# The Probability of Ischaemic Stroke Prediction with a Multi-Neural-Network Model

**Yan Liu [1], Bo Yin [1,2,]\* and Yanping Cong [1]**

[1]   College of Information Science and Engineering, Ocean University of China, Qingdao 266000, China; yanliukd@163.com (Y.L.); congyp@ouc.edu.cn (Y.C.)
[2]   Pilot National Laboratory for Marine Science and Technology, Qingdao 266000, China
\*   Correspondence: ybfirst@126.com or ybfirst@ouc.edu.cn

check for updates

**Abstract:** As is known, cerebral stroke has become one of the main diseases endangering people's health; ischaemic strokes accounts for approximately 85% of cerebral strokes. According to research, early prediction and prevention can effectively reduce the incidence rate of the disease. However, it is difficult to predict the ischaemic stroke because the data related to the disease are multi-modal. To achieve high accuracy of prediction and combine the stroke risk predictors obtained by previous researchers, a method for predicting the probability of stroke occurrence based on a multi-model fusion convolutional neural network structure is proposed. In such a way, the accuracy of ischaemic stroke prediction is improved by processing multi-modal data through multiple end-to-end neural networks. In this method, the feature extraction of structured data (age, gender, history of hypertension, etc.) and streaming data (heart rate, blood pressure, etc.) based on a convolutional neural network is first realized. A neural network model for feature fusion is then constructed to realize the feature fusion of structured data and streaming data. Finally, a predictive model for predicting the probability of stroke is obtained by training. As shown in the experimental results, the accuracy of ischaemic stroke prediction reached 98.53%. Such a high prediction accuracy will be helpful for preventing the occurrence of stroke.

**Keywords:** convolutional neural network; stroke; feature extraction; disease diagnosis

## 1. Introduction

### 1.1. Related Work

Stroke is one of the main causes of death and disability worldwide [1]. Due to lack of effective treatment, it is difficult to cure stroke patients completely. On the other hand, even if the patient is cured, they still have to face harsh realities: permanent disability, incapacity [2], reduced social activities [3], etc. Hence, the disease puts a heavy burden on patients, the health care system and society. According to lots of studies, there is a certain eclipse period before the onset of stroke and early prediction and prevention can effectively reduce incidence rate of the disease. Actually, some premonitory symptoms appear during the eclipse period of stroke. For example, Zhang designed a questionnaire to put forward that a series of symptoms will appear during the eclipse period before the onset of stroke [4], such as chronic yawning, frequent choking coughs and a habit of biting the tongue. Goldstein [5] has divided risk factors into nonmodifiable risk factors (age, gender, race and genetic, etc.), well-documented and modifiable risk factors (high blood pressure, smoking, diabetes, atrial fibrillation, some other heart diseases, etc.) and less well-documented or potentially modifiable risk factors (metabolic syndrome, alcoholism, drug abuse, etc.), and it was suggested that changing well-documented and modifiable risk factors or less well-documented or potentially modifiable risk

factors could reduce the risk of stroke onset. Therefore, early detection and prevention can effectively reduce the risk of ischaemic stroke and increase the success rate of cure during the eclipse period.

In the medical field, a large-scale independent electronic health record (EHR) database has been established, which provides a large quantity of clinical diagnoses, and imaging and laboratory data [6]. It makes a considerable contribution to predicting the occurrence of diseases by artificial intelligence, which has been widely used for analysis and prediction, with remarkable results. In this field, it is used to deal with complex disease prediction tasks [7]. For example, Czabanski used the Lagrangian support vector machine to predict atrial fibrillation (AF). The results obtained during the test stage showed that the classification accuracy was 98.86%; it can effectively detect AF and provide more reliable information for the processing stage after the onset of AF [8]. Osman used an automatic epilepsy diagnostic method based on a self-organization map (SOM) method to discover epilepsy [9], and the detection accuracy of the model reached 97.47%; it could effectively detect epilepsy. Therefore, the combination of deep learning and big data has made remarkable achievements in the field of disease prediction.

In terms of the prediction of stroke, many researchers used artificial intelligence technology to predict stroke. For example, Songhee [10] used a deep neural network based on extended PCA to extract features from medical service usage and health behavior data and predicted stroke; the area under the curve (AUC) value of our method was 83.48%. It can be used by both patients and doctors to prescreen for possible strokes; however, the risk factors considered are not comprehensive, and the predictive performance of the model is average. Chen-Ying [11] used a deep learning network model to perform feature extraction and stroke prediction on electronic medical claim records. The area under the curve (AUC) value of the method was 0.915. The prediction performance of the prediction model is good. However, there is still a problem of insufficient consideration of factors related to stroke. Although the above-mentioned researchers proved that using deep learning techniques can predict stroke, they all used single-modal data for model training and prediction, and the accuracy of the model is average.

In addition, many researchers have explored many different predictors related to stroke, which provide a sufficient feasibility basis for the prediction methods of stroke. Flint, a stroke specialist at Kaiser Permanente medical center in the United States, performed 36 million blood pressure measurements on more than one million people. This study brought a large amount of data and provided a clear answer to the basic question about blood pressure and stroke: "diastolic blood pressure" and "systolic blood pressure" are independent predictors of stroke risk [12]. Wesley collected left ventricular hypertrophy (LVH) detected by electrocardiography (ECG-LVH) and LVH detected by echocardiography to assess the risk of stroke. Finally, ECG and echocardiography were found to be predictive factors for stroke [13]. Di [14] recruited 11 post-stroke patients and 20 healthy control subjects and performed an elbow sinusoidal trajectory tracking experiment. The experimental results showed that the EMG's fApEn (fuzzy approximate entropy) values of the experimental group and the control group were significantly different, so stroke can induce neurological changes in paretic muscles. Bodapati examined that whether 24-hour heart rate variability (HRV) added predictive value to the Cardiovascular Health Study clinical stroke risk score (CHS-SCORE) [15]. The value of adding HRV to the CHS-SCORE was assessed with stepwise Cox regression analysis. They found that two HRV parameters, CV% (coefficient of variance of NN intervals) and power law slope, emerged as significantly associated with incident stroke when added to a validated clinical risk score. Chantamit-o-pas et al. integrated the icd-10 code into the health records and other potential risk factors in Electronic Healthcare Records (EHRs) into the patterns and models to predict stroke [16].

*1.2. Novelty and Contributions*

Current forecasting methods mainly use single-modal data, and the field of stroke prediction is no exception. If the features in text, image or stream data are needed, some methods are used to extract the required features from the unstructured data. How to process multi-modal data with the

help of multiple end-to-end neural networks, and to make fusions and predictions, are very important technical challenges.

To overcome those challenges, we combined the stroke risk predictors obtained by previous researchers and propose a multi-model fusion convolutional neural network architecture to predict the occurrence probability of ischemic stroke (here the stroke referred to is ischemic stroke). In this method, a convolutional neural network is the first part used for feature extraction. In fact, there are two models here. One of them is convolutional neural network based on the VGG16 model used to extract features from an electrocardiogram (ECG), an electromyogram (EMG), a blood pressure graph and a heart rate graph. The other is a one-dimensional convolutional neural network model, which is used to extract features from personal health information (smoking, drinking, history of atrial fibrillation, history of hyperlipidemia, etc). The second part fuses all the features acquired in the first part and makes stroke predictions. This study used multiple end-to-end models to fuse and predict multi-modal data of all stroke-related predictors; we also solved the problem of multi-modal data fusion in disease prediction.

## 2. Materials and Method

### 2.1. Dataset

Since stroke is accompanied by dynamic cerebral automatic regulation injury, the factors related to the state of dynamic cerebral automatic regulation injury, such as blood pressure, heart rate, ECG and EMG, can be used to predict stroke. The public dataset called cerebral vasoregulation in elderly with stroke [17], published by Goldberger [18], is used here. It contains data from a large multi-modal study that investigated the effects of ischaemic stroke on cerebral vascular regulation. The cross-sectional study compared 60 subjects who suffered strokes to 60 control subjects, collecting the following data for each patient and normal person across multiple days: transcranial doppler of cerebral arteries; 24-hour blood pressure numerics; high resolution waveforms (ECG, blood pressure, $CO_2$ and respiration) during various movement tasks; 24-hour ECG, EMG, and accelerometer recordings; and gait pressure recordings during a walking test. The parts of the human body detected by ME6000 are shown in Table 1. As the information of some research subjects is incomplete, the data from 39 patients with ischaemic stroke and 40 normal persons were chosen. Demographic characteristics among the two groups are shown in Table 2. Blood pressure, heart rate, ECG and EMG and personal health information (smoking, drinking, history of atrial fibrillation, history of hyperlipidaemia, etc.) are used in this paper; baseline information about stroke patients and normal persons is shown in Tables A1–A12.

These factors can reflect the impacts of stroke on cardiovascular and cerebrovascular diseases. Hence, the current state of cardiovascular and cerebrovascular diseases can be judged by observing the factors. On the other hand, since personal health data such as smoking [19], alcoholism [20], history of hypertension [21] and history of hyperlipidaemia have strong correlations with stroke [22], it is necessary to combine personal health data to infer the possibility and probability of stroke attack. Table 3 lists the predictors related to stroke used in this article.

**Table 1.** The parts of the human body detected by ME6000.

| Type | Position |
| --- | --- |
| ECG 1 | CH1 V5/V6-L clavicle |
| ECG 2 | CH2 V1/V2 L clavicle |
| EMG 1 | gastrocnemius right |
| EMG 2 | gastrocnemius left |

**Table 2.** Demographic characteristics among the two groups.

| Group | Stroke | Control | *p* |
|---|---|---|---|
| Age(years) | 64.21 (±8.94) | 64.48 (± 8.07) | 0.87 |
| Sex (male, female) | 20,19 (39) | 17,23 (40) | N=79 |
| Race (W, A, AA) | 33,1,5 | 33,3,4 | |
| Body mass index(kg/m2) | 27.53 (± 4.74) | 27.59(± 6.48) | 0.95 |
| Years after stroke | 6.05 (±4.88) | - | - |
| Stroke side (right, left) | 24,19 | - | - |
| Infarct volume (cm3) | 18.69 (±34.06) | - | - |
| NIHSS | 2.71 (±2.72) | - | - |
| MRS | 1.2 (±1.14) | - | - |

Continuous variables are presented as mean ± SD, Ordinal variables are presented as mean ± SD (range), Nominal variables are presented as numbers, Comparison is not significantly different if $p > 0.05$, Race: W—White, A—Asian, AA—African American.

The data used in the experiment can be divided into two types, streaming data and structured data. The streaming data contain blood pressure, heart rate, ECG and EMG data. Twenty-four-hour beat-to-beat heart rate and BP monitoring using Dynapulse were measured at 20 min intervals during daytime and at 30 min intervals at night; 24-hour ECG and EMG monitoring was done using ME6000 devices. ECG and EMG data were sampled at 1000 Hz. ECG and EMG were done for 24 h during sleep and daily activities, such as walking; therefore, both ECG and EMG data have two types, static state data and dynamic data. The structured data are the personal health information (age, gender, height, history of hypertension, alcoholism, smoking, etc.) of 79 subjects.

For streaming data, graphical and integrated processing was carried out. First, streaming data were converted into graphs, such as a blood pressure graph (as shown in Figure A2), a heart rate graph (as shown in Figure A3), an ECG and an EMG (as shown in Figure A1). In order to easily show the characteristics of all streaming data, the blood pressure graph (upper left), heart rate graph (upper right), an ECG and an EMG (lower left and right) are shown in Figure 1.
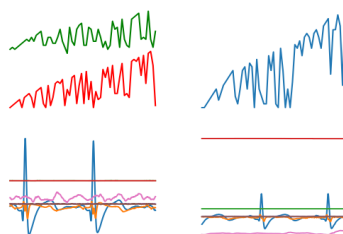


**Figure 1.** The composite image contains a blood pressure graph, a heart rate graph, an ECG and an EMG.

One-dimensional data were obtained by transforming and processing the structured data. Some of the information (gender, history of hypertension, family members with histories of hypertension, etc.) has two or more values, so it needs to be processed by one-hot encoding [23]. For example, in Table 4, index values from 7 to 9 indicate race types (African American, White, and Asian). The details are shown in Tables 4 and 5, which respectively represent data before and after personal health conversion.

**Table 3.** Influencing factors of stroke.

| Human Characteristic Data | Age Mass/kg Gender | Height/m BMI Race |
|---|---|---|
| Personal medical history | Htn patient medical history<br>Dizziness autonomic symptoms<br>DM/on-DM stroke<br>OHspecific autonomic symptoms<br>HTN years patient medical history<br>Stroke patient medical history<br>Heart failure =CHF /ifaction=-MI<br>patient medical history | Neuropathy autonomic symptoms<br>Numbness autonomic symptoms<br>Syncope autonomic symptoms<br>Atrial fibrillation patient medical history<br>Cancer patient medical history<br>DM patient medical history |
| Behavioral | Current tobacco use<br>Previous alcohol use<br>ALCOHOL Dose/Week | Pevious tobacco use<br>Pack tobacco years |
| Family medical history | Cancer family history<br>HeartDisease family history<br>HTN family history<br>DM family history<br>Stroke family history | Cancerspecific family history<br>Hdspecific family history<br>HTNspecific family history<br>Dmspecific family history<br>StrokeSpecific family history |
| Life sign parameters | Heart rate<br>ECG | blood pressure<br>EMG |

**Table 4.** One example of one-hot encoding of personal health data information.

| Index | Value | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 0–9 | 1.0 | 0.0 | 64.0 | 1.63 | 72.6 | 27.5 | 0.0 | 1.0 | 0.0 | 0.0 |
| 10–19 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 20–29 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 30–39 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 40–49 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 50–59 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 60–69 | 0.0 | 0.0 | 4.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

## 2.2. Experimental Design

### 2.2.1. Overall Architecture of Proposed Model

The network model proposed in this paper mainly includes two parts. In the first part, a convolutional neural network model based on VGG16 was built to extract the features of blood pressure, heart rate, EMG and ECG, because the VGG16 model has a better classification effect in image classification [24,25]. For feature extraction of personal health data, the one-dimensional conolutional neural network model was built. In the second part, a model for feature fusion was built to train the prediction model that can predict the occurrence probability of stroke. The entire network structure system is shown in Figure 2.

### 2.2.2. The First Part: Feature Extraction

The first model used a model based on convolutional neural networks to distinguish the waveforms of sign-of-life parameters of stroke patients. The model includes a VGG16 model; a fully connected layer with 256 neurons and the ReLU, which is used as the activation function; and a fully connected layer with a neuron and the softmax, which is used as the activation function. The work flow of this model is as follows: First, input the waveform graphs with the shape of 150×150×3 into the model to train the model. Features of the waveform graphs are extracted through the VGG16

model, and then input into the fully connected layer of the next layer to get a feature layer with a shape of 256×1 for next step.

The second model was built using a one-dimensional convolutional neural network model [26] to identify the personal health data of stroke patients. The model consists of two layers of one-dimensional convolutional layers containing 16 and 32 neurons, and ReLU is used as the activation function; two pooling layers—one fully connected layer containing a neuron, and sigmoid is used as the activation function. The work flow of this model is as follows: First, the convolutional layer and the pooling layer are used to convolve and pool the text to extract features, and then the features are input into the fully connection layer of the next layer, and the feature layer with a shape of 32×1 is obtained for the next step.

**Table 5.** The values of personal health data information before conversion.

| Factor | Value | Used One-Hot |
|---|---|---|
| Htn patient medical history | YES | NO |
| Age | 70 | NO |
| Alcohol Dose/Week | 0 | NO |
| Neuropathy autonomic symptoms | YES | NO |
| Previous Tobacco Use | YES | NO |
| Current Tobacco Use | NO | NO |
| HeartDisease family history | 1 | NO |
| HdspeciFIc family history | f | YES |
| Stroke year patient medical history | 0 | NO |
| Atrial FIbtrillation patient medical history | NO | NO |
| BMI | 26.7 | NO |
| Gender | F | NO |
| Painful feet autonomic symptoms | NO | NO |
| Syncope autonomic symptoms | NO | NO |
| cancSpec family history | NULL | NO |
| HTN years patient medical history | 4 | NO |
| DM patient history | 0 | NO |
| DmspeciFIc patient history | NULL | YES |
| DM patient medical history | NO | NO |
| Height/m | 1.64 | NO |
| Mass/kg | 71.67 | NO |
| Dizziness autonomic symptoms | NO | NO |
| Numbness autonomic symptoms | NO | NO |
| Pack years | 20 | NO |
| Previous alcohol use | YES | NO |
| HTN family history | 0 | YES |
| HTNspeciFIc family history | NULL | YES |
| Heart failure =CHF/ifaction=-MI patient medical history | NO | NO |
| Race | White | YES |
| DM Non-DM stroke | Non-DM | NO |
| OH autonomic symptoms | NO | NO |
| Cancer family history | 0 | YES |
| Cancer patient medical history | NO | NO |
| Stroke patient medical history | NO | NO |
| Stroke family history | 0 | NO |
| Stroke Specific family history | NULL | YES |
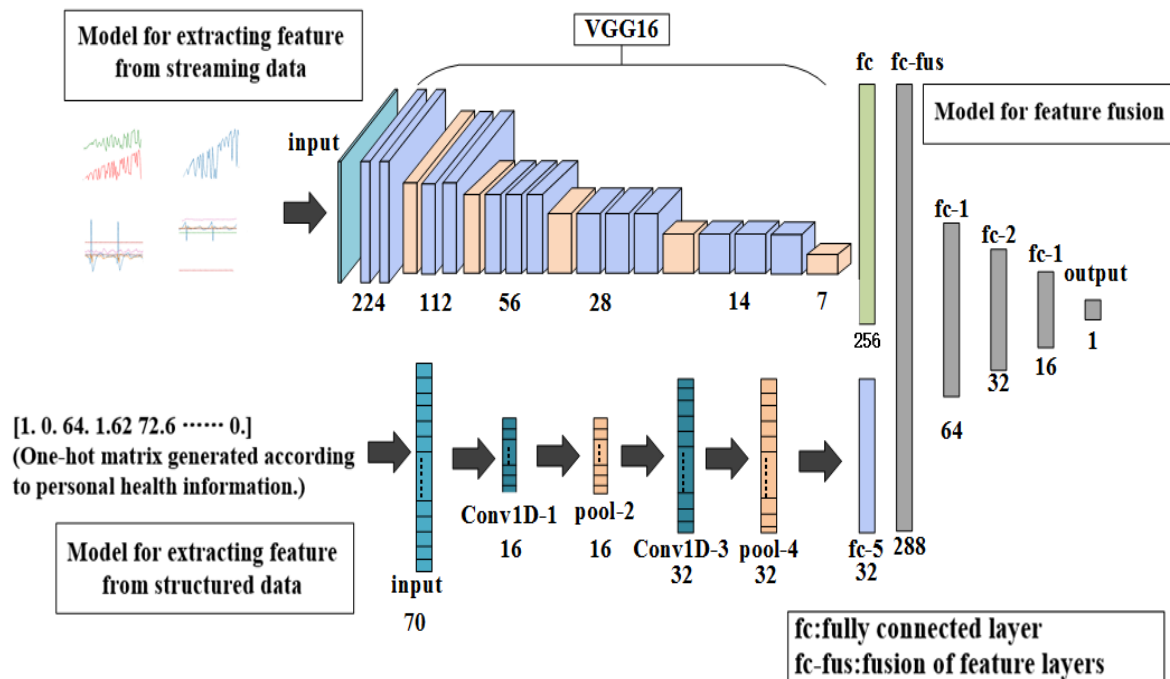
**Figure 2.** Overall architecture of proposed model.

### 2.2.3. The Second Part: Prediction of Incidence Probability of Stroke

The third model was composed of four layers of fully connected layer. The numbers of neurons in the first three layers of the fully connected layer were 64, 32 and 16. ReLU is used as the activation function. The number of neurons in the last layer was 1, and the activation function was softmax. The two models obtained from the training in first part were used to obtain feature layers with shapes of $256 \times 1$ and $32 \times 1$. Then the two feature layers were fused [27] to obtain the 189,679 one-dimensional dataset with a shape of $288 \times 1$, which was then used to train the model for predicting the incidence probability of stroke.

## 3. Results

During the experiment, graphs of sign-of-life parameters and 79 one-dimensional data with the shape of $70 \times 1$ were used to train the model for extracting the feature from stream data and the model for extracting the feature from structured data in the first part. In the second part, the two feature layers were combined as the training data input layer of the model for feature fusion. Finally, a prediction model was obtained to predict the probability of stroke.

### 3.1. Results of Training a Model for Extracting Features from Streaming Data

Graphs of sign-of-life parameters were divided into three sets, a training set (60%), a verification set (20%) and a testing set (20%). The training set and verification set were used to train the model for extracting features from streaming data. The optimizer was RMSProp, the learning rate was $1 \times 10^{-5}$, the loss function was binary cross entropy and the number of iterations of the training was 30. Finally, the model for extracting features from streaming data was able to identify the waveform graphs of sign-of-life parameters of stroke patients and normal persons. The curves of accuracy and loss rate of the feature extraction model are shown in Figures 3 and 4.

### 3.2. Results of Training a Model for Extracting Features from Structured Data

The one-dimensional dataset was divided into three sets, a training set (60%), a verification set (20%) and a testing set (20%). The training set and verification set were used to train the model for extracting features from structured data. The optimizer was RMSProp, the learning rate was $1 \times 10^{-4}$,

the loss function was binary cross entropy and the number of iterations of the training was 100. Finally, a classified prediction model that could identify personal health data of stroke patients and normal persons was obtained. The curves of accuracy and loss rate of the feature extraction model are shown in Figures 5 and 6.



**Figure 3.** Accuracy of the model for extracting features from streaming data during training.



**Figure 4.** Loss of the model for extracting features from streaming data during training.

### 3.3. Results of Training a Model for Feature Fusion

On the one hand, graphs of signs of life were input into the feature extraction model for streaming data; a 256×1 feature dataset was then obtained. On the other hand, the one-hot encoded personal health data of 79 subjects were input into the feature extraction model for structured data to obtain the 32×1 feature data. Then the feature layer of the sign-of-life parameters of the research subjects was combined with the corresponding feature layer of the personal health data to obtain a set of feature layers with a shape of 288×1, which were used as the training data of the model for feature fusion.

The combined feature layer set obtained above was divided into three categories: a training set (60%), a verification set (20%) and a testing set (20%). The training set and verification set were used to train the feature fusion model. The optimizer was Adam and the loss function was categorical cross entropy. The number of training iterations was 3000. Finally, a prediction model for feature fusion

capable of predicting the incidence probability of stroke was obtained. The curves of the accuracy and loss rate of the feature fusion model are shown in Figures 7 and 8.
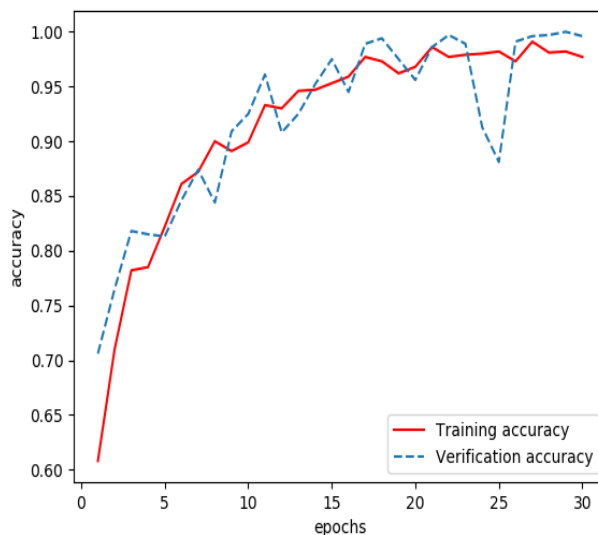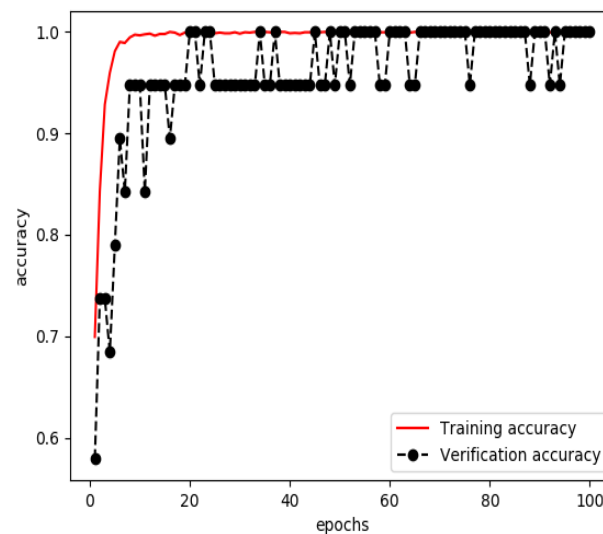


**Figure 5.** Accuracy of the model for extracting features from structured data during training.
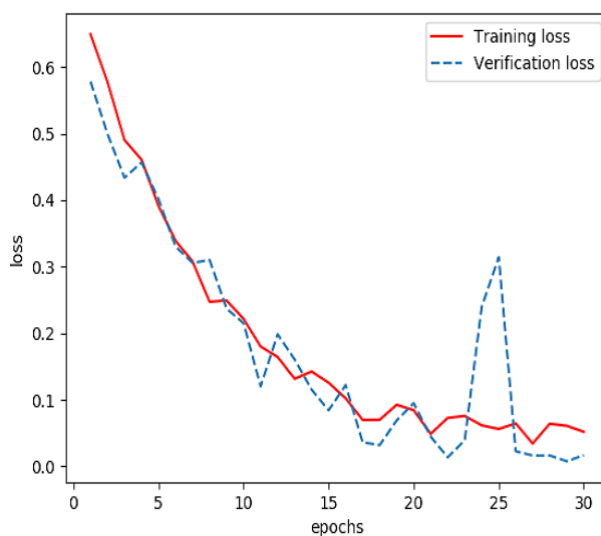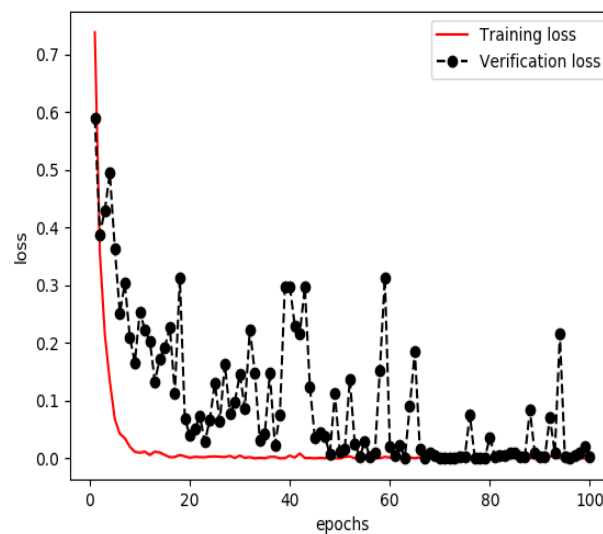


**Figure 6.** Loss of the model for extracting features from structured data during training.

*3.4. Model Evaluation*

The testing group containing 38,808 samples was used to evaluate the model for feature fusion, and the confusion matrix obtained is shown in Figure 9. Due to the value of AUC being up to 0.99, the prediction performance of the proposed model was proven, as shown in Figure 10. The index values of the model evaluation (precision, recall, accuracy, AUC and f1-score) are shown in Table 6; the precision, recall and f1-score were obtained by Equations (1)–(3). As os known, the accuracy rate is expressed as the proportion of positive samples predicted as positive samples. There are two cases for predicting positive samples. One case labeled as TP is to predict positive samples as positive samples, and the other labeled as FP is to predict negative samples as positive samples. In this paper, a positive sample represents a stroke patient and a negative sample represents a normal person. Obviously, the precision of the proposed model was 98.59%. In the other hand, the accuracy of the proposed model, which is expressed as the proportion of samples that are correctly predicted, was up to 98.53%. In short, these results justify the fact that the proposed model has a good performance in terms of

distinguishing the sign-of-life parameter waveforms and personal health information of stroke patients and normal people, and predicting the probability of a stroke patient, that is, the occurrence probability of stroke.
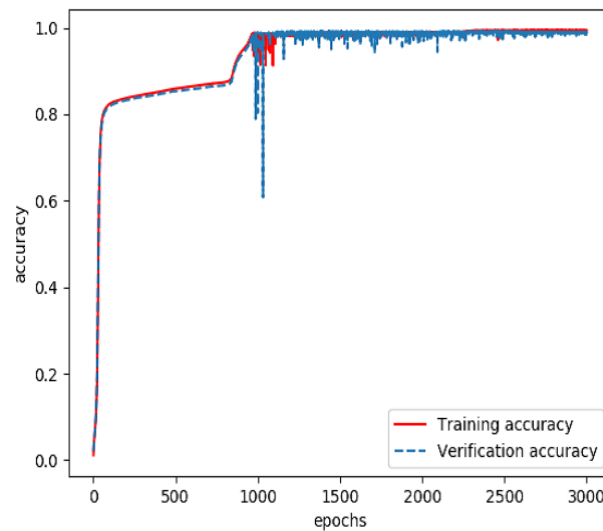
**Figure 7.** Accuracy of the model for feature fusion during training.
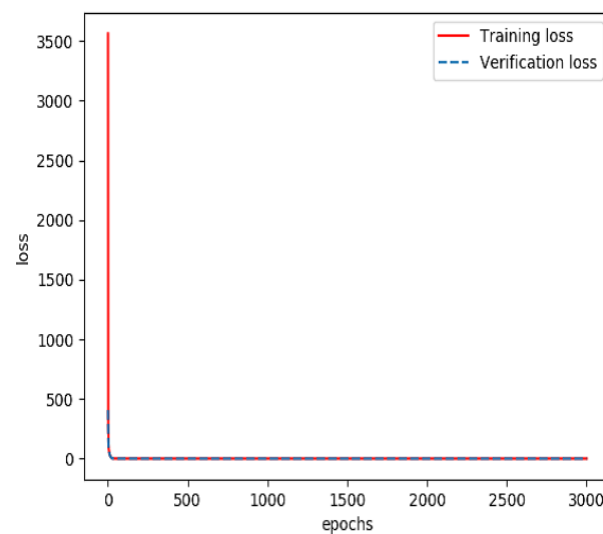
**Figure 8.** Loss rate of the model for feature fusion during training.

Once the predictive model recognizes that the sign-of-life parameters and personal health data of the current test subject have the characteristics of a stroke patient, the current test subject is judged to be a stroke patient. If the current test subject has not had a stroke, the result will be used as a pre-stroke warning. At this time, the subject should take corresponding preventive measures in time to reduce the harm caused by stroke.
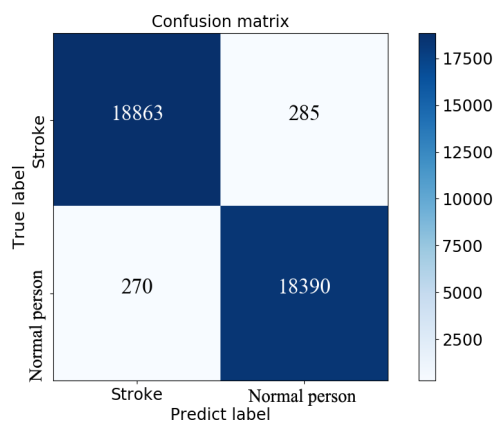
**Figure 9.** Confusion matrix obtained by model evaluation on the testing set.
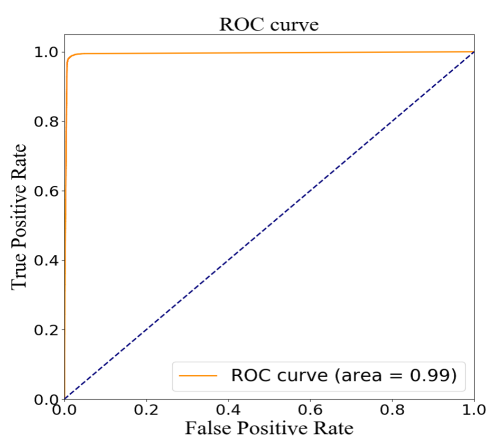


**Figure 10.** ROC curve by model evaluation on the testing set.

$$Precision = TP/(TP + FP) \tag{1}$$

$$Recall = TP/(TP + FN) \tag{2}$$

$$f1 - score = (2 \times Precision \times Recall)/(Precision + Recall) \tag{3}$$

**Table 6.** Evaluation index values of the model.

| TP | FN | FP | TN | Precision | Recall | Accuracy | AUC | *f*1-Score (0) | *f*1-Score (1) |
|---|---|---|---|---|---|---|---|---|---|
| 18863 | 285 | 270 | 18390 | 98.59% | 98.51% | 98.53% | 0.99 | 0.96 | 0.96 |

## 4. Discussion

In the model for extracting features from streaming data, the convolutional neural network model based on VGG16 was used to extract the features of ECG, EMG, blood pressure and heart rate to identify the waveform graphs of sign-of-life parameters of stroke patients. It should be noted that VGG16 did have perform better when extracting features compared with other models, such as VGG19, DenseNet201 and ResNet50, as shown in Figure 11 and Table 7. According to the results, the recognition accuracy of VGG19 was the worst. Further, the training time of VGG16 was the shortest compared with DenseNet201 and ResNet50. Meanwhile, the number of parameters was the least. Due to these reasons, VGG16 was selected as the basic network model for extracting features from the streaming data.

The multiple end-to-end network models proposed in this paper realized the feature fusion of multi-modal data and stroke prediction. We compared the method proposed in this paper with the current stroke prediction methods [10,11], as shown in Table 8. First, the method proposed in this paper has made perfect measures in terms of input data, changing from universal single-modal data to multi-modal data. Secondly, optimization was made on the network model, and a prediction model based on multi-model fusion was used to extract and fuse multi-modal data. Finally, a stroke prediction model with better classification performance than other methods was obtained. This model is used to identify the abnormal characteristics of stroke in the sign-of-life parameters and personal health data in time, so as to prepare for stroke prevention measures in advance to reduce the harm caused by stroke.
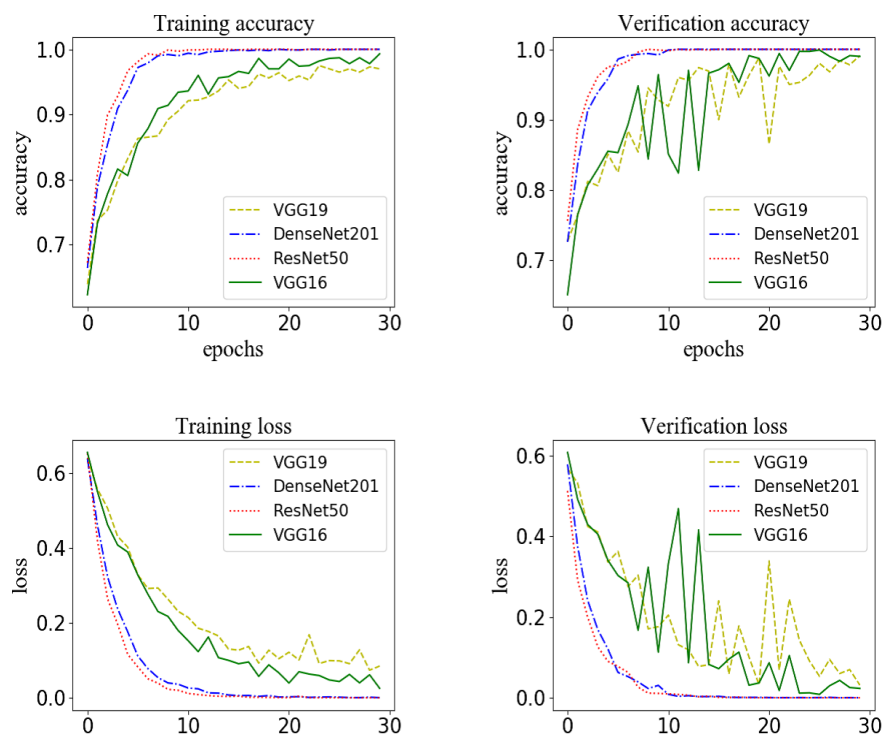


**Figure 11.** The model for extracting features from streaming data based on different network structures: the accuracy and loss for the training set and the accuracy and loss for the verification set.

**Table 7.** Comparison of training time, number of parameters and accuracy using different network structures on testing sets.

|  | Accuracy | Training Time (Second) | Total Parameters |
|---|---|---|---|
| VGG19 | 0.96 | 1678 | 122122049 |
| DenseNet201 | 0.97 | 34271 | 26186817 |
| ResNet50 | 0.97 | 20162 | 23638913 |
| VGG16 | 0.97 | 12689 | 16812353 |

**Table 8.** The method proposed in this paper with the current stroke prediction methods.

| Methods | Input Data | Model Structure | AUC |
|---|---|---|---|
| DNN with scaled PCA | Medical service use and health behavior data | DNN | 83.48% |
| Deep neural network | Electronic medical claims (EMCs) | DNN | 91.5% |
| Multi model | Streaming data (Blood pressure etc.), structured data (EHRs) | Multi model fusion | 99% |

## 5. Conclusions

The purpose of the current study was to estimate the probability of stroke occurrence. Hence, a convolutional neural network based on multi-model fusion was proposed. First, feature extraction of streaming data and structured data was carried out in combination with a convolutional neural network to expand ischaemic stroke-related factors and enhance feature extraction ability. Second, this paper proposed the processing of multi-modal data by multiple end-to-end neural architectures to achieve feature fusion and stroke prediction, and solved a major technical problem in disease prediction, which effectively improved upon the prediction accuracy of traditional models. To verify the effectiveness of the proposed method, the personal health data of 79 subjects were used in experiments that were carried out (based on a public dataset). The prediction accuracy reached 98.53%. This study contributes to our understanding of the impacts of risk factors on the occurrence of stroke. It could be used to help detect the disease early and thereby help institute appropriate control measures.

## Appendix A. Additional Tables

**Table A1.** Baseline information about subjects numbered S0030 to S208.

| Subject Number | HTN Patient Medical History | Group | Age | Height | Mass | BMI | Gender | Race | DM/ Non-DM STROKE |
|---|---|---|---|---|---|---|---|---|---|
| S0030 | YES | CONTROL | 64 | 1.6256 | 72.5747792 | 27.46365545 | F | White | Non-DM |
| S0064 | YES | CONTROL | 76 | 1.7018 | 68.0388555 | 23.49308018 | M | White | Non-DM |
| S0068 | NO | CONTROL | 79 | 1.5748 | 64.86370891 | 26.15477364 | F | White | Non-DM |
| S0121 | NO | CONTROL | 65 | 1.8288 | 72.5747792 | 21.69967838 | M | White | Non-DM |
| S0153 | NO | CONTROL | 71 | 1.7018 | 66.67807839 | 23.02321858 | F | White | Non-DM |
| S0154 | YES | CONTROL | 71 | 1.7526 | 80.73944186 | 26.28573518 | M | White | Non-DM |
| S0160 | NO | CONTROL | 72 | 1.8288 | 106.594207 | 31.87140263 | F | White | Non-DM |
| S0163 | YES | CONTROL | 73 | 1.651 | 84.82177319 | 31.11810921 | F | White | Non-DM |
| S0164 | NO | CONTROL | 60 | 1.651 | 63.5029318 | 23.29698015 | F | White | Non-DM |
| S0165 | YES | CONTROL | 75 | 1.7018 | 53.97749203 | 18.63784361 | F | White | Non-DM |
| S0166 | NO | CONTROL | 76 | 1.651 | 73.48196394 | 26.95793418 | M | White | Non-DM |
| S0172 | NO | CONTROL | 71 | 1.49 | 58.9670081 | 26.56051894 | F | Asian | Non-DM |
| S0174 | YES | CONTROL | 71 | 1.7018 | 67.58526313 | 23.33645965 | M | White | Non-DM |
| S0175 | YES | STROKE | 64 | 1.778 | 77.1107029 | 24.39220991 | M | White | STROKE |
| S0176 | NO | CONTROL | 68 | 1.778 | 90.718474 | 28.69671754 | M | White | Non-DM |
| S0183 | NO | CONTROL | 60 | 1.6256 | 64.41011654 | 24.37399422 | F | White | Non-DM |
| S0184 | NO | CONTROL | 68 | 1.7018 | 64.41011654 | 22.24011591 | F | White | Non-DM |
| S0185 | YES | STROKE | 72 | 1.778 | 77.1107029 | 24.39220991 | M | White | STROKE |
| S0187 | NO | CONTROL | 65 | 1.6002 | 60.78137758 | 23.73679105 | F | White | Non-DM |
| S0194 | NO | CONTROL | 64 | 1.778 | 95.2543977 | 30.13155341 | M | White | Non-DM |
| S0197 | NO | CONTROL | 65 | 1.5494 | 58.9670081 | 24.56303288 | F | White | Non-DM |
| S0199 | YES | STROKE | 77 | 1.7526 | 87.08973504 | 28.35315255 | M | White | STROKE |
| S0200 | NO | CONTROL | 70 | 1.8034 | 76.20351816 | 23.43100365 | M | White | Non-DM |
| S0203 | YES | CONTROL | 72 | 1.8034 | 70.30681735 | 21.61789027 | M | White | Non-DM |
| S0204 | YES | CONTROL | 80 | 1.7653 | 67.13167076 | 21.54221785 | M | White | Non-DM |
| S0205 | YES | STROKE | 63 | 1.6002 | 53.97749203 | 21.07968757 | F | Asian | STROKE |
| S0207 | YES | CONTROL | 75 | 1.6764 | 54.88467677 | 19.52971055 | F | AA | Non-DM |
| S0208 | NO | CONTROL | 67 | 1.651 | 70.30681735 | 25.79308517 | F | White | Non-DM |

Race: AA—African American.

**Table A2.** Baseline information about subjects numbered S0210 to S0402.

| Subject Number | HTN Patient Medical History | Group | Age | Height | Mass | BMI | Gender | Race | DM/ Non-DM STROKE |
|---|---|---|---|---|---|---|---|---|---|
| S0210 | NO | CONTROL | 72 | 1.6256 | 71.66759446 | 27.12035976 | F | White | Non-DM |
| S0212 | YES | CONTROL | 65 | 1.6002 | 53.97749203 | 21.07968757 | F | White | Non-DM |
| S0213 | NO | CONTROL | 71 | 1.8542 | 95.2543977 | 27.70587572 | M | White | Non-DM |
| S0214 | YES | STROKE | 62 | 1.7018 | 61.23496995 | 21.14377217 | F | White | STROKE |
| S0215 | YES | CONTROL | 61 | 1.6002 | 90.718474 | 35.42804634 | F | White | Non-DM |
| S0218 | NO | CONTROL | 70 | 1.6002 | 90.718474 | 35.42804634 | F | White | Non-DM |
| S0221 | NO | CONTROL | 71 | 1.6 | 54.88467677 | 21.43932686 | F | White | Non-DM |
| S0225 | NO | CONTROL | 66 | 1.524 | 64.41011654 | 27.73218897 | F | AA | Non-DM |
| S0227 | YES | CONTROL | 66 | 1.5494 | 67.13167076 | 27.9640682 | F | AA | Non-DM |
| S0228 | NO | CONTROL | 61 | 1.8034 | 79.37866475 | 24.40729546 | M | White | Non-DM |
| S0230 | YES | STROKE | 76 | 1.7272 | 72.5747792 | 24.32766712 | M | White | STROKE |
| S0231 | YES | STROKE | 64 | 1.6764 | 76.20351816 | 27.11563117 | M | White | STROKE |
| S0232 | YES | STROKE | 70 | 1.6256 | 63.5029318 | 24.03069852 | F | White | STROKE |
| S0239 | YES | STROKE | 77 | 1.7018 | 70.30681735 | 24.27618286 | M | White | STROKE |
| S0240 | NO | STROKE | 60 | 1.6764 | 99.7903214 | 35.50856463 | F | White | STROKE |
| S0242 | YES | CONTROL | 61 | 1.778 | 77.1107029 | 24.39220991 | M | AA | Non-DM |
| S0243 | YES | CONTROL | 62 | 1.7 | 77.11 | 26.6816609 | M | Asian | Non-DM |
| S0244 | NO | STROKE | 71 | 1.57 | 72.57 | 29.44135665 | F | AA | STROKE |
| S0247 | YES | STROKE | 79 | 1.695 | 72.25 | 25.14771017 | M | White | STROKE |
| S0248 | YES | STROKE | 80 | 1.68 | 69.85 | 24.74844104 | F | White | STROKE |
| S0277 | YES | STROKE | 67 | 1.6 | 99.6 | 38.90625 | F | AA | STROKE |
| S0305 | NO | CONTROL | 52 | 1.63 | 83.65 | 31.48406037 | M | Asian | Non-DM |
| S0321 | NO | STROKE | 54 | 1.75 | 112.35 | 36.68571429 | M | White | STROKE |
| S0322 | YES | STROKE | 78 | 1.61 | 65.9 | 25.42340187 | F | White | STROKE |
| S0324 | YES | STROKE | 62 | 1.7 | 84.5 | 29.23875433 | M | White | STROKE |
| S0332 | YES | STROKE | 73 | 1.67 | 66.5 | 23.84452652 | F | White | STROKE |
| S0334 | NO | STROKE | 59 | 1.57 | 63 | 25.5588462 | F | White | STROKE |
| S0337 | NO | STROKE | 67 | 1.68 | 78.45 | 27.7954932 | M | White | STROKE |
| S0340 | NO | STROKE | 50 | 1.68 | 75.85 | 26.87429138 | F | AA | STROKE |
| S0343 | YES | CONTROL | 66 | 1.63 | 66.2 | 24.91625579 | F | White | Non-DM |
| S0348 | YES | STROKE | 72 | 1.68 | 51.5 | 18.24688209 | M | white | STROKE |
| S0351 | YES | STROKE | 53 | 1.66 | 85.9 | 31.17288431 | M | white | STROKE |
| S0352 | NO | STROKE | 66 | 1.47 | 49.89 | 23.08760239 | F | white | STROKE |
| S0353 | YES | STROKE | 65 | 1.56 | 79.25 | 32.56492439 | F | white | STROKE |
| S0354 | NO | STROKE | 54 | 1.65 | 66.3 | 24.35261708 | F | white | STROKE |
| S0355 | YES | STROKE | 67 | 1.75 | 83.91 | 27.39918367 | M | AA | STROKE |
| S0358 | YES | STROKE | 80 | 1.57 | 61.2 | 24.82859345 | F | white | STROKE |
| S0361 | YES | STROKE | 71 | 1.74 | 81.45 | 26.90249703 | M | WHITE | STROKE |
| S0363 | YES | STROKE | 55 | 1.56 | 94.55 | 38.85190664 | F | white | STROKE |
| S0364 | NO | CONTROL | 63 | 1.8 | 106.55 | 32.88580247 | M | WHITE | Non-DM |
| S0371 | YES | STROKE | 66 | 1.67 | 96.2 | 34.49388648 | M | WHITE | STROKE |
| S0374 | YES | STROKE | 64 | 1.57 | 61 | 24.74745426 | F | WHITE | STROKE |
| S0376 | YES | CONTROL | 70 | 1.83 | 77.27 | 23.07324793 | M | WHITE | Non-DM |
| S0378 | NO | STROKE | 58 | 1.68 | 78.2 | 27.7069161 | M | WHITE | STROKE |
| S0379 | YES | STROKE | 58 | 1.74 | 75.38 | 24.89760867 | M | AA | STROKE |
| S0380 | YES | STROKE | 69 | 1.74 | 85.15 | 28.12458713 | M | WHITE | STROKE |
| S0388 | YES | STROKE | 61 | 1.62 | 67.9 | 25.8725804 | F | WHITE | STROKE |
| S0389 | YES | STROKE | 50 | 1.67 | 107.2 | 38.43809387 | F | WHITE | STROKE |
| S0397 | YES | STROKE | 74 | 1.83 | 90.7 | 27.08351996 | M | WHITE | STROKE |
| S0399 | NO | CONTROL | 51 | 1.81 | 68.75 | 20.98531791 | F | WHITE | Non-DM |
| S0402 | NO | STROKE | 54 | 1.82 | 96.45 | 29.11786016 | M | WHITE | STROKE |

Race: AA—African American.

**Table A3.** Baseline information about subjects numbered S0030 to S0343.

| Subject Number | Previous Tobacco Use | Current Tobacco Use | Pack Years Alcohol Use | Previous Dose(Week) | Alcohol Symptoms | Neuropathy Autonomic Symptoms |
|---|---|---|---|---|---|---|
| S0030 | NO | NO | 0 | YES | 0 | NO |
| S0064 | NO | NO | 0 | YES | 0 | NO |
| S0068 | NO | NO | 0 | NO | 0 | NO |
| S0121 | NO | NO | 0 | YES | 3 | NO |
| S0153 | NO | NO | 0 | NO | 0 | NO |
| S0154 | NO | NO | 0 | NO | 0 | NO |
| S0160 | NO | NO | 0 | NO | 0 | NO |
| S0163 | NO | NO | 0 | YES | 0 | NO |
| S0164 | YES | NO | 0 | YES | 15 | NO |
| S0165 | YES | NO | 0 | YES | 7 | NO |
| S0166 | NO | NO | 0 | NO | 0 | NO |
| S0172 | NO | NO | 0 | NO | 0 | NO |
| S0174 | NO | NO | 0 | NO | 0 | NO |
| S0175 | YES | NO | 35 | YES | 20 | NO |
| S0176 | YES | NO | 15 | YES | 5 | NO |
| S0183 | YES | YES | 41 | NO | 0 | NO |
| S0184 | NO | NO | 0 | YES | 0 | NO |
| S0185 | YES | NO | 60 | YES | 70 | NO |
| S0187 | NO | NO | 0 | YES | 0 | NO |
| S0194 | NO | NO | 0 | YES | 3 | YES |
| S0197 | NO | NO | 0 | NO | 0 | NO |
| S0199 | YES | NO | 56 | NO | 0 | NO |
| S0200 | NO | NO | 0 | NO | 0 | NO |
| S0203 | NO | NO | 0 | NO | 0 | NO |
| S0204 | NO | NO | 0 | YES | 7 | NO |
| S0205 | YES | NO | 0 | YES | 0 | NO |
| S0207 | NO | NO | 0 | YES | 1 | NO |
| S0208 | YES | NO | 27 | YES | 0 | NO |
| S0210 | YES | NO | 160 | NO | 0 | YES |
| S0212 | NO | NO | 0 | YES | 0 | YES |
| S0213 | YES | NO | 9 | YES | 0 | NO |
| S0214 | YES | NO | 28.57 | YES | 7 | NO |
| S0215 | YES | NO | 30 | YES | 0 | NO |
| S0218 | YES | YES | 11 | NO | 0 | YES |
| S0221 | NO | NO | 0 | YES | 7 | NO |
| S0225 | NO | NO | 0 | YES | 0 | NO |
| S0227 | NO | NO | 0 | YES | 1 | NO |
| S0228 | NO | NO | 0 | YES | 0 | NO |
| S0230 | YES | NO | 8.6 | YES | 4 | NO |
| S0231 | NO | NO | 0 | YES | 3 | NO |
| S0232 | NO | NO | 0 | YES | 2 | NO |
| S0239 | NO | NO | 0 | NO | 0 | NO |
| S0240 | YES | NO | 48 | NO | 0 | YES |
| S0242 | NO | NO | 0 | YES | 2 | NO |
| S0243 | YES | NO | 1 | YES | 1 | NO |
| S0244 | NO | NO | 0 | YES | 1 | NO |
| S0247 | YES | YES | 60 | YES | 2.5 | NO |
| S0248 | NO | NO | 0 | NO | 0 | NO |
| S0277 | NO | NO | 0 | YES | 0 | NO |
| S0305 | NO | NO | 0 | NO | 0 | NO |
| S0321 | NO | YES | 30.86 | YES | 7 | YES |
| S0322 | NO | NO | 0 | YES | 4 | NO |
| S0324 | YES | NO | 86 | YES | 0 | NO |
| S0332 | YES | NO | 30 | YES | 2 | NO |
| S0334 | YES | NO | 60 | NO | 0 | NO |
| S0337 | YES | NO | 66 | YES | 49 | NO |
| S0340 | YES | NO | 0.1667 | YES | 2 | NO |
| S0343 | YES | NO | 10.5 | YES | 0 | NO |

**Table A4.** Baseline information about subjects numbered S0030 to S0343.

| Subject Number | Dizziness Autonomic Symptoms | Numbness Autonomic Symptoms | Painful Feet Autonomic Symptoms | Syncope Autonomic Symptoms |
|---|---|---|---|---|
| S0030 | NO | NO | NO | NO |
| S0064 | NO | NO | NO | NO |
| S0068 | NO | NO | NO | NO |
| S0121 | YES | YES | NO | NO |
| S0153 | NO | NO | NO | YES |
| S0154 | NO | NO | NO | YES |
| S0160 | NO | NO | NO | NO |
| S0163 | NO | NO | NO | NO |
| S0164 | NO | NO | YES | NO |
| S0165 | NO | NO | NO | NO |
| S0166 | NO | YES | NO | NO |
| S0172 | NO | NO | NO | NO |
| S0174 | NO | NO | NO | NO |
| S0175 | YES | NO | NO | NO |
| S0176 | YES | NO | NO | NO |
| S0183 | NO | NO | NO | NO |
| S0184 | YES | NO | NO | NO |
| S0185 | YES | NO | NO | NO |
| S0187 | NO | NO | NO | NO |
| S0194 | YES | YES | YES | NO |
| S0197 | NO | NO | NO | NO |
| S0199 | NO | NO | NO | NO |
| S0200 | YES | NO | NO | NO |
| S0203 | NO | NO | NO | NO |
| S0204 | NO | NO | NO | NO |
| S0205 | NO | NO | NO | NO |
| S0207 | NO | NO | NO | NO |
| S0208 | NO | NO | NO | YES |
| S0210 | NO | NO | NO | NO |
| S0212 | NO | NO | NO | NO |
| S0213 | YES | NO | NO | NO |
| S0214 | NO | NO | NO | NO |
| S0215 | NO | NO | NO | NO |
| S0218 | YES | NO | NO | YES |
| S0221 | YES | NO | NO | YES |
| S0225 | NO | NO | NO | NO |
| S0227 | NO | NO | NO | NO |
| S0228 | NO | NO | NO | NO |
| S0230 | YES | NO | NO | NO |
| S0231 | YES | NO | NO | NO |
| S0232 | YES | NO | NO | NO |
| S0239 | NO | NO | NO | NO |
| S0240 | YES | YES | NO | NO |
| S0242 | NO | NO | NO | NO |
| S0243 | NO | NO | NO | NO |
| S0244 | NO | NO | NO | NO |
| S0247 | NO | NO | NO | NO |
| S0248 | NO | NO | NO | NO |
| S0277 | YES | NO | NO | NO |
| S0305 | NO | NO | NO | NO |
| S0321 | NO | YES | NO | NO |
| S0322 | YES | NO | NO | NO |
| S0324 | YES | NO | NO | NO |
| S0332 | NO | NO | NO | NO |
| S0334 | NO | NO | NO | NO |
| S0337 | YES | NO | NO | YES |
| S0340 | NO | NO | NO | NO |
| S0343 | NO | NO | NO | NO |

**Table A5.** Baseline information about subjects numbered S0030 to S0343.

| Subject Number | OH Autonomic Symptoms | Cancer Family History | CancSpec Family History | HeartDisease Family History | Hdspecific Family History | HTN Family History |
|---|---|---|---|---|---|---|
| S0030 | NO | 0 | | 0 | | 0 |
| S0064 | NO | 0 | | 2 | b | 0 |
| S0068 | NO | 3 | f, m, si | 0 | | 0 |
| S0121 | NO | 0 | | 0 | | 1 |
| S0153 | YES | 2 | gp, f | 1 | m | 0 |
| S0154 | YES | 0 | | 0 | | 0 |
| S0160 | YES | 0 | | 4 | gp | 0 |
| S0163 | NO | 0 | | 1 | f | 2 |
| S0164 | NO | 1 | f | 1 | gp | 0 |
| S0165 | YES | 0 | | 1 | f | 2 |
| S0166 | YES | 1 | gp | 0 | | 0 |
| S0172 | YES | 0 | | 0 | | 0 |
| S0174 | YES | 0 | | 0 | | 0 |
| S0175 | NO | 2 | f, si | 0 | | 0 |
| S0176 | YES | 3 | f, m, si | 1 | gp | 0 |
| S0183 | YES | 1 | m | 1 | f | 1 |
| S0184 | YES | 2 | gp, si | 1 | f | 1 |
| S0185 | YES | 0 | | 1 | f | 0 |
| S0187 | NO | 0 | | 0 | | 1 |
| S0194 | YES | 3 | f, m, si | 1 | b | 0 |
| S0197 | YES | 1 | f | 1 | f | 0 |
| S0199 | YES | 0 | | 0 | | 0 |
| S0200 | YES | 1 | si | 1 | f | 0 |
| S0203 | YES | 0 | | 1 | f | 1 |
| S0204 | YES | 0 | | 1 | f | 1 |
| S0205 | YES | 1 | m | 1 | gp | 0 |
| S0207 | NO | 0 | | 1 | m | 3 |
| S0208 | NO | 0 | | 1 | gp | 1 |
| S0210 | NO | 2 | gp,m | 1 | gp | 0 |
| S0212 | NO | 1 | gp | 1 | gp | 2 |
| S0213 | NO | 0 | | 1 | f | 0 |
| S0214 | NO | 2 | gp,si | 1 | m | 3 |
| S0215 | NO | 0 | | 0 | | 0 |
| S0218 | NO | 2 | gp,m | 1 | gp | 0 |
| S0221 | NO | 1 | gp | 1 | f | 1 |
| S0225 | NO | 0 | | 0 | | 1 |
| S0227 | NO | 0 | | 1 | m | 1 |
| S0228 | NO | 1 | m | 0 | | 0 |
| S0230 | NO | 0 | | 0 | | 0 |
| S0231 | YES | 3 | f, m, si | 1 | gp | 2 |
| S0232 | YES | 0 | | 1 | f | 0 |
| S0239 | NO | 0 | | 0 | | 0 |
| S0240 | YES | 0 | | 1 | b | 0 |
| S0242 | NO | 0 | | 0 | | 1 |
| S0243 | NO | 0 | | 1 | b | 1 |
| S0244 | NO | 1 | b | 1 | si | 1 |
| S0247 | NO | 0 | | 0 | | 0 |
| S0248 | NO | 0 | | 1 | f | 0 |
| S0277 | NO | 1 | gp | 1 | f | 2 |
| S0305 | NO | 0 | | 0 | | 0 |
| S0321 | NO | 0 | | 0 | | 1 |
| S0322 | YES | 1 | m | 1 | f | 2 |
| S0324 | NO | 0 | | 1 | gp | 1 |
| S0332 | NO | 0 | | 1 | f | 0 |
| S0334 | NO | 0 | | 1 | f | 0 |
| S0337 | NO | 2 | si,si | 1 | si | 0 |
| S0340 | NO | 0 | | 1 | gp | 1 |
| S0343 | NO | 2 | b,si | 1 | gp | 2 |

Family member: f—father, si—sister, m—mother, gp—grandparent, so—son, b—brother.

**Table A6.** Baseline information about subjects numbered S0030 to S0343.

| Subject Number | HTNspecific Family History | DM Family History | Dmspecific Family History | StrokeFAMILY Family History | StrokeSpecific Family History | HTN Years Patient Medical History |
|---|---|---|---|---|---|---|
| S0030 | | 0 | | 0 | | 4 |
| S0064 | | 0 | | 0 | | 0 |
| S0068 | | 0 | | 0 | | 0 |
| S0121 | b | 1 | f | 1 | f | 0 |
| S0153 | | 0 | | 0 | | 0 |
| S0154 | | 1 | f | 0 | | 2 |
| S0160 | | 0 | | 0 | | 0 |
| S0163 | f, m | 0 | | 0 | | 50 |
| S0164 | | 0 | | 0 | | 0 |
| S0165 | gp, m | 0 | | 1 | f | 4 |
| S0166 | | 2 | gp, b | 0 | | 0 |
| S0172 | | 0 | | 0 | | 0 |
| S0174 | | 0 | | 0 | | 15 |
| S0175 | | 0 | | 0 | | 0 |
| S0176 | | 0 | | 1 | si | 0 |
| S0183 | m | 1 | m | 0 | | 0 |
| S0184 | si | 0 | | 0 | | 0 |
| S0185 | | 0 | | 1 | m | 3 |
| S0187 | m | 0 | | 1 | m | 0 |
| S0194 | | 1 | gp | 0 | | 0 |
| S0197 | | 1 | m | 0 | | 0 |
| S0199 | | 0 | | 0 | | 0 |
| S0200 | | 0 | | 0 | | 0 |
| S0203 | f | 0 | | 1 | m | 8 |
| S0204 | so | 1 | so | 0 | | 6 |
| S0205 | | 0 | | 0 | | 0 |
| S0207 | gp,f,m | 2 | gp,m | 3 | f,b,si | 10 |
| S0208 | f | 0 | | 0 | | 0 |
| S0210 | | 1 | gp | 0 | | 0 |
| S0212 | gp,si | 0 | | 1 | si | 3 |
| S0213 | | 0 | | 0 | | 0 |
| S0214 | f,m,si | 1 | b | 1 | f | 0 |
| S0215 | | 0 | | 0 | | 3 |
| S0218 | | 1 | gp | 0 | | 0 |
| S0221 | m | 2 | gp,f | 0 | | 0 |
| S0225 | gp | 0 | | 1 | gp | 0 |
| S0227 | m | 0 | | 1 | m | 16 |
| S0228 | | 1 | f | 1 | f | 0 |
| S0230 | | 0 | | 0 | | 0 |
| S0231 | m, si | 1 | si | 2 | m, si | 6 |
| S0232 | | 3 | f, b, si | 0 | | 6 |
| S0239 | | 0 | | 0 | | 21 |
| S0240 | | 0 | | 1 | gP | 0 |
| S0242 | b | 0 | | 0 | | 25 |
| S0243 | b | 0 | | 0 | | 26 |
| S0244 | b | 4 | m,b,so,si | 1 | m | 0 |
| S0247 | | 0 | | 0 | | 1 |
| S0248 | | 0 | | 0 | | 2 |
| S0277 | f,m | 1 | m | 0 | | 24 |
| S0305 | | 1 | f | 0 | | 0 |
| S0321 | f | 0 | | 1 | f | 0 |
| S0322 | m,b | 0 | | 1 | b | 20 |
| S0324 | gp | 0 | | 0 | | 47 |
| S0332 | | 0 | | 0 | | 1 |
| S0334 | | 1 | gp | 0 | | 0 |
| S0337 | | 0 | | 1 | m | 0 |
| S0340 | m | 1 | f | 1 | m | 0 |
| S0343 | gp,m | 1 | gp | 0 | | 1 |

Family member: f—father, si—sister, m—mother, gp—grandparent, so—son, b—brother.

**Table A7.** Baseline information about subjects numbered S0030 to S0343.

| Subject Number | Cancer Patient Medical History | Stroke Patient Medical History | Stroke Years | Atrial Fibtrillation Patient Medical History | Heart Failure = CHF /Ifarction = -MI Patient Medical History | DM Patient Medical History |
|---|---|---|---|---|---|---|
| S0030 | NO | NO | 0 | NO | NO | NO |
| S0064 | NO | NO | 0 | NO | NO | NO |
| S0068 | NO | NO | 0 | NO | NO | NO |
| S0121 | NO | NO | 0 | NO | NO | NO |
| S0153 | NO | NO | 0 | NO | NO | NO |
| S0154 | NO | NO | 0 | NO | NO | NO |
| S0160 | YES | NO | 0 | NO | NO | NO |
| S0163 | NO | NO | 0 | NO | NO | NO |
| S0164 | NO | NO | 0 | NO | NO | NO |
| S0165 | NO | NO | 0 | NO | NO | NO |
| S0166 | NO | NO | 0 | NO | NO | NO |
| S0172 | NO | NO | 0 | NO | NO | NO |
| S0174 | NO | NO | 0 | NO | NO | NO |
| S0175 | NO | YES | 16 | NO | NO | NO |
| S0176 | NO | NO | 0 | NO | NO | NO |
| S0183 | NO | NO | 0 | NO | NO | NO |
| S0184 | NO | NO | 0 | NO | NO | NO |
| S0185 | NO | YES | 3 | NO | NO | NO |
| S0187 | NO | NO | 0 | NO | NO | NO |
| S0194 | NO | NO | 0 | NO | NO | NO |
| S0197 | NO | NO | 0 | NO | NO | NO |
| S0199 | YES | YES | 16 | NO | NO | NO |
| S0200 | NO | NO | 0 | NO | NO | NO |
| S0203 | NO | NO | 0 | NO | NO | NO |
| S0204 | YES | NO | 0 | NO | NO | NO |
| S0205 | NO | YES | 11 | YES | NO | NO |
| S0207 | NO | NO | 0 | NO | NO | NO |
| S0208 | NO | NO | 0 | NO | NO | NO |
| S0210 | NO | NO | 0 | NO | NO | NO |
| S0212 | NO | NO | 0 | NO | NO | NO |
| S0213 | YES | NO | 0 | NO | NO | NO |
| S0214 | NO | NO | 0 | NO | NO | NO |
| S0215 | NO | NO | 0 | NO | NO | NO |
| S0218 | NO | NO | 0 | NO | NO | NO |
| S0221 | NO | NO | 0 | NO | NO | NO |
| S0225 | NO | NO | 0 | NO | NO | NO |
| S0227 | NO | NO | 0 | NO | NO | NO |
| S0228 | NO | NO | 0 | NO | NO | NO |
| S0230 | YES | YES | 1 | YES | NO | NO |
| S0231 | YES | YES | 6 | NO | NO | YES |
| S0232 | NO | YES | 1 | NO | NO | NO |
| S0239 | NO | YES | 4 | NO | NO | NO |
| S0240 | NO | YES | 12 | NO | NO | NO |
| S0242 | YES | NO | 0 | NO | NO | NO |
| S0243 | NO | NO | 0 | NO | NO | NO |
| S0244 | NO | YES | 1 | NO | NO | NO |
| S0247 | YES | YES | 8 | NO | NO | NO |
| S0248 | NO | YES | 2 | NO | NO | NO |
| S0277 | NO | YES | 4 | NO | NO | NO |
| S0305 | NO | NO | 0 | NO | NO | NO |
| S0321 | NO | YES | 2 | NO | NO | NO |
| S0322 | NO | YES | 5 | NO | NO | NO |
| S0324 | NO | YES | 1 | NO | YES | NO |
| S0332 | NO | YES | 1 | NO | NO | NO |
| S0334 | NO | YES | 8 | NO | NO | NO |
| S0337 | NO | YES | 16 | NO | NO | NO |
| S0340 | NO | YES | 2 | NO | NO | NO |
| S0343 | NO | NO | 0 | NO | NO | YES |

**Table A8.** Baseline information about subjects numbered S0348 to S0402.

| Subject Number | Previous Tobacco Use | Current Tobacco Use | Pack Years Alcohol Use | Previous Dose(Week) | Alcohol Symptoms | Neuropathy Autonomic Symptoms |
|---|---|---|---|---|---|---|
| S0348 | YES | YES | 0 | YES | 0.5 | NO |
| S0351 | NO | NO | 0 | YES | 7 | NO |
| S0352 | YES | NO | 33 | YES | 2 | NO |
| S0353 | YES | NO | 15 | YES | 0.25 | YES |
| S0354 | YES | YES | 0 | YES | 3 | NO |
| S0355 | YES | NO | 10 | NO | 0 | YES |
| S0358 | NO | NO | 0 | YES | 1 | NO |
| S0361 | YES | NO | 42 | YES | 0 | NO |
| S0363 | YES | NO | 33 | YES | 0 | NO |
| S0364 | YES | NO | 12 | YES | 3 | NO |
| S0371 | YES | YES | 57 | YES | 42 | YES |
| S0374 | YES | NO | 10 | YES | 2 | NO |
| S0376 | YES | NO | 60 | YES | 3 | NO |
| S0378 | YES | YES | 96 | YES | 24 | YES |
| S0379 | YES | YES | 14 | YES | 0 | NO |
| S0380 | YES | NO | 70 | YES | 20 | YES |
| S0388 | YES | YES | 6.75 | YES | 0 | NO |
| S0389 | YES | NO | 24 | YES | 1 | NO |
| S0397 | YES | NO | 15 | YES | 7 | YES |
| S0399 | NO | NO | 0 | YES | 0 | NO |
| S0402 | YES | YES | 0 | YES | 4 | NO |

**Table A9.** Baseline information about subjects numbered S0348 to S0402.

| Subject Number | Dizziness Autonomic Symptoms | Numbness Autonomic Symptoms | Painful Feet Autonomic Symptoms | Syncope Autonomic Symptoms |
|---|---|---|---|---|
| S0348 | YES | NO | YES | NO |
| S0351 | NO | NO | NO | NO |
| S0352 | NO | YES | NO | NO |
| S0353 | YES | YES | NO | NO |
| S0354 | YES | NO | NO | NO |
| S0355 | NO | YES | NO | NO |
| S0358 | YES | NO | NO | NO |
| S0361 | NO | NO | NO | NO |
| S0363 | NO | NO | NO | NO |
| S0364 | NO | NO | NO | NO |
| S0371 | YES | YES | NO | YES |
| S0374 | NO | NO | NO | YES |
| S0376 | YES | NO | NO | YES |
| S0378 | YES | YES | YES | YES |
| S0379 | NO | YES | NO | YES |
| S0380 | NO | NO | YES | NO |
| S0388 | YES | NO | NO | NO |
| S0389 | NO | NO | NO | NO |
| S0397 | YES | YES | NO | YES |
| S0399 | NO | NO | NO | NO |
| S0402 | NO | YES | NO | NO |

**Table A10.** Baseline information about subjects numbered S0348 to S0402.

| Subject Number | OH Autonomic Symptoms | Cancer Family History | CancSpec Family History | HeartDisease Family History | Hdspecific Family History | HTN Family History |
|---|---|---|---|---|---|---|
| S0348 | NO | 1 | b | 1 | b | 0 |
| S0351 | NO | 2 | gp,f | 2 | gp | 3 |
| S0352 | NO | 1 | si | 1 | f | 0 |
| S0353 | NO | 0 | | 1 | m | 1 |
| S0354 | NO | 3 | gp, m, b | 1 | gp | 0 |
| S0355 | NO | 0 | | 1 | f | 0 |
| S0358 | YES | 2 | b, si | 0 | | 0 |
| S0361 | NO | 2 | gp, m | 1 | f | 0 |
| S0363 | NO | 2 | gp, f | 1 | b | 1 |
| S0364 | YES | 0 | | 1 | f | 1 |
| S0371 | NO | 1 | m | 1 | gp | 0 |
| S0374 | YES | 0 | | 3 | gp | 3 |
| S0376 | YES | 0 | | 3 | gp | 3 |
| S0378 | NO | 1 | gp | 1 | m | 0 |
| S0379 | NO | 5 | gp, m, si, so, d | 1 | f | 4 |
| S0380 | NO | 0 | | 1 | gp | 3 |
| S0388 | YES | 1 | f | 0 | | 1 |
| S0389 | NO | 1 | gp | 2 | f | 0 |
| S0397 | NO | 0 | | 1 | b | 3 |
| S0399 | NO | 1 | m | 0 | | 0 |
| S0402 | NO | 1 | m | 0 | | 0 |

Family member: f—father, si—sister, m—mother, gp—grandparent, so—son, b—brother.

**Table A11.** Baseline information about subjects numbered S0348 to S0402.

| Subject Number | HTNspecific Family History | DM Family History | Dmspecific Family History | StrokeFAMILY Family History | StrokeSpecific Family History | HTN Years Patient Medical History |
|---|---|---|---|---|---|---|
| S0348 | | 0 | | 0 | | 3 |
| S0351 | gp,f,m | 1 | f | 1 | gp | 12 |
| S0352 | | 0 | | 2 | gp, m | 0 |
| S0353 | b | 0 | | 2 | gp, f | 4 |
| S0354 | | 0 | | 0 | | 0 |
| S0355 | | 0 | | 0 | | 13 |
| S0358 | | 0 | | 0 | | 7 |
| S0361 | | 2 | gp, b | 1 | gp | 36 |
| S0363 | b | 3 | gp, m, f | 1 | b | 0 |
| S0364 | f | 1 | f | 0 | | 0 |
| S0371 | | 1 | d | 0 | | 4 |
| S0374 | gp, m, b | 0 | | 1 | gp | 3 |
| S0376 | gp, f, m | 0 | | 2 | gp, f | 4 |
| S0378 | | 2 | m, si | 0 | | 0 |
| S0379 | f, m, b, si | 0 | | 1 | si | 1 |
| S0380 | m,b,si | 0 | | 2 | f,m | 39 |
| S0388 | m | 1 | m | 1 | m | 1 |
| S0389 | | 0 | | 0 | | 1 |
| S0397 | f,m,b | 0 | | 2 | f,m | 0 |
| S0399 | | 0 | | 0 | | 0 |
| S0402 | | 0 | | 1 | f | 0 |

Family member: f—father, si—sister, m—mother, gp—grandparent, so—son, b—brother.

**Table A12.** Baseline information about subjects numbered S0348 to S0402.

| Subject Number | Cancer Patient Medical History | Stroke Patient Medical History | Stroke Years | Atrial Fibtrillation Patient Medical History | Heart Failure = CHF /Ifarction = -MI Patient Medical History | DM Patient Medical History |
|---|---|---|---|---|---|---|
| S0348 | YES | YES | 5 | NO | NO | NO |
| S0351 | NO | YES | 8 | NO | NO | NO |
| S0352 | NO | YES | 6 | NO | NO | NO |
| S0353 | NO | YES | 2 | NO | NO | NO |
| S0354 | NO | YES | 2 | NO | NO | NO |
| S0355 | NO | YES | 13 | NO | NO | NO |
| S0358 | NO | YES | 5 | NO | NO | NO |
| S0361 | NO | YES | 1 | NO | NO | NO |
| S0363 | NO | YES | 1 | NO | NO | NO |
| S0364 | YES | NO | 0 | NO | NO | NO |
| S0371 | NO | YES | 1 | NO | NO | NO |
| S0374 | YES | YES | 1 | NO | NO | NO |
| S0376 | NO | NO | 0 | YES | NO | NO |
| S0378 | NO | YES | 1 | NO | NO | NO |
| S0379 | NO | YES | 1 | NO | NO | NO |
| S0380 | NO | YES | 1 | NO | NO | NO |
| S0388 | NO | YES | 1 | NO | NO | NO |
| S0389 | NO | YES | 1 | NO | NO | NO |
| S0397 | NO | YES | 1 | YES | NO | NO |
| S0399 | NO | NO | 0 | NO | NO | NO |
| S0402 | NO | YES | 1 | NO | YES | NO |

## Appendix B. Additional Figures
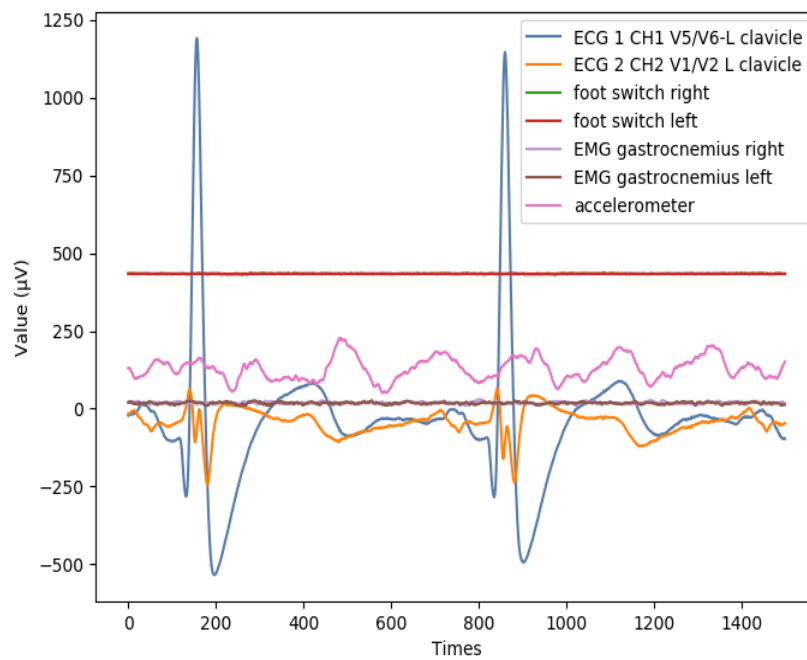


**Figure A1.** An example of ECG and EMG for a subject.
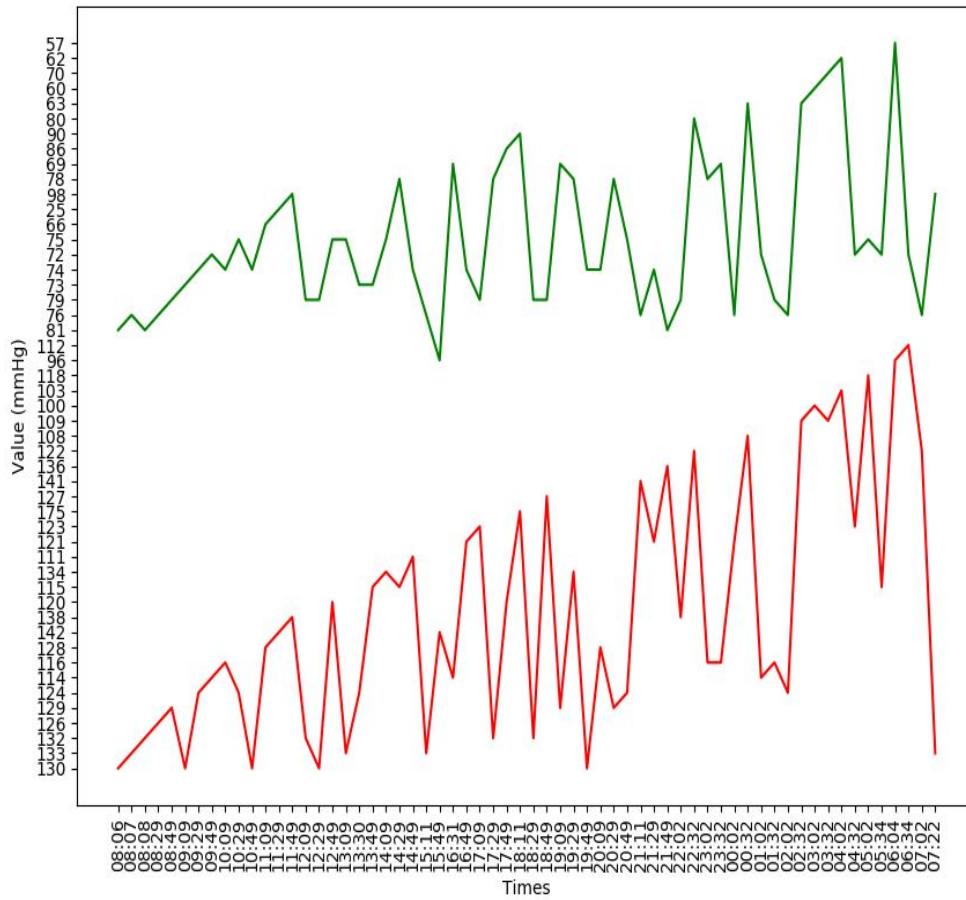
**Figure A2.** An example of the blood pressure curve of a subject in the experiment. The lower curve in the figure is the systolic pressure curve, and the upper curve is the diastolic pressure curve.
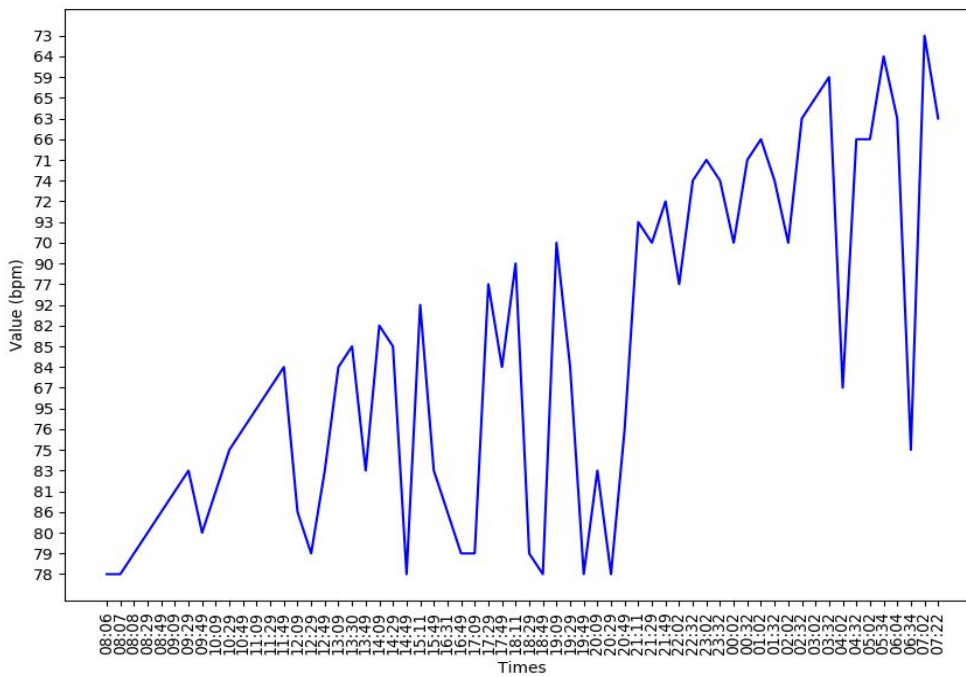


**Figure A3.** An example of the heart rate of a subject in the experiment.

## References

1. Mijajlovic, M.; Aleksic, V.; Sternic, N.; Mirkovic, M.; Bornstein, N. Role of prediabetes in stroke. *Neuropsychiatr. Dis. Treat.* **2017**, *13*, 259–267. [CrossRef]

2. Ackerley, S.J.; Stinear, C.M.; Barber, P.A.; Byblow, W.D. Priming sensorimotor cortex to enhance task-specific training after subcortical stroke. *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neuro-Physiol.* **2017**, *125*, 1451–1458. [CrossRef] [PubMed]

3. Michele, A.; Robol, E.; Fiaschi, A.; Manganotti, P. A double blind placebo RCT to investigate the effects of serotonergic modulation on brain excitability and motor recovery in stroke patients. *J. Neurol.* **2009**, *256*, 1152. [CrossRef]

4. Zhang, X.; Zhao, Q.; Zuo, Q.; Cheng, S.; Ma, J.; Liu, S.; Wang, Y. The development, valitidy and reliability of Stroke Premonitory Symptoms Alert Questionnaire. *Chin. Nurs. Manag.* **2015**, *15*, 416–419.

5. Goldstein, L.B.; Adams, R.; Alberts, M.J.; Appel, L.J.; Brass, C.D.; Culebras, A.; DeGraba, T.J.; Gorelick, P.B.; Guyton, J.R.; Hart, R.G.; et al. Primary prevention of ischemic stroke: a guideline from the American Heart Association/American Stroke Association Stroke Council: cosponsored by the Atherosclerotic Peripheral Vascular Disease Interdisciplinary Working Group; Cardiovascular Nursing Council; Clinical Cardiology Council; Nutrition, Physical Activity, and Metabolism Council; and the Quality of Care and Outcomes Research Interdisciplinary Working Group. *Circulation* **2006**, *113*, e873–e923. [CrossRef] [PubMed]

6. Liebeskind, D.S.; Albers, G.W.; Crawford, K.; Derdeyn, C.P.; George, M.S.; Palesch, Y.Y.; Toga, A.W.; Warach, S.; Zhao, W.; Brott, T.G.; et al. Imaging in strokenet: Realizing the potential of big data. *Stroke J. Cereb. Circ.* **2015**, *125*, 2000–2006. [CrossRef]

7. He, F.; Hu, Z.; Zhang, W.; Cai, L.; Cai, G.; Aoyagi, K. Construction and evaluation of two computational models for predicting the incidence of influenza in Nagasaki Prefecture, Japan. *Sci. Rep.* **2017**, *7*, 7192. [CrossRef]

8. Czabanski, R.; Horoba, K.; Wrobel, J.; Matonia, A.; Martinek, R.; Kupka, T.; Jezewski, M.; Kahankova, R.; Jezewski, J.; Leski, J.M. Detection of atrial fibrillation episodes in long-term heart rhythm signals using a support vector machine. *Sensors* **2020**, *20*, 765. [CrossRef]

9. Osman, A.H.; Alzahrani, A.A. New approach for automated epileptic disease diagnosis using an integrated self-organization map and radial basis function neural network algorithm. *IEEE Access.* **2019**, *7*, 4741–4747. [CrossRef]

10. Cheon, S.; Kim, J.; Lim, J. The use of deep learning to predict stroke patient mortality. *Int. J. Env. Res. Public. Health.* **2019**, *16*, 1876. [CrossRef]

11. Hung, C.; Chen, W.; Lai, P.; Lin, C.; Lee, C. Comparing deep neural network and other machine learning algorithms for stroke prediction in a large-scale population-based electronic medical claims database. In Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Seogwipo, South Korea, 11–15 July 2017.

12. Flint, A.C.; Conell, C.; Ren, X.; Banki, N.M.; Bhatt, D.L. Effect of systolic and diastolic blood pressure on cardiovascular outcomes. *N. Engl. J. Med.* **2019**, *381*, 243–251. [CrossRef] [PubMed]

13. Wesley, T.O.; Mohamed, F.A.; Waqas, T.Q.; Elsayed, Z.S. Electrocardiographic and echocardiographic left ventricular hypertrophy in the prediction of stroke in the elderly. *J. Stroke Cerebrovasc. Dis.* **2015**, *24*, 1991–1997. [CrossRef]

14. Ao, D.; Sun, R.; Tong, K.Y.; Song, R. Characterization of stroke- and aging-related changes in the complexity of EMG signals during tracking tasks. *J. Ann. Biomed. Eng.* **2015**, *43*, 990–1002. [CrossRef] [PubMed]

15. Bodapati, R.K.; Kizer, J.R.; Kop, W.J.; Kamel, H.; Stein, P.K. Addition of 24-hour heart rate variability parameters to the cardiovascular health study stroke risk score and prediction of incident stroke: The cardiovascular health study. *J. Am. Heart Assoc.* **2017**, *6*, e004305. [CrossRef] [PubMed]

16. Chantamit-O-Pas, P.; Goyal, M. Long short-term memory recurrent neural network for stroke prediction. *Springer Int. Publ.* **2018**, 312–323. [CrossRef]

17. Novak, V.; Hu, K.; Desrochers, L.; Novak, P.; Caplan, L. Cerebral flow velocities during daily activities depend on blood pressure in patients with chronic ischemic infarctions. *Stroke* **2009**, *41*, 61–66. [CrossRef] [PubMed]

18. Goldberger, A.L.; Amaral, L.A.N.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.-K.; Stanley, H.E. PhysioBank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. *Circulation* **2000**, *101*, e215–e220. [CrossRef]

19. Larsson, S.C.; Burgess, S.; Michaëlsson, K. Smoking and stroke: A mendelian randomization study. *Ann. Neurol.* **2019**, *86*, 468–471. [CrossRef]

20. Yang, S.; Zhao, Y.; Chen, X.; Lu, X.; Chen, Y.; Zhao, X.; Zhu, L.; Fang, Z.; Zhao, H.; Yao, Y.; et al. The ACTB Variants and Alcohol Drinking Confer Joint Effect to Ischemic Stroke in Chinese Han Population. *J. Atheroscler. Thromb.* **2019**, *27*. [CrossRef]

21. Shah, T.G.; Sutaria, J.M.; Vyas, M.V. The association between pulmonary hypertension and stroke: A systematic review and meta-analysis. *Int. J. Cardiol.* **2019**, *7*, S0167527319310757. [CrossRef]

22. Subha, P.P.; Pillai Geethakumari, S.M.; Athira, M.; Nujum, Z.T. Pattern and risk factors of stroke in the young among stroke patients admitted in medical college hospital, Thiruvananthapuram. *Ann. Indian Acad. Neurol.* **2015**, *18*, 20–23. [CrossRef]

23. Pham, D.H.; Le, A.C. Exploiting multiple word embeddings and one-hot character vectors for aspect-based sentiment analysis. *Int. J. Approx. Reason.* **2018**, *103*, S0888613X17304139. [CrossRef]

24. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroundy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; et al. Recent advances in convolutional neural networks. *arXiv* **2015**, arXiv:1512.07108.

25. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *Comput. Sci.* **2014**, arXiv:1409.1556.

26. Malek, S.; Melgani, F.; Bazi, Y. One-dimensional convolutional neural networks for spectroscopic signal regression. *J. Chemom.* **2017**, *5*, e2977. [CrossRef]

27. Eitel, A.; Springenberg, J.T.; Spinello, L.; Riedmiller, M.; Burgard, W. Multimodal deep learning for robust rgb-d object recognition. *Oalib J.* **2015**, 681–687. [CrossRef]