Prediction of stroke diagnosis through a classification model based on cerebral autoregulation: a preliminary study

R Romanelli[1], ASM Salinet[2], RC. Nogueira[2], J Salinet[1]

[1]HEartLab, Federal University of ABC, Sao Bernardo do Campo, Brazil
[2]Neurology Department, University of São Paulo, Hospital das Clínicas, São Paulo, Brazil

Abstract

Ischemic stroke can severely impact the brain perfusion. Cerebral autoregulation (CA) maintains a constant cerebral blood flow during changes in arterial blood pressure. Previous studies demonstrated that CA are compromised with increasing occurrence of stroke. Considering this, this paper aims to evaluate the feasibility of employing K-nearest neighbor techniqueto automatically predict stroke outcomes, based on CA evaluation by Transfer Function analysis method. One-way ANOVA was performed to verify differences between means of the groups (stroke levels and control). The results presented in this study add to existing evidence that cerebral hemodynamics are negatively affected in stroke patients. The classification algorithm showed promising results, particularly using Transfer Function Analysis with gain (Acc. 68% for very low frequency and Acc 62% for low frequency) and phase parameters (Acc. 62% for very low frequency). But most metrics presented AUC values close to 50% (mean of 50.7 ± 7.3 %), which indicates that the results have a large percentage of false outcomes. Although further analyzes require improvements in the algorithm, this study showed a promising path in the development of more accurate and reliable diagnostic tools for stroke patients, which can lead to better clinical outcomes.

1. Introduction

The human brain relies on a consistent supply of oxygen and energy substrates through the cerebral blood flow (CBF) [1]. The occurrence of ischemic stroke can severely impact this supply by altering brain perfusion, which is dependent on cerebral arteries, collaterals, and compensatory mechanisms [2]. The mechanism of cerebral autoregulation (CA) maintains a constant CBF during changes in arterial blood pressure (ABP) between 60 to 140 mmHg [1]. However, if compromised, CBF becomes dependent on ABP resulting in critical conditions for the brain [3].

Previous studies demonstrated that CA are compromised with increasing the occurrence of stroke [2]. Therefore, this paper aims to evaluate the feasibility of employing K-nearest neighbor (KNN) technique to automatically predict stroke outcomes, based on the Transfer Function Analysis's metrics (TFA). This method is one of the most accomplished methods which models the CA as a linear system in which spontaneous ABP variations (system input) reflect in the CBF velocity (CBFv) (system output) to report gain, phase and coherence values [4].

2. Methods

2.1. Data Collection

The study includes 110 subjects, provided by two world reference centres in cerebral hemodynamic research, with 45 subjects from Hospital das Clínicas, in São Paulo, Brazil (local ethics committee number 982.280) and 65 subjects from University of Leicester, in Leicester, England (The Nottingham Research Ethics Committee 1, United Kingdom, Ref: 11/EM/0016).

CBFv was obtained through bilateral middle cerebral artery using a Transcranial Doppler (DWL Doppler-Box, Germany) for at least 5 minutes beat-to-beat. Simultaneously, for each individual, beat-to-beat non-invasive ABP was collected using a Finometer (Finapres Medical Systems®, Amsterdam, The Netherlands). Patients were ranked by two stroke scales, NIHSS the National Institutes of Health Stroke Scale (NIHSS) which categorizes stroke as mild (≤4), moderate (5-15) and severe (≥16) and modified Rankin Scale (mRS) score that defines poor functional outcome as mRS >2.

2.2. Pre-processing

The pre-processing of CBFv and ABP signals is conducted according to evidence-based recommendations from CARNet – Cerebrovascular Research Network [5]. First, the signals are calibrated based on brachial sphygmomanometry and synchronized with a fixed ABP

delay of 0.9 s. Then, a Butterworth low-pass filter (cutoff frequency: 20 Hz, order 3) is applied and a surrogated beat-to-beat signals are calculated: first, the fiducial points related to the cardiac systole (RR mark) are identified in the bilateral ABP and CBFv signals. The RR mark from the signals is detected using automatic multiscale-peak detection (AMPD) [3]. Afterwards, the average beat-to-beat value is calculated, and a surrogate signal is generated for each signal, with the Y-axis values being the average beat-to-beat values and the X-axis representing the central time instant of the beat. Then, the surrogate signals are resampled at a frequency of 5 Hz using cubic spline interpolation. Finally, a triangular moving average three-point window with coefficients [0.25, 0.5 and 0.25] is applied to smooth the signal [3].

2.3 Cerebral autoregulation Analysis: Transfer Function Analysis

TFA method decomposes a stationary input and output signals into sums of sines and cosines of multiple frequencies using the classical discrete Fourier transform combined with Welch's method [3]. The input sinusoids are transformed into output sinusoids of the same frequency, but with a different amplitude (or gain) and shifted in time (phase shift of the response). The coherence parameter helps to identify conditions where estimates of gain and phase may not be reliable. In the frequency domain the autoregulation system works as a high-pass filter and it is mostly assessed in the frequency bands of 0.02-0.07 Hz (Very Low Frequency, VLF), 0.07-0.2 Hz (Low Frequency, LF) and 0.2-0.5 Hz (High Frequency, HF) [4].

Starting with Welch's method, the window used is the Hanning window, with 100 s of length and 50% of overlap between the windows. The auto-spectrum of ABP S_{xx} and cross spectrum of ABP with CBFv S_{xy} are obtained from equations (1) and (2), respectively. Where X_w and Y_w are the windowed discrete Fourier transform of ABP and CBFv segments and $E\{\cdot\}$ is the expected value operator among the segments. The transfer function is estimated through the relation between S_{xx} and S_{xy}, as presented in equation (3).

$$S_{xx}(f) = E\{X_w(f)X_w^*(f)\} \quad (1)$$

$$S_{xy}(f) = E\{X_w(f)Y_w^*(f)\} \quad (2)$$

$$H(f) = \left(\frac{S_{xx}(f)}{S_{xy}(f)}\right)^* \quad (3)$$

From the transfer function represented by $H(f)$, the gain $G(f)$, phase $\Phi(f)$ and coherence $MQC(f)$ can be obtained in the previously cited frequency ranges (i.e VLF, LF and HF), throughout equations 4 to 7, where $u(f)$ and $v(f)$ are the real and imaginary parts of the transfer function.

$$H(f) = u(f) + iv(f) \quad (4)$$

$$G(f) = |H(f)| = \sqrt{u(f)^2 + v(f)^2} \quad (5)$$

$$\Phi(f) = \tan^{-1}\left(\frac{v(f)}{u(f)}\right) \quad (6)$$

$$MQC(f) = \frac{|S_{xy}(f)|^2}{S_{xx}(f)S_{yy}(f)}, \quad (7)$$

2.3. K-nearest neighbor classifier

After pre-processing the CBFv and ABP signals and evaluating them using TFA method, a predictive model for classifying the results was developed with the goal of categorizing stroke outcome using a well know machine learning method. The classification model is based on comparing the normative CA findings obtained by NIHSS index with TFA parameters (gain and phase). Coherence was not considered for the classification, because it is a measure of gain and phase reliability. Individuals were defined for this stage as divided into two groups, control group (including control individuals and patients with mild level of stroke) and injured group (including patients with moderate or severe stroke level).

The KNN model is used as the classification method, which is a supervised classier, that uses the proximity to make predictions about the grouping of an individual data point. The algorithm uses CA results as inputs to predict stroke status of each individual as output.

To define the value of K-neighbours (of KNN model), values from 1 to 30 are tested, and the smallest value of K that achieves the highest accuracy of the results is selected to proceed. The data was randomly divided into 75% training and 25% testing. The model's results were evaluated through accuracy between the predicted and training data, by ROC curve area under the curve (AUC).

2.4. Statistical analysis

Mean and standard deviation were initially calculated for the raw signals of ABP and CBFv, and for TFA parameters (gain, phase and coherence). Then, the Shapiro-Wilk normality test and Levene's test for homogeneity of variances were performed. For groups that met at least one of the conditions, a one-way ANOVA test was applied to verify if there is a difference between the means of the groups. For metrics that showed a statistically significant difference between the groups, Tukey's Post Hoc test was applied to identify where the statistical difference is located.

3. Results

From the 110 subjects studied (50 control and 60 stroke), CA is evaluated based on the middle artery of the left and right side, respectively, resulting in a total of 220 CA measurements evaluation. From these measurements, only 142 (67 control and 75 ischemic stroke) presented acceptable results (i.e., signals and results that did not show physiological values were rejected). Table 1 summarize the clinical characteristics and CA results.

3.1. Cerebral autoregulation

The results shows that the increase in stroke level based on NIHSS score also increases the mRS score. The stroke groups presented higher ABP (systolic and diastolic) and heart rate in moderate and severe groups.

Table 1. Clinical characteristics and cerebral autoregulation results.

	Stroke			
	Mild N = 20	Moderate N = 49	Severe N= 6	Control N = 67
	Mean ± std.	Mean ± std.	Mean ± std.	Mean ± std.
Stroke scales				
NIHSS	2.9 ± 1.1	8.6 ± 2.5	18.6 ± 2.7	-
mRS	0.8 ± 0.6	2.4 ± 1.2	3.5 ± 2.0	-
Hemodynamics				
Systolic ABP, mmHg	156.58 ± 17.49	157.38 ± 14.02	165.01 ± 12.46	150.25 ± 25.40
Diastolic ABP, mmHg	86.53 ± 20.41	90.70 ± 7.99	95.25 ± 4.49	72.91 ± 13.37
Heart rate, bpm	67.33 ± 4.65	64.99 ± 8.57	69.71 ± 9.98	56.84 ± 3.77
CBFv left hem., cm/s	46.88 ± 6.60	45.87 ± 5.02	59.60 ± 5.19	57.12 ± 5.23
CBFv right hem., cm/s	40.68 ± 2.86	51.13 ± 4.63	54.07 ± 4.41	72.44 ± 4.92
Cerebral autoregulation				
TFA - Gain, cm/s mmHg				
VLF	1.71 ± 0.96[a]	1.38 ± 0.61[b]	1.60 ± 0.25	0.98 ± 0.53
LF	1.61 ± 0.74	1.54 ± 0.59	1.63 ± 0.52	1.32 ± 0.57
HF	2.12 ± 1.24[a]	1.73 ± 0.65	1.63 ± 0.47	1.54 ± 0.54
TFA - Phase, rad				
VLF	0.63 ± 0.59	0.96 ± 0.78	0.26 ± 0.33	0.84 ± 0.81
LF	0.72 ± 0.35[d]	0.52 ± 0.28	0.26 ± 0.18	0.58 ± 0.38
HF	0.15 ± 0.27	0.08 ± 0.27	0.12 ± 0.16	-0.02 ± 0.27
TFA - Coherence				
VLF	0.73 ± 0.19[a]	0.62 ± 0.18b	0.73 ± 0.17	0.53 ± 0.18
LF	0.69 ± 0.15	0.70 ± 0.15	0.65 ± 0.21	0.68 ± 0.14
HF	0.68 ± 0.16	0.77 ± 0.14b	0.81 ± 0.09[c]	0.64 ± 0.16

[a]$p<.05$ for Tukey Post-hoc test differences between stroke mild and control. [b]$p<.05$ for Tukey Post-hoc test differences between stroke moderate and control. [c]$p<.05$ for Tukey Post-hoc test differences between stroke severe and control.

Results of CA assessment showed that gain (VLF and LF) and phase (HF) presented higher differences between the stroke and control groups. Coherence presented values higher than 0.5 for all frequency ranges, indicating that the estimation of gain and phase parameters were reliable.

3.2. Classification

Significantly higher accuracy is seen in gain (VLF and LF) compared with the other results. Gain (VLF) and phase (VLF) presented the highest values of AUC. Although most of the methods presented AUC close to the minimum acceptable (50%).

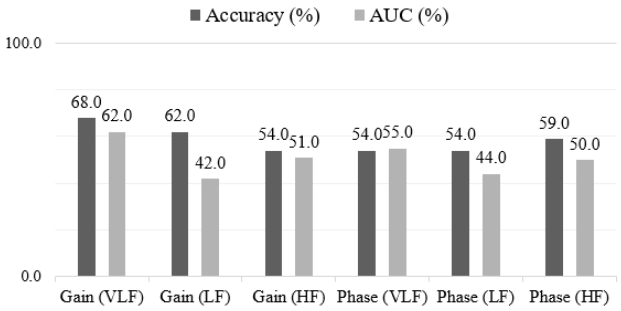

Figure 1. Classification performance measurement for TFA parameters. Font: Author.

4. Discussion

The results presented in this study add to existing evidence that cerebral hemodynamics are negatively affected in stroke patients, with increase stroke severity leading to greater CA impairment [2]. Besides that, stroke severity can also be associated negatively with mRS outcome, indicating poor functioning recovery in patient with CA impairment.

These findings suggests that the use of CA results, may be useful in predicting outcomes for stroke patients and identifying those who are at the higher risk of poor recovery.

The results of the classification algorithm showed that TFA gain in VLF and HF and phase in VLF may be particularly useful in diagnosing stroke patients, as they showed significantly higher accuracy and AUC compared with the other parameters. However, most parameters result in AUC values close to 50%, which indicates that the classification algorithm has a large percentage of false negative and false positive outcomes. One possible explanation for this could be due to asymmetric classification error, in other words, if the model is having difficulty identifying instances of a class but is correctly classifying instances of the other class, the accuracy may be high, but the AUC may be low. Currently, we are testing other machine learning methods to improve classification, such: support vector machine (SVM) with linear classification and artificial neural networks (ANN). SVM a model that divides data in hyperplanes the maximize the margin between classes, it can be useful in the stroke outcome classification context because it can handle well the high-dimensional feature space and can find optimal boundary between the classes. While ANN, is a model inspired by the structure of the human brain that uses layers of neurons to learn patters in data, this algorithm can learn and represent non-linear relationships between the input features and the class labels. Nevertheless, it is recommended to test multiple models and compare their performances to choose the best one for the task.

Although further analyzes require improvements in the classification algorithm, the findings of this study showed a promising path in the development of more accurate and reliable diagnostic tools for stroke patients, which can lead to better clinical outcomes. Overall, this preliminary study highlights the importance of CA and its role in the pathophysiology of stroke and show that the use of CA results can be helpful in predicting outcomes for stroke patients.

Acknowledgments

RR is supported by grant #2021/10288-8 FAPESP. JS is supported by grant 2018/25606-2, FAPESP. ASMS was funded by PNPD/UFABC of the CAPES (2018– 2019), (UFABC—2020–2021) and CNPq #426440/2018-8.

References

[1] Payne S. “Cerebral Autoregulation: Control of Blood Flow in the Brain. 2016.

[2] Salinet AS et al. Impaired cerebral autoregulation and neurovascular coupling in middle cerebral artery stroke: Influence of severity?. *J Cereb Blood Flow Metab*. 2019;39(11):2277-2285.

[3] Salinet, J et al. “CAAos platform: an integrated platform for analysis of cerebral hemodynamics data.” Physiological measurement vol. 42,10.

[4] Zhang R Giller CA Levine BD. Transfer function analysis of dynamic cerebral autoregulation in humans Am J Physiol. 1998;274:(1 Pt 2):H233–H241.

[5] Claassen JA et al. “Transfer function analysis of dynamic cerebral autoregulation: A white paper from the International Cerebral Autoregulation Research Network 2016;36(4):665-80.

Renata Romanelli da Costa
Biomedical Engineering - CECS
Federal University of ABC - UFABC
Street: Av.Anchieta, Sao Bernardo do Campo - SP, Brazil
E-mail address: romanelli.c@aluno.ufabc.edu.br