardial infarction. *Ann Intern Med.* (1991) 115:843–8. doi: 10.7326/0003-4819-115-11-843
- Celik G, Baykan OK, Kara Y, Tireli H. Predicting 10-day mortality in patients with strokes using neural networks and multivariate statistical methods. *J Stroke Cerebrovasc Dis.* (2014) 23:1506–12. doi: 10.1016/j.jstrokecerebrovasdis.2013.12.018
- Edwards DF, Hollingsworth H, Zazulia AR, Diringner MN. Artificial neural networks improve the prediction of mortality in intracerebral hemorrhage. *Neurology.* (1999) 53:351–7. doi: 10.1212/WNL.53.2.351
- Shi HY, Hwang SL, Lee KT, Lin CL. In-hospital mortality after traumatic brain injury surgery: a nationwide population-based comparison of mortality predictors used in artificial neural network and logistic regression models. *J Neurosurg.* (2013) 118:746–52. doi: 10.3171/2013.1.JNS121130
- Langarizadeh M, Moghbeli F. Applying naive Bayesian networks to disease prediction: a systematic review. *Acta Inform Med.* (2016) 24:364–9. doi: 10.5455/aim.2016.24.364-369
- Orru G, Pettersson-Yeo W, Marquand AF, Sartori G, Mechelli A. Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neurosci Biobehav Rev.* (2012) 36:1140–52. doi: 10.1016/j.neubiorev.2012.01.004
- Sayin R, Keskin S, Hamamci M. Evaluation of several classification methods in carpal tunnel syndrome. *J Pak Med Assoc.* (2017) 67:1654–7.
- Yang J, Fu JH, Chen XY, Chen YK, Leung TW, Mok V, et al. Validation of the ABCD2 score to identify the patients with high risk of late stroke after a transient ischemic attack or minor ischemic stroke. *Stroke.* (2010) 41:1298–300. doi: 10.1161/STROKEAHA.110.578757
- Moneta GL, Edwards JM, Chitwood RW, Taylor LM, Lee RW, Cummings CA, et al. Correlation of North American Symptomatic Carotid Endarterectomy Trial (NASCET) angiographic definition of 70% to 99% internal carotid artery stenosis with duplex scanning. *J Vasc Surg.* (1993) 17:152–9. doi: 10.1016/0741-5214(93)90019-1
- Lynn M, Chimowitz M, Howlett-smith H, Frankel M, Hertzberg V, Chaturvedi S, et al. Warfarin vs aspirin for symptomatic intracranial stenosis: subgroup analyses from WASID. *Stroke.* (2005) 36:457. doi: 10.1212/01.wnl.0000238506.76873.2f
- Bentley P, Ganesalingam J, Jones ALC, Mahady K, Epton S, Rinne P, et al. Prediction of stroke thrombolysis outcome using CT brain machine learning. *Neuroimage Clin.* (2014) 4:635–40. doi: 10.1016/j.nicl.2014.02.003
- Zhou Z, Folkert M, Cannon N, Iyengar P, Westover K, Zhang Y, et al. Predicting distant failure in early stage NSCLC treated with SBRT using clinical parameters. *Radiother Oncol.* (2016) 119:501–4. doi: 10.1016/j.radonc.2016.04.029
- Hoshino T, Uchiyama S, Wong LKS, Sissani L, Albers GW, Bornstein NM, et al. Differences in characteristics and outcomes between Asian and non-Asian patients in the TIAregistry.org. *Stroke.* (2017) 48:1779–87. doi: 10.1161/STROKEAHA.117.016874
- Basheer IA, Hajmeer M. Artificial neural networks: fundamentals, computing, design, and application. *J Microbiol Methods.* (2000) 43:3–31. doi: 10.1016/S0167-7012(00)00201-3

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Chan, Leng, Zhang, Dong, Qiu, Yang, Soo, Wong, Leung and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.