Bull. Egypt. Soc. Physiol. Sci. *Vol. (44), Issue (1)*, 17-28

Bull. of Egyp. Soc. Physiol. Sci.

(Official Journal of Egyptian Society for Physiological Sciences)

(pISSN: 1110-0842; eISSN: 2356-9514)

Potential Role of Logistic Regression Analysis to Identify Significant Risk Factors Associated with Stroke

Moutasem Salih Aboonq MD, Ph.D,

Department of Physiology, Taibah College of Medicine, Taibah University, Al-Madinah Al-Munawwarah, Saudi Arabia.

Submit Date : 17 Sept. 2023
Accept Date: 28 Oct. 2023

Keywords

- Stroke
- Logistic Regression
- Risk Factors

Abstract

Objectives**:** This research paper aims to clarify and analyse the various risk factors contributing to the occurrence of stroke in a specific population.***Material and methods*****:** This study employed a cross-sectional analysis of the 2015 Behavioral Risk Factor Surveillance System (BRFSS) dataset. The BRFSS is an annual telephone-based survey system designed to gather information about behavioural risk factors among adults across the United States. The dataset used in this study consisted of 70,692 observations obtained from the 2015 BRFSS. It included information on 21 potential risk factors and a binary outcome variable indicating the presence or absence of a stroke. The data analysis was conducted using Google Colab, a cloud-based platform that supports the programming language Python and its libraries. ***Results*****:** The logistic regression analysis revealed that the strongest associations with stroke were observed for heart disease or heart attack ($p <0.001$), high blood pressure ($p < 0.001$), high cholesterol ($p < 0.001$) and difficulties in walking ($p < 0.001$). Other risk factors that showed significant associations with stroke were diabetes, smoking, fruit consumption, vegetable consumption, general health perception, mental health, physical health, age, education and income. It is important to note that some risk factors, including cholesterol check, physical activity, access to healthcare and absence of doctor visits, did not exhibit statistically significant associations with stroke.***Conclusion*****:** The findings revealed that heart disease or heart attack, high blood pressure, high cholesterol and difficulties in walking exhibited the strongest associations with stroke.

Corresponding author: Moutasem S Aboonq M.D. Ph.D, Department of Medical Physiology, Taibah Faculty of Medicine, Taibah University, Al-Madinah Al-Munawwarah, Saudi Arabia 40001. Tel: +966-50-532-7177. E-mail: aboonq@yahoo.co.uk

Introduction

Stroke is a serious medical condition that poses a significant burden on individuals and healthcare systems worldwide. Understanding the risk factors associated with stroke is crucial for prevention, early detection and effective management. The present study explores the association between stroke and several potential risk factors, including heart disease or heart attack, age, high blood pressure, difficulties in walking, diabetes, high cholesterol, body mass index (BMI), fruit consumption, education, gender (sex), mental health, smoking, general health perception, heavy alcohol consumption, veggies consumption, physical health, income, cholesterol check, physical activity, any healthcare and no doctor because of cost (1). To determine the significance of these risk factors, logistic regression analysis was performed, considering the odds ratios, coefficients and p-values. The odds ratios provide insights into the strength of the association between each risk factor and stroke, while the coefficients and p-values help evaluate their statistical significance. Based on the analysis, the findings revealed that certain risk factors exhibited a stronger association with stroke. The risk factors with the highest coefficients and significant p-values included heart disease or heart attack, age, high blood pressure, difficulties in walking, diabetes, high cholesterol and BMI. These factors have consistently been identified in previous studies as significant contributors to stroke risk (2). However, it is important to note that other risk factors, although exhibiting relatively lower coefficients and p-values, should not be disregarded. Each individual risk factor may contribute to the overall risk of stroke and may have different impacts on various populations or subgroups. Therefore, a comprehensive understanding of all the identified risk factors is essential for a holistic approach to stroke prevention and management.

The aim of this research paper is to provide valuable insights into the relationship between these risk factors and stroke. By elucidating the significant associations, healthcare professionals, policymakers and individuals can gain a deeper understanding of the modifiable and non-modifiable risk factors for stroke. This knowledge can guide targeted interventions, public health campaigns and personalized healthcare strategies aimed at reducing the incidence and burden of stroke. In the subsequent sections of this paper, we will delve into the methodology employed, present the results of the analysis and discuss the implications and potential applications of the findings. A comprehensive exploration of these risk factors will contribute to the existing body of knowledge on stroke prevention and provide valuable insights for future research and clinical practice. Overall, this study sheds light on the complex interplay between various risk factors and stroke, highlighting the importance of a multidimensional approach to stroke prevention. By identifying and understanding these risk factors, we can pave the way for improved preventive strategies, early detection and enhanced patient care, ultimately striving towards a reduction in the incidence and impact of stroke on individuals and society. This research paper aims to clarify and analyse the various risk factors contributing to the occurrence of strokes in a specific population.

Materials and Methods

Study Design

In this research, a cross-sectional analysis was performed using secondary data obtained from the 2015 Behavioral Risk Factor Surveillance System (BRFSS), which is maintained by the Centers for Disease Control and Prevention (CDC) in the United States. The BRFSS database is freely accessible to the public online and is released under the CC0 1.0 Universal (CC0 1.0) Public Domain Dedication license.

Data Source

The dataset used in this study consisted of 70,692 observations obtained from the 2015 BRFSS. It included information on 21 potential risk factors, as well as a binary outcome variable indicating the presence or absence of stroke (Table 1). The data analysis was conducted using Google Colab, a cloud-based platform that supports the programming language Python and its libraries.

Table -1: Variables of risk factors

Variable	Definition
Diabetes_binary	0 (No Diabetes), 1 (Diabetes)
HighBP	0 (No High Blood Pressure), 1 (High Blood Pressure)
HighChol	0 (No High Cholesterol), 1 (High Cholesterol)
CholCheck	0 (No Check), 1 (Cholesterol Check Done)
BMI	1: Underweight, 2: Normal weight, 3: Overweight, 4: Obese
Smoker	0 (No), 1 (Yes)
Stroke	0 (No), 1 (Yes)
HeartDiseaseorAttack	0 (No), 1 (Yes)
PhysActivity	0 (No), 1 (Yes)
Fruits	0 (No), 1 (Yes)
Veggies	0 (No), 1 (Yes)
HvyAlcoholConsump	0 (No), 1 (Yes)
AnyHealthcare	0 (No), 1 (Yes)
NoDocbcCost	0 (No), 1 (Yes)
GenHlth	1 (Excellent), 2 (Very Good), 3 (Good), 4 (Fair), 5 (Poor)
MentHlth	0 to 30 (Number of Days)
PhysHlth	0 to 30 (Number of Days)
DiffWalk	0 (No Difficulty), 1 (Serious Difficulty)
Sex	0 (Female), 1 (Male)
Age	13-level category
Education	6-level category
Income	1: <$10 K, 2: $10–$15 K, 3: $15–$20 K, 4: $20–$25 K, 5: $25–$35 K, 6: $35–$50 K, 7: $50–$75 K, 8: >$75 K

Diabetes_binary: This variable indicates whether or not an individual has diabetes, with 0 indicating no diabetes, 1 indicating prediabetes and 2 indicating diabetes.HighBP: This variable indicates whether or not an individual has high blood pressure, with 0 indicating no high blood pressure and 1 indicating high blood pressure. HighChol: This variable indicates whether or not an individual has high cholesterol, with 0 indicating no high cholesterol and 1 indicating high cholesterol. CholCheck: This variable indicates whether or not an individual has had a cholesterol check in the past five years, with 0 indicating no check and 1 indicating that a check has been done. BMI: This variable indicates if an individual is (i) underweight, (ii) normal weight,

(iii) overweight or (iv) obese. Smoker: This variable indicates whether or not an individual has smoked at least 100 cigarettes in their lifetime, with 0 indicating no and 1 indicating yes. Stroke: This variable indicates whether or not an individual has ever had a stroke, with 0 indicating no and 1 indicating yes. Heart Disease Attack: This variable indicates whether an individual has had coronary heart disease with 0 indicating no and 1 indicating yes. PhysActivity: This variable indicates whether or not an individual has engaged in physical activity in the past 30 days (not including job-related activity), with 0 indicating no and 1 indicating yes. Fruits: This variable indicates whether or not an individual consumes fruit one or more times per day, with 0 indicating no and 1 indicating yes. Veggies: This variable indicates whether or not an individual consumes vegetables one or more times per day, with 0 indicating no and 1 indicating yes. HvyAlcoholConsump: This variable indicates whether or not an individual consumes heavy amounts of alcohol (adult men ≥ 14 drinks per week and adult women ≥ 7 drinks per week), with 0 indicating no and 1 indicating yes. AnyHealthcare: This variable indicates whether or not an individual has any kind of health care coverage, including health insurance and prepaid plans such as Health Maintenance Organization (HMO,) with 0 indicating no and 1 indicating yes. NoDocbcCost: This variable indicates whether or not an individual has not been able to see a doctor in the past 12 months because of cost, with 0 indicating no and 1 indicating yes.GenHlth: This variable is a scale from 1 to 5 indicating how an individual would rate their general health, with 1 indicating excellent, 2 indicating very good, 3 indicating good, 4 indicating fair and 5 indicating poor. MentHlth: This feature represents the number of days in the past 30 days that an individual reported poor mental health. The unique value is 31, indicating that the scale for this feature ranges from 0–30 days. PhysHlth: This feature represents the number of days in the past 30 days an individual has reported physical illness or injury. DiffWalk: This feature represents whether an individual has serious difficulty walking or climbing stairs, with 0 indicating no difficulty and 1 indicating serious difficulty. Sex: This feature represents the individual’s gender, with 0 indicating female and 1 indicating male. Age: This feature represents the individual’s age in a 13-level category. Education: This feature represents an individual’s education level. There are six possible levels of education.Income: This variable indicates how much annual household income is.

Data Analysis

The data analysis was performed using Colaboratory, an online Google platform with various Python libraries. Pandas was used for data manipulation, allowing us to preprocess and clean the dataset. Matplotlib was utilized for data visualization, enabling us to generate informative graphs and plots. Scipy was employed for statistical analysis, specifically for conducting logistic regression.

Logistic Regression Analysis

Logistic regression analysis was employed to model the association between stroke and the identified risk factors. The logistic regression model estimated the probability of stroke occurrence based on the presence or absence of various risk factors.

Model Estimation and Interpretation

The logistic regression model estimated the coefficients associated with each risk factor. The coefficient values represent the change in the log odds of stroke for a one-unit increase in the corresponding risk factor.

Results

As demonstrated in Table 2 and Figure 1. The logistic regression analysis revealed statistically significant associations between stroke and several risk factors. The strongest associations with stroke were observed for heart disease or heart attack (coef = 0.9853, $p < 0.001$), high blood pressure (coef = 0.4207, $p < 0.001$), high cholesterol (coef = 0.2316, $p < 0.001$), and difficulties in walking (coef = 0.5368, $p < 0.001$). Other risk factors that showed significant associations with stroke were diabetes (coef = 0.2171, $p < 0.001$), BMI (coef = -0.2154, $p < 0.001$), smoking (coef = 0.1634, $p < 0.001$), fruit consumption (coef = 0.0911, $p = 0.010$), vegetable consumption (coef = -0.2013, $p < 0.001$), general health perception (coef = 0.2443, $p < 0.001$), mental health (coef = 0.0083, $p < 0.001$), physical health (coef = 0.0070, $p < 0.001$), age (coef = 0.1336, $p < 0.001$), education (coef = 0.0508, $p = 0.003$), and income (coef = -0.0894, $p < 0.001$).

It Is Important to note that some risk factors,including cholesterol check, physical activity, access to healthcare, and absence of doctor visits due to cost, did not exhibit statistically significant associations with stroke.

The odds ratios further demonstrate the magnitude of these associations. The risk factors with the highest odds ratios for stroke were heart disease or heart attack (OR = 2.678611), high blood pressure (OR = 1.523022), high cholesterol (OR = 1.260623), and difficulties in walking (OR = 1.710512). Other significant risk factors included diabetes (OR = 1.242460), BMI (OR = 0.806249), smoking (OR = 1.177540), fruit consumption (OR = 1.095361), vegetable consumption (OR = 0.817663), general health perception (OR = 1.276671), mental health (OR = 1.008335), physical health (OR = 1.007052), age (OR = 1.142951), education (OR = 1.052161), and income (OR = 0.914484).

These findings emphasize the importance of addressing modifiable risk factors such as heart disease, high blood pressure, high cholesterol, and difficulties in walking in stroke prevention and management strategies.

Table-2: Logistic Regression Results

Variable	Coefficient	Standard Error	z-value	p-value	95% Confidence Interval
const	-5.2067	0.237	-21.938	0.000	[-5.672, -4.742]
Diabetes_binary	0.2171	0.040	5.365	0.000	[0.138, 0.296]
HighBP	0.4207	0.044	9.550	0.000	[0.334, 0.507]
HighChol	0.2316	0.038	6.140	0.000	[0.158, 0.306]
CholCheck	0.1220	0.154	0.793	0.428	[-0.179, 0.423]
BMI	-0.2154	0.041	-5.293	0.000	[-0.295, -0.136]
Smoker	0.1634	0.034	4.747	0.000	[0.096, 0.231]
HeartDiseaseorAttack	0.9853	0.036	27.401	0.000	[0.915, 1.056]
PhysActivity	-0.0105	0.036	-0.295	0.768	[-0.080, 0.059]
Fruits	0.0911	0.035	2.588	0.010	[0.022, 0.160]
Veggies	-0.2013	0.038	-5.235	0.000	[-0.277, -0.126]
HvyAlcoholConsump	-0.1436	0.105	-1.373	0.170	[-0.348, 0.061]
AnyHealthcare	-0.0170	0.087	-0.196	0.845	[-0.187, 0.153]
NoDocbcCost	0.0873	0.053	1.660	0.097	[-0.016, 0.190]
GenHlth	0.2443	0.021	11.615	0.000	[0.203, 0.285]
MentHlth	0.0083	0.002	4.556	0.000	[0.005, 0.012]
PhysHlth	0.0070	0.002	4.052	0.000	[0.004, 0.010]
DiffWalk	0.5368	0.040	13.400	0.000	[0.458, 0.615]
Sex	0.0613	0.035	1.747	0.081	[-0.007, 0.130]
Age	0.1336	0.008	16.818	0.000	[0.118, 0.149]
Education	0.0508	0.017	2.980	0.003	[0.017, 0.084]
Income	-0.0894	0.009	-10.049	0.000	[-0.107, -0.072]

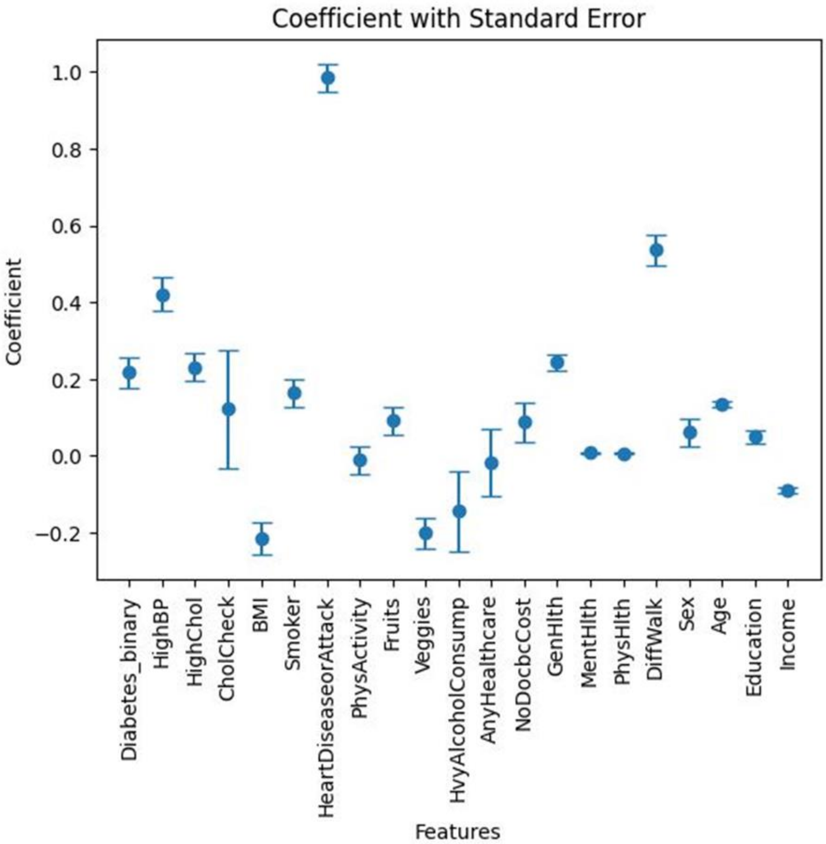

Figure 1: Risk Factors Associated with Stroke

Discussion

The present study aimed to clarify and analyse the association between stroke and various potential risk factors using logistic regression analysis of the 2015 BRFSS dataset. The results revealed several significant risk factors that were strongly associated with stroke, providing important insights into the prevention and management of this debilitating condition.

One of the key findings of this study was the strong association between heart disease or heart attack and stroke. Individuals with a history of heart disease or heart attack were found to be at a significantly higher risk of experiencing a stroke. This finding is consistent with previous research that identified cardiovascular diseases as important risk factors for stroke. The underlying mechanisms linking these conditions may include shared pathophysiological processes, such as atherosclerosis, hypertension and endothelial dysfunction. These results were supported by (3). These results emphasize the need for comprehensive management and control of cardiovascular risk factors to reduce the risk of stroke.

High blood pressure was also identified as a significant risk factor for stroke. Hypertension is a well-established risk factor for various cardiovascular diseases, including strokes. Elevated blood pressure can lead to damage to blood vessels in the brain, increasing the likelihood of stroke occurrence. The results of this study emphasize the importance of regular blood pressure monitoring and effective management of hypertension to reduce the burden of stroke. These results were supported by (4).

In addition to heart disease and high blood pressure, high cholesterol levels were found to be significantly associated with an increased risk of stroke. High cholesterol levels contribute to the development of atherosclerosis, a condition characterized by the buildup of plaque in the arteries. This plaque can restrict blood flow to the brain, leading to a higher risk of stroke. These findings highlight the importance of cholesterol management through lifestyle modifications, such as adopting a healthy diet and engaging in regular physical activity, as well as appropriate medication when necessary (5).

Furthermore, difficulties in walking were identified as a strong risk factor for stroke. This association may be explained by the fact that walking difficulties are often indicative of underlying neurological or musculoskeletal conditions that can increase the risk of stroke. Impaired mobility may be associated with a sedentary lifestyle, reduced physical activity and other comorbidities that contribute to stroke risk. These findings highlight the importance of promoting physical activity and mobility, as well as early identification and management of conditions affecting walking ability, to prevent strokes (6). Moreover, the practice of walking is vital for social factors and quality of life. Because walking is of great concern to patients and their families, outdoor walking is the cornerstone goal in managing and rehabilitating stroke patients (7).

Moreover, several other risk factors showed significant associations with stroke in this study. Diabetes was found to increase the risk of stroke, which aligns with previous research (8) highlighting the link between diabetes and cardiovascular diseases. Effective management of diabetes through glycaemic control and lifestyle modifications is crucial in reducing the risk of stroke. Diabetes is associated with an increased risk of cardiovascular complications, including stroke. A greater risk is observed among women than men. Several mechanisms associated with diabetes lead to stroke, including cardiac embolism, cerebral small vessel disease and large artery atherosclerosis. Hyperglycaemia increases the risk of a worse prognosis in people presenting with acute ischemic stroke compared with people with a euglycemic state. Moreover, people with diabetes may have poorer outcomes after a stroke and a higher risk of stroke recurrence than those without diabetes. Appropriate control of diabetes and other vascular risk factors may improve stroke outcomes and reduce the risk of recurrence. Secondary stroke prevention guidelines recommend screening for diabetes following a stroke. The diabetes medications pioglitazone and glucagon-like peptide-1 receptor agonists have demonstrated protection against stroke in randomized controlled trials; this protective effect is believed to be independent of glycaemic control. Specialists are often involved in the management of modifiable risk factors for stroke (hypertension, atrial fibrillation and hyperlipidaemia) but less often in the direct management of diabetes. This should aid neurologists in diabetes-related decision-making when treating people with acute or recurrent stroke (9).

BMI was also identified as a significant risk factor, with higher BMI values associated with an increased risk of stroke. This finding underscores the importance of maintaining a healthy weight through balanced nutrition and regular physical activity (10).

Smoking emerged as a significant risk factor for stroke, consistent with previous studies. Smoking can significantly increase the possibility of stroke due to blood vessel trauma, weakening blood vessels, or initiating plaque or clot formation. This finding emphasizes the urgent need for tobacco control measures and smoking cessation interventions to reduce the burden of stroke in the population (11). After the first stroke attack, persistent smoking increases the risk of stroke recurrence. There exists a dose–response relationship between the risk of stroke recurrence and smoking quantity (12).

The study also revealed that certain lifestyle factors, such as fruit consumption, vegetable consumption, general health perception, mental health and physical health, were significantly associated with stroke. In particular, a diet rich in fruits and vegetables is highly recommended because it meets the requirements without adding substantially to overall energy consumption. However, the consumption of low fruits and vegetables is high worldwide, especially in low- and middle-income countries. According to the World Health Organization, increasing individual vegetables and fruits consumption to ≥600 gm per day could decrease the possibility of ischemic stroke by 19% worldwide and 15% among members of the European Union (13). These findings highlight the importance of adopting a healthy lifestyle, including a balanced diet and regular exercise, to reduce the risk of stroke. They also underscore the significance of mental health and overall well-being in stroke prevention and management (14).

Age, education and income were identified as significant demographic factors associated with stroke. The incidence of stroke is directly proportional to advancing age. The chance of having a stroke doubles every 10 years after age 55. Although strokes are common among older adults, people younger than 65 years also have strokes (15). Advanced age has consistently been recognized as a major risk factor for stroke due to the accumulation of risk factors over time and the cardiovascular system with ageing. The complex network of the adult brain vasculature measures approximately 370 miles, receives about 20% of total cardiac output and exchanges 20% of total blood glucose and oxygen. With advanced age, both cerebral micro- and macro circulations undergo morphological and pathophysiological alterationsMicrocirculatory changes with aging may be mediated by impaired cerebral autoregulation, endothelial dysfunction and combined neurovascular factors. Whenever endothelial dysfunction initiates impaired cerebral autoregulation, neuro-inflammation may lead to microvascular injury, and impaired neurovascular coupling may promote a decline in cortical function, potentially resulting in future therapeutic interventions (16). Otherwise, ageing in healthy persons is associated with many observed changes in the human intracranial and extracranial cerebral arteries that predict the future risk for stroke (17). Higher education and income levels were found to be associated with a lower risk of stroke, potentially due to better access to healthcare, healthier lifestyle choices and greater awareness of stroke risk factors. Individuals with higher education and income may have better access to healthcare resources, including preventive screenings and early detection and management of risk factors, which can contribute to a lower risk of stroke (18).

Notably, some risk factors did not exhibit statistically significant associations with stroke in this study. Cholesterol checks, physical activity, access to healthcare and absence of doctor visits due to cost did not show significant associations with stroke. These findings may be influenced by various factors, including sample characteristics, measurement limitations and potential interactions among risk factors. Further research is needed to explore the relationship between these factors and stroke risk in more detail.

The strengths of this study include the utilization of a large and nationally representative dataset, allowing for robust analysis and generalizability of the findings to the adult population in the United States. The use of logistic regression analysis enabled the identification of significant risk factors and the estimation of their coefficients and odds ratios, providing valuable insights into the magnitude and direction of the associations.

However, several limitations should be considered when interpreting the results. *First*, the cross-sectional nature of the study design prevents the establishment of causality. The identified risk factors may be predictors or consequences of stroke, and longitudinal studies are necessary to elucidate temporal relationships. *Second*, reliance on self-reported data introduces the possibility of recall bias and social desirability bias, potentially affecting the accuracy and reliability of the obtained information. *Third*, unmeasured confounders and residual confounding may be present, which could influence the observed associations.

Despite these limitations, the findings of this study contribute to the existing body of knowledge on stroke risk factors and provide valuable insights for public health interventions and clinical practice. The identification of significant risk factors, such as heart disease or heart attack, high blood pressure, high cholesterol, difficulties in walking, diabetes, BMI, smoking and various lifestyle factors, underscores the importance of comprehensive stroke prevention strategies. Efforts should be directed towards promoting cardiovascular health, managing risk factors and raising awareness about stroke prevention among individuals at risk. Public health initiatives should focus on improving access to healthcare, implementing effective tobacco control measures and promoting healthy lifestyles through education and targeted interventions.

Future Work

While the present study has provided valuable insights into the risk factors associated with stroke using logistic regression analysis, there are opportunities for further research and the development of more advanced machine learning models to improve prediction and risk assessment. Here, we discuss potential avenues for future work in the field of stroke prediction using machine learning techniques.

Feature Selection and Model Optimization: Future studies can explore advanced feature selection techniques to identify the most relevant risk factors for stroke prediction. Techniques such as recursive feature elimination, genetic algorithms or least absolute shrinkage and selection operator; also Lasso or LASSO regularization can be employed to select a subset of the most informative features. Furthermore, different machine learning algorithms, such as support vector machines, random forests or neural networks, can be evaluated and optimized to improve the predictive performance of the models.

Incorporating Temporal Information: The current study utilized cross-sectional data, providing a snapshot of the risk factors and stroke status at a specific point in time. Future work could consider incorporating temporal information by utilizing longitudinal datasets that allow for the examination of risk factors over time and their dynamic relationship with stroke occurrence. Longitudinal data can provide a more comprehensive understanding of the temporal dynamics and trajectories of stroke risk.

Integration of Clinical and Imaging Data: While the present study focused on behavioural risk factors, future research can explore the integration of clinical data, such as medical history, laboratory results and imaging findings, to enhance the predictive models. Advanced imaging techniques, such a MRIor CT, can provide additional insights into the structural and functional changes in the brain associated with stroke risk.

Development of Risk Prediction Models: Building on the findings of this study, future research can aim to develop comprehensive risk prediction models that combine multiple risk factors, including both behavioural and clinical variables. These models can provide individualized risk estimates for stroke and aid in targeted interventions and preventive strategies. The use of machine learning algorithms, such as ensemble methods or deep learning, can facilitate the development of accurate and reliable risk prediction models.

External Validation and Generalizability: To ensure the generalizability and robustness of the developed models, future studies should consider external validation using independent datasets from different populations and geographical regions. This will help to evaluate the performance of the models in diverse populations and ensure their applicability in real-world settings.

Clinical Implementation and Decision Support Systems: Once reliable and accurate prediction models are developed, future work should focus on the implementation of these models in clinical practice. Decision support systems can be designed to assist healthcare providers in assessing individual stroke risk and providing tailored recommendations for prevention and management strategies.

Integration of Real-time Data and Wearable Devices: The advent of wearable devices and real-time monitoring technologies opens new possibilities for stroke prediction and prevention. Future research can explore the integration of data from wearable devices, such as heart rate monitors, activity trackers or blood pressure cuffs, to continuously monitor and update individual risk profiles. Real-time data integration can enable timely interventions and personalized preventive measures.

Conclusion

In conclusion, this study utilized logistic regression analysis to identify significant risk factors associated with stroke using the 2015 BRFSS dataset. The findings revealed that heart disease or heart attack, high blood pressure, high cholesterol and difficulties in walking exhibited the strongest associations with stroke. Other significant risk factors included diabetes, BMI, smoking, fruit consumption, vegetable consumption, general health perception, mental health, physical health, age, education and income. These findings contribute to the understanding of the risk factors associated with stroke and highlight the importance of targeted interventions and preventive strategies to reduce the burden of stroke in the population. Future research should consider longitudinal studies and explore additional potential risk factors to further enhance our understanding of stroke prevention and management.

Recommendations

Future work in the field of stroke prediction using machine learning techniques should focus on feature selection, model optimization, incorporation of temporal information, integration of clinical and imaging data, development of risk prediction models, external validation, clinical implementation and integration of real-time data. These efforts will contribute to improving stroke risk assessment, personalized interventions and, ultimately, the prevention and management of stroke.

Conflicts of Interest

The author 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.

References

1- **Mozzafarian D, Benjamin EJ, Go AS, et al**, on behalf of the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics–2016 update: a report from the American Heart Association. Circulation. 2016;133:e38–e360.

2- **Go AS, Mozaffarian D, Roger VL, et al**; American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics–2013 update: a report from the American Heart Association. Circulation. 2013; 127:e6–e245. doi: 10.1161/CIR.0b013e31828124ad..

3- **Connie W. Tsao, Aaron W. Aday, Zaid I. Almarzooq, Cheryl A.M. Anderson, Pankaj Arora, Christy L. Avery, Carissa M. Baker-Smith, Andrea Z. Beaton, Amelia K. Boehme, Alfred E. Buxton, Yvonne Commodore-Mensah, Mitchell S.V. Elkind, Kelly R. Evenson, Chete Eze-Nliam, Setri Fugar, Giuliano Generoso, Debra G. Heard, Swapnil Hiremath, Jennifer E. Ho, Rizwan Kalani, Dhruv S. Kazi, Darae Ko, Deborah A. Levine, Junxiu Liu, Jun Ma, Jared W. Magnani, Erin D. Michos, Michael E. Mussolino, Sankar D. Navaneethan, Nisha I. Parikh, Remy Poudel, Mary Rezk-Hanna, Gregory A. Roth, Nilay S. Shah, Marie-Pierre St-Onge, Evan L. Thacker, Salim S. Virani, Jenifer H. Voeks, Nae-Yuh Wang, Nathan D. Wong, Sally S. Wong, Kristine Yaffe and Seth S. Martin.** Heart Disease and Stroke Statistics—2023 Update: A Report from the American Heart Association. 2023; Vol. 147, No. 8.

4- **Mario R. Capecchi, PhD**. Hypertension, Opening Keynote Lecture. University of Utah School of Medicine. 2023

5- **Daniel G. Hackam and Robert A**. Hegele. Cholesterol Lowering and Prevention of Stroke. 2019; Vol. 50, No. 2.; 50:537–541.

6- **Kento Muto, Daijo Shiratsuchi, Kazuki Nanbu Hayato Sakamoto, Naohiro Furuya, Kazushi Nakamura, Mitani Yushi, Nako**

Tsujita and Hyuma Makizako . Ability to walk 10 m within the first week of stroke predicts independent outdoor walking and destination. August 2023; Volume 32, Issue 8, 107145,

7- **B Piernik-Yoder.** Rehabilitation outcomes of stroke patients with and without diabetes.Arch Phys Med Rehabil(2013).

8- **Ofri Mosenzon, Alice Yy Cheng, Alejandro A.** Rabinstein and Simona Sacco. Diabetes and Stroke: What Are the Connections? 2023 Jan;25(1):26-38. doi: 10.5853/jos.2022.02306. Epub 2023 Jan 3.

9- **Mohammed Alluhidan, reem f. Alsukait, Taghreed Alghaith, Meera Shekar, Nahar Alazemi, and Christopher h. Herbst.** Overweight and Obesity in Saudi Arabia; Consequences and Solutions. 2022; Volume 9 Issue 3 10.3390/healthcare9030311.

10- **Jingjing Chen, Shun Li, Kuo Zheng, Huaiming Wang, Yi Xie, Pengfei Xu, Zhengze Dai, Mengmeng Gu, Yaqian Xia, Min Zhao, Xinfeng Liu, and Gelin Xu.** Impact of Smoking Status on Stroke Recurrence. J Am Heart Assoc. 2019 Apr.; 16;8(8): e011696. doi: 10.1161/JAHA.118.011696.

11-**Kernan WN, Ovbiagele B, Black HR, Bravata DM, Chimowitz MI, Ezekowitz MD, Fang MC, Fisher M, Furie KL, Heck DV, Johnston SC, Kasner SE, Kittner SJ, Mitchell PH, Rich MW, Richardson D, Schwamm LH, Wilson JA;** on behalf of the American Heart Association Stroke Council, Council on Cardiovascular and Stroke Nursing, Council on Clinical Cardiology, and Council on Peripheral Vascular Disease. Guidelines for the prevention of stroke in patients with stroke and transient ischemic attack: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. 2014; Stroke. 45:2160–2236.

12-**Dan Hu, Junqian Huang, Yuchun Wang, Dongfeng Zhang and Yan Qu.** Fruits and Vegetables Consumption and Risk of Stroke. A Meta-Analysis of Prospective Cohort Studies. 2014https://doi.org/10.1161/STROKEAHA.114.004836Stroke. 2014; Vol. 45, No. 6, 45:1613–1619.

13-**Hall JN, Moore S, Harper SB, Lynch JW**. Global variability in fruit and vegetable consumption. Am J Prev Med. 2009; 36:402–409.e5.

14- **Mohammed Yousufuddin and Nathan Young**. Aging and ischemic stroke. Aging (Albany NY). 2019 May 15; v.11(9); 2542–2544. doi: 10.18632/aging.101931 2019 May 15 PMC6535078.

15-**Benjamin EJ, et al.** Circulation. 2018; 137: e67–492. 10.1161/CIR.0000000000000558

16- **Wen Xiuyun,Wu Qian, Xie Minjun, Li Weidong**, **and Liao Lizhen**. Education and stroke: evidence from epidemiology and Mendelian randomization study. Sci Rep. 2020 Dec 3; 10: 21208. doi: 10.1038/s41598-020-78248-8.