Arch Iran Med. December 2019;22(12):708-715

www.aimjournal.ir

ARCHIVES OF IRANIAN MEDICINE

Original Article

Open Access

Partitioning Stroke Patients, Determining Related Factors, and Comparing Derived Clusters Based on 12-Month Health Outcomes

Ali Soroush, MD[1]; Payam Sariaslani, MD[2]; Nadya Baharirad, BS[1]; Nasim Shams-Alizadeh, PhD[1]; Saeid Komasi, MSc[3*]

[1]Lifestyle Modification Research Center, Imam Reza Hospital, Kermanshah University of Medical Sciences, Kermanshah, Iran
[2]Department of Neurology, Imam Reza Hospital, Kermanshah University of Medical Sciences, Kermanshah, Iran
[3]Clinical Research Development Center, Imam Reza Hospital, Kermanshah University of Medical Sciences, Kermanshah, Iran

Abstract

Background: (*i*) Cluster analysis and partitioning samples based on cardio-cerebrovascular histories and length of stay (LOS); (*ii*) Determining related demographic and medical factors in individual clusters; and (*iii*) Comparing clusters based on 12-month health outcomes.

Methods: The statistical population of the study included 2,293 stroke patients hospitalized in Imam Reza hospital of Kermanshah city from January 1, 2015, to December 31, 2016. After a one-year follow-up, the data collection window was closed on December 31, 2017. The patients' data were extracted from the electronic hospital information system (HIS). Two-step cluster analysis (TSCA), chi-square, Fisher exact, Kruskal-Wallis, and Mann-Whitney U tests, as well as multinomial logistic regression analysis were the analysis methods.

Results: This model suggested five distinct clusters: the patients (*i*) without any cardio-cerebrovascular history and LOS = 5 days (36.2%); (*ii*) without any cardio-cerebrovascular history and LOS = 6 days (21.6%); (*iii*) with cerebrovascular history and LOS = 6 days (18.6%); (*iv*) with cardiovascular history and LOS = 6 days (16.1%); and (v) with cardio-cerebrovascular history and LOS = 6 days (7.5%). Hypertension, diabetes, and smoking were respectively the most significant modifiable risk factors, while sex, cerebrovascular diseases in the family, and age were respectively the most significant non-modifiable risk factors in high-risk clusters and LOS = 6 days. Compared to Cluster 1 (reference), during a one-year follow-up, a larger number of members in Clusters 3 and 5 were readmitted and/or expired.

Conclusion: Considering the modifiable risk factors identified in the current study, providing programs for preventing readmission and potential death caused by stroke for Clusters 3 and 5 seems essential.

Keywords: Clustering, Hospitalization, Medical history taking, Mortality, Patient readmission, Stroke

Cite this article as: Soroush A, Sariaslani P, Baharirad N, Shams-Alizadeh N, Komasi S. Partitioning stroke patients, determining related factors, and comparing derived clusters based on 12-month health outcomes. Arch Iran Med. 2019;22(12):708–715.

Received: March 17, 2019, Accepted: September 18, 2019, ePublished: December 1, 2019

Introduction

Stroke is one of the most serious health threats in the world, such that it is considered the second cause of mortality and the third cause of disability.[1,2] It imposes a heavy financial burden on the health system.[2] Although in recent decades, stroke and deaths from it are decreasing in the US, 975 thousand strokes occur annually in this country, leading to death in 57%–68% of cases.[2,3] Stroke is a common disease in Iran, as well[4]; recent reports show a prevalence of 139 per 100 000 people, leading to death in 11.4%–40.6% of cases.[5,6] The results of two review studies show that the prevalence of stroke in Iran ranges from 23 to 139 cases per 100 000 people.[7,8] These studies introduce factors such as aging, hypertension, diabetes, dyslipidemia, cardiovascular diseases (CVDs), and a history of stroke as the main risk factors for increasing the likelihood of stroke.[6,7,9]

Stroke, as a chronic disease, is sometimes accompanied by readmission and mortality. Based on recent reports, gender, aging, diabetes, history of stroke and/or cerebrovascular complications, CVDs, type of stroke, and longer length of stay (LOS) in the hospital are the most important measures for predicting readmission and mortality caused by stroke.[10,11] Besides histories related to cardio-cerebrovascular diseases, LOS is one of the factors related to mortality, and it can increase the likelihood of mortality 1.45 times.[12,13] Moreover, readmission increases the risk of mortality.[10,11]

While the abovementioned studies have addressed the risk factors of readmission and mortality among stroke patients, there are few studies on partitioning patients based on their cardio-cerebrovascular history as well as LOS. On the other hand, recent studies have shown that cluster analysis is a good method for categorizing and identifying high-risk stroke patients.[14,15] Therefore,

***Corresponding Author:** Saeid Komasi, MSc; Clinical Research Development Center, Imam Reza Hospital, Kermanshah University of Medical Sciences, Zakarya Razi Boulevard, Kermanshah, Iran; Tel: +988334276299, Fax: +988334276299, E-mail: S_komasi63@yahoo.com

systematic categorization of stroke patients for identifying groups in risk of readmission and mortality is likely to be effective in planning and delivering timely service to these patients. Based on these considerations, the current study was carried out in order to realize three objectives: (*i*) cluster analysis and partitioning the participants based on cardio-cerebrovascular history and LOS; (*ii*) determining related demographic and medical factors in individual clusters; and (*iii*) comparing clusters based on 12-month health outcomes.

Materials and Methods

Design and Context

In this existing data study, we evaluated stroke patients hospitalized in Imam Reza hospital of Kermanshah city (Iran) from 2015 to 2016. Kermanshah city, the capital of Kermanshah province in western Iran, is located 326 miles west of Tehran. According to the 2016 census, the total population of Kermanshah province was 1952434 people (988015 men).[16] More than 947000 people of this province are living in Kermanshah city. Imam Reza hospital of Kermanshah city is a general public hospital with 841 beds and more than 20 health wards. The neurology ward is one of the most active departments of this hospital, providing healthcare services for a large number of hospitalized stroke patients.

Participants and Data Collection

The statistical population of the study includes all patients diagnosed with stroke who were hospitalized in Imam Reza hospital of Kermanshah city from January 1, 2015, to December 31, 2016. The initial sample of this existing data study included 2347 individuals, whose information was extracted from the electronic hospital information system (HIS). Of these, 54 individuals were eliminated from the sample due to missing records. The final sample included 2293 patients. From February to April 2018, the patients' data were extracted by a nurse using the electronic HIS and recorded in the research forms designed for the current study. The electronic HIS includes the current medical records of the patients as well as all their outpatient and hospitalization records. In the next stage, the data related to the readmission of patients and in-hospital mortality within one year of their release from the hospital was recorded. Therefore, the data collection window for the study was closed on December 31, 2017. The data for the study were extracted under the supervision of a neurologist and the accuracy of the information for each patient was later rechecked by another nurse. The recorded data included gender, age, histories related to hypertension, diabetes, hyperlipidemia, smoking, drug addiction, CVDs, cerebrovascular diseases (CeVDs), and CeVDs in the family, LOS in the hospital, type of release (recovery/death), readmission, waiting time for readmission, LOS after readmission, and type of release after readmission (recovery/death).

Data Analysis

The data related to continuous variables were reported as mean and standard deviation, while the discrete data were reported as values (percentages) or median (interquartile range). Before performing the main analysis, the patients' records were first coded as no (=0)/yes (=1). The LOS score was entered into the analysis in the form of mean and standard deviation. In the main analysis, in order to identify the clusters, two-step cluster analysis (TSCA) was performed. This method was used due to the large size of the sample and the presence of continuous and discrete variables. TSCA determined the significance rank of the classification variables playing a role in predicting the model and identified the number of clusters automatically. The fitting of the model was achieved based on Schwarz's Bayesian information criterion (BIC) using the average silhouette coefficient. The silhouette coefficient is a measure of internal validity, which ranges from 0 to 1. A score closer to 1 indicates a better model.[17] Cluster analysis was performed using cerebrovascular history, cardiovascular history, and LOS. In the next stage, the clusters were compared with regard to cerebrovascular and cardiovascular histories using the Fisher exact test (because of the absence of the main assumption of the chi-square test). Due to the skewed distribution in the LOS, the clusters were compared using Kruskal-Wallis test.

In the next step, multinomial logistic regression analysis was performed for identifying the associated factors of the derived clusters. Gender and age along with all the other medical variables including the histories related to hypertension, diabetes, hyperlipidemia, smoking, addiction, and CeVDs in the family[6,7,9] were entered into the model simultaneously. Since five clusters were identified, Cluster 1 (healthier subjects) was considered as the reference cluster and analyses were adjusted for gender and age. The results of the analysis were presented in the form of adjusted odds ratios (OR) with 95% confidence intervals (CIs).

In the final stage, the type of patient discharge, readmission related to CeVDs and the type of discharge, the waiting period between initial hospitalization and readmission, and LOS after the readmission were explored in the total population and based on individual clusters. In order to evaluate the significance of the difference between individual clusters and the reference cluster, Mann-Whitney U test and one-variable chi-square test were used. In order to do this, the frequency proportion of each cluster compared to the reference cluster was calculated, and after assigning weights to the frequencies, the significance of the difference between the two clusters was computed. The discrete variables (waiting time for readmission and LOS in readmission) were compared using Mann-Whitney U test. All the statistical analyses were carried out using

SPSS20 (IBM Corp., Armonk, NY, USA). All the tests had two domains and statistical significance was defined as P value < 0.05.

Results

Identified Clusters

Table 1 depicts the results of the TSCA as well as a summary of the model. As can be seen, the silhouette measure of cohesion and separation is completely acceptable. All three factors have played a significant role in identifying the clusters. Based on the results in this table, there is a significant difference (P<0.001) between the clusters with regard to cerebrovascular and cardiovascular histories (since in three cells, frequencies are less than 25% of the total frequency, the Fisher exact test was used). In addition, there is a significant difference between the clusters with regard to the LOS in the hospital (Kruskal-Wallis test = 11.919; P = 0.018). This model has suggested five clusters, as follows: the patients (*i*) without any cardio-cerebrovascular history and LOS = 5 days (36.2%); (*ii*) without any cardio-cerebrovascular history and LOS = 6 days (21.6%); (*iii*) with cerebrovascular history and LOS = 6 days (18.6%); (*iv*) with cardiovascular history and LOS = 6 days (16.1%); and (*v*) with cardio-cerebrovascular history and LOS = 6 days (7.5%).

Demographic and Medical Factors Related to the Clusters

Table 2 presents the characteristics of the participants based on individual clusters. Moreover, this table depicts the results of multinomial logistic regression after adjustment for age and sex. The data in the table indicate that there is a significant relationship between sex, age, histories related to hypertension, diabetes, smoking, and history of CeVDs in the family and the clusters in the study. Compared to the reference cluster (Cluster 1), Clusters 2 (P < 0.001), 3 (P = 0.011), and 4 (P < 0.001) contain fewer women and more men. Moreover, Cluster 2 includes younger patients (P = 0.002), and Cluster 4 (P = 0.012) includes older patients. The results of this table also indicate that the CeVDs in the family (P = 0.013) and smoking (P < 0.001) in Cluster 2 are significantly higher compared to the patients in the reference cluster. Moreover, history of hypertension and diabetes in Clusters 3 (P = 0.008; P = 0.018), 4 (P = 0.001; P = 0.007), and 5 (P = 0.002; P = 0.024) are significantly higher compared to the reference cluster.

Comparing Clusters with Regard to 12-Month Health Outcomes

Table 3 compares clusters based on 12-month health outcomes. As can be seen, compared to the reference cluster, the frequency of readmission and mortality caused by it is higher in Clusters 3 (P = 0.021; P = 0.047) and 5 (P = 0.002; P = 0.040). In Clusters 2, waiting time for readmission is shorter than samples in the reference cluster (P = 0.049). There is no significant difference with regard to other variables between the other clusters and the reference cluster (P > 0.05).

Discussion

Main Findings

- Five clusters were identified based on cerebrovascular history, cardiovascular history, and LOS: two clusters without any cardio-cerebrovascular history and LOS less than or equal to six days, and three clusters with cardiovascular history, cerebrovascular history, or both and LOS equal to 6 days.
- The results show that 42.2% of the patients in the sample (Clusters 3, 4, and 5) have cardio-cerebrovascular history and LOS equal to 6 days. Only 36.2% of the participants (Cluster 1) do not have a history of the disease and have a LOS less than six days.
- The history of hypertension and diabetes, sex, smoking, history of CeVDs in the family, and age are

Table 1. Cardio-cerebrovascular History and LOS Profile Derived from Two-Stage Cluster Analysis (n = 2293)

Factors	Total (n = 2293)	Cluster 1 (n = 830; 36.2%) Patients without Cardio-Cerebrovascular History and LOS = 5 Day	Cluster 2 (n = 495; 21.6%) Patients without Cardio-Cerebrovascular History and LOS = 6 Day	Cluster 3 (n = 427; 18.6%) Patients with Cerebrovascular History and LOS = 6 Day	Cluster 4 (n = 370; 16.1%) Patients with Cardiovascular History and LOS = 6 Day	Cluster 5 (n = 171; 7.5%) Patients with Cardio-cerebrovascular History and LOS = 6 Day	*P* Value
Cerebrovascular history (%)[a]	598 (26.1)	0 (0)	0 (0)	427 (100)	0 (0)	171 (100)	<0.001
Cardiovascular history (%)[a]	541 (23.6)	0 (0)	0 (0)	0 (0)	370 (100)	171 (100)	<0.001
Length of stay, day (median & interquartile range)[b]	6 (4-10)	5 (3-9)	6 (4-10)	6 (4-11)	6 (4-10)	6 (4-10)	0.018

Summary of model: Silhouette measure of cohesion and separation is 0.7; Ratio of sizes for largest to the smallest cluster is 4.85; the most important predictors are: cerebrovascular and cardiovascular history = 1.0, Length of stay = 0.9.
NA, Not applicable.
[a] Considering that in three cells, frequencies are less than 25% of the total frequency (=zero), the Fisher's exact test was used.
[b] Kruskal-Wallis test.

Table 2. Results of Multinomial Logistic Regression for Identifying Correlates

Predictors	Total (n = 2293)	Cluster 1 (n = 830; 36.2%) Reference		Cluster 2 (n = 495; 21.6%)		Cluster 3 (n = 427; 18.6%)		Cluster 4 (n = 370; 16.1%)		Cluster 5 (n = 171; 7.5%)	
		Patients without Cardio-Cerebrovascular History and LOS = 5 Day		Patients without Cardio-Cerebrovascular History and LOS = 6 Day		Patients with Cerebrovascular History and LOS = 6 Day		Patients with Cardiovascular History and LOS = 6 Day		Patients with Cardio-Cerebrovascular History and LOS = 6 Day	
Sex, female (%)	1070 (46.7)	409 (49.3)	1	225 (45.5)	0.13 [0.04 – 0.36] $P < 0.001$	178 (43.8)	0.20 [0.06 – 0.69] $P = 0.011$	168 (45.4)	0.10 [0.03 – 0.38] $P < 0.001$	81 (47.4)	0.38 [0.05 – 2.55] $P = 0.316$
Age, y (Mean SD)	67.3 ± 14.3	65.4 ± 15.2	1	65.1 ± 15.5	0.98 [0.97 – 0.99] $P = 0.002$	69.4 ± 12.5	1.0 [1.0 – 1.2] $P = 0.163$	70.0 ± 12.2	1.0 [0.99 – 1.02] $P = 0.323$	71.3 ± 11.4	1.02 [1.0 – 1.04] $P = 0.012$
Medical history (%)											
HTN	1358 (59.2)	446 (53.7)	1	263 (53.1)	1.0 [0.79 – 1.27] $P = 0.996$	276 (64.6)	1.41 [1.10 – 1.82] $P = 0.008$	249 (67.3)	1.56 [1.19 – 2.05] $P = 0.001$	124 (72.5)	1.83 [1.25 – 2.67] $P = 0.002$
DM	537 (23.4)	162 (19.5)	1	107 (21.6)	1.19 [0.89 – 1.59] $P = 0.235$	113 (26.5)	1.42 [1.06 – 1.89] $P = 0.018$	102 (27.6)	1.51 [1.12 – 2.04] $P = 0.007$	53 (31.0)	1.57 [1.06 – 2.31] $P = 0.024$
HLP	228 (9.9)	83 (10.0)	1	41 (8.3)	0.81 [0.54 – 1.21] $P = 0.296$	44 (10.3)	0.93 [0.62 – 1.39] $P = 0.722$	32 (8.6)	0.74 [0.48 – 1.16] $P = 0.189$	28 (16.4)	1.47 [0.91 – 2.39] $P = 0.117$
Cerebrovascular disease in family	30 (1.3)	19 (2.3)	1	2 (0.4)	0.16 [0.4 – 0.68] $P = 0.013$	4 (0.9)	0.40 [0.13 – 1.20] $P = 0.102$	3 (0.8)	0.37 [0.11 – 1.27] $P = 0.114$	2 (1.2)	0.50 [0.11 – 2.23] $P = 0.366$
Smoking	210 (9.2)	59 (7.1)	1	68 (13.7)	2.09 [1.38 – 3.17] $P < 0.001$	41 (9.6)	1.19 [0.75 – 1.89] $P = 0.467$	27 (7.3)	1.01 [0.60 – 1.70] $P = 0.958$	15 (8.8)	1.10 [0.57 – 2.12] $P = 0.781$
Drug addiction	197 (8.6)	60 (7.2)	1	48 (9.7)	0.93 [0.59 – 1.46] $P = 0.757$	44 (10.3)	1.37 [0.87 – 2.17] $P = 0.173$	26 (7.0)	0.97 [0.57 – 1.63] $P = 0.898$	19 (11.1)	1.69 [0.92 – 3.11] $P = 0.093$

HTN, hypertension; DM, diabetes mellitus; HLP, hyperlipidemia.

The socio-demographics and other factors in this table were all included as covariates in the generation of the multinomial logistic regression model. The results were adjusted for age and sex (using interaction effects between the two variables on the model). Boldface indicates statistical significance ($P < 0.05$).

Summary of model: The model fitting information is: Chi-square = 170.805, $P < 0.0005$; Pseudo R-square based on McFadden and Nagelkerke = 0.025 to 0.076.

Table 3. The 12-Month Health Consequences for Separate Clusters

Components	Total (n = 2293)	Cluster 1(n = 830; 36.2%) Reference	Cluster 2 (n = 495; 21.6%)		Cluster 3 (n = 427; 18.6%)		Cluster 4 (n = 370; 16.1%)		Cluster 5 (n = 171; 7.5%)	
		Patients without Cardio-Cerebrovascular History and LOS = 5 Days	Patients without cardio-Cerebrovascular History and LOS = 6 Days	*P* Value	Patients with Cerebrovascular History and LOS = 6 Days	*P* Value	Patients with Cardiovascular History and LOS = 6 Days	*P* Value	Patients with Cardio-cerebrovascular History and LOS = 6 Days	*P* Value
Hospital discharge (%)[a]										
Improved	1936 (84.4)	705 (84.9)	407 (82.2)	0.816	367 (85.9)	0.939	311 (84.1)	0.939	146 (85.4)	0.999
Dead	357 (15.6)	125 (15.1)	88 (17.8)	0.602	60 (14.1)	0.480	59 (15.9)	0.715	25 (14.6)	0.999
Hospital readmission (%)[a]	104 (4.5)	17 (2.0)	10 (2.0)	0.999	44 (10.3)	0.021	8 (2.2)	0.999	25 (14.6)	0.002
Readmission discharge (%)[a]										
Improved	112 (4.9)	13 (1.6)	12 (2.4)	0.999	50 (11.7)	0.008	13 (3.5)	0.414	24 (14.0)	0.003
Dead	22 (1.0)	4 (0.5)	2 (0.4)	0.999	10 (2.3)	0.047	3 (0.8)	0.999	3 (1.8)	0.040
Waiting time for readmission, day (Median & interquartile range)[b]	72 (19.5–191)	76 (17.5–260)	21 (6.5–50)	0.049	92.5 (20.25–194.25)	0.854	93.5 (27.75–315.25)	0.652	85 (31–189)	0.971
LOS in readmission, day (Median & interquartile range)[b]	6 (3–10.25)	6 (3–12)	3 (1.75–9.25)	0.328	5 (3-10)	0.805	10.5 (6.5–16.75)	0.056	5 (4–10)	0.856

LOS, length of stay.

[a] *P* value of chi-square test; [b] *P* value of Mann-Whitney U test for comparing all the groups with the reference group; Boldface indicates statistical significance ($P < 0.05$) compared to the reference group.

respectively the most significant predictors for high-risk clusters and LOS equal to 6 days.

- During a one-year follow-up, a larger number of members of Clusters 3 and 5, compared to the reference cluster, were readmitted and died after readmission. Moreover, samples of Cluster 2 had a shorter waiting time for readmission.

In the current study, five clusters were identified and suggested based on cardio-cerebrovascular histories and LOS. In line with previous studies,[14,15] our study was able to successfully categorize the patients into different groups, providing valuable consistent information. Van Rheenen et al[18] were able to successfully classify stroke patients based on the type of stroke, mortality, and risk factors. Partitioning the stroke risk factors and its related family history can be effective in identifying individuals vulnerable to the risk of stroke.[19]

The results of the current study show that hypertension, diabetes, and smoking are the most significant modifiable predictors for risky clusters and LOS equal to 6 days. Recent original and review studies have comprehensively highlighted the importance of these modifiable factors and their related mechanisms for preventing stroke.[20-22] Hypertension can be mitigated through appropriate medication, and diabetes can be controlled through correct diet and nutrition, regular exercise, and stress control.[23-26] Controlling these risk factors can, in turn, reduce the risk of stroke or its fatal outcomes.[21] While the reports related to the effects of educational and behavioral interventions on controlling the risk factors of stroke are contradictory,[27,28] these interventions may be useful if they are accurately designed and cost-effective.[29] Moreover, smoking is one of the most significant risk factors for stroke and quitting smoking can significantly reduce the likelihood of this disease.[20]

Furthermore, a history of CeVDs in the family, aging, and being male are the most important non-modifiable predictors for high-risk clusters. Recently, a review study provided a comprehensive exploration of the genetic risk of stroke and the mechanisms related to it.[21] A family history of stroke is an independent risk factor for development of stroke in healthy populations.[30] Studies on large populations have shown that early stroke in the family and a history stroke in parents or siblings can significantly increase the likelihood of this event for an individual.[30,31] In addition, a family history of stroke is related to poor levels of physiological factors such as hypertension.[32]

With regard to age, previous studies clearly confirm the role of advanced age in increasing the risk of stroke and the mortality caused by it.[33,34] However, the findings related to the risk of stroke in men and women are contradictory; i.e. the findings of some studies show that stroke is more prevalent and severe among women,[35,36] while other studies have argued that men are more vulnerable to the risk of stroke.[37,38] These differences are most likely related to sex hormones and their changes during various life stages, as well as the complex interaction between these hormones and the immune system.[39] Reinforcing this argument, the findings of another study show that while the prevalence of stroke, in general, is higher among women, young men will experience stroke more than women.[40]

In the one-year follow-up, clusters containing patients with a history of CeVDs or cardio-cerebrovascular diseases and LOS equal to six days (26.1% in total) were more prone to readmission and mortality compared to the participants in the reference cluster. A history of cardio-cerebrovascular diseases is one of the most important risk factors for repeated strokes; indeed, roughly 14%–45% of patients with this sort of history will develop repeated strokes.[41,42] On the other hand, increased LOS, which is generally due to complications such as infection, pneumonia, and constipation,[12] is directly related to readmission and mortality.[13]

Limitations

The retrospective nature of the study and lack of direct interviews with the patients and their family members were the most important limitations of this study. Moreover, the current study does not explore various types and intensity of strokes. Finally, the small number of cases in some of the subgroups in terms of events such as cerebrovascular disease in family, HLP, smoking, and drug addiction may have resulted in sparse-data bias. This makes it difficult to interpret logistic regression results.[43,44] Therefore, one of the limitations of the study is the likelihood of sparse-data bias because of the small number of events in the subgroups. Mitigating these challenges in future studies can facilitate the generalization of the findings to other populations.

Conclusion

In conclusions, partitioning patients based on cardio-cerebrovascular history and LOS will provide useful information about the potential predictors and outcomes of stroke. Hypertension, diabetes, and smoking were respectively the most significant modifiable risk factors, while sex, cerebrovascular diseases in the family, and age were respectively the most significant non-modifiable risk factors in high-risk clusters and LOS equal to six days. Considering the modifiable risk factors identified in the current study, providing programs for preventing readmission and potential death for Clusters 3 and 5 seems essential.

Authors' Contribution

AS, PS, NB, NSA, and SK prepared study proposal. NB and NSA are involved in the data gathering process. SK generated study hypotheses and drafted the manuscript. SK analyzed data. Critical revisions were done by AS and PS. All authors read and approved the final manuscript.

Conflict of Interest Disclosures
There are no conflicts of interest.

Ethical Statement
Kermanshah University of Medical Sciences (ID: KUMS.REC.1395.325).

Financial Support
Kermanshah University of Medical Sciences (ID: 95282).

Acknowledgments

We extend our appreciation to the Kermanshah University of Medical Sciences for funding support for this project. Ali Soroush received a financial support from the Kermanshah University of Medical Sciences (grant number: 95282).

References

1. Thrift AG, Thayabaranathan T, Howard G, Howard VJ, Rothwell PM, Feigin VL, et al. Global stroke statistics. Int J Stroke. 2017;12(1):13-32. doi: 10.1177/1747493016676285.
2. Koton S, Rexrode KM. Trends in stroke incidence in the United States. Will women overtake men? Neurology. 2017;89(10):982-983. doi:10.1212/WNL.0000000000004342.
3. Koton S, Schneider AL, Rosamond WD, Shahar E, Sang Y, Gottesman RF, et al. Stroke incidence and mortality trends in us communities, 1987 to 2011. JAMA. 2014;312(3):259-68. doi: 10.1001/jama.2014.7692.
4. Bahonar A, Khosravi A, Khorvash F, Maracy M, Oveisgharan S, Mohammadifard N, et al. Ten-year trend in stroke incidence and its subtypes in Isfahan, Iran during 2003-2013. Iran J Neurol. 2017;16(4):201-9.
5. Azarpazhooh MR, Etemadi MM, Donnan GA, Mokhber N, Majdi MR, Ghayour-Mobarhan M, et al. Excessive incidence of stroke in Iran: evidence from the Mashhad Stroke Incidence Study (MSIS), a population-based study of stroke in the Middle East. Stroke. 2010;41(1):e3-e10. doi: 10.1161/STROKEAHA.109.559708.
6. Daneshfard B, Izadi S, Shariat A, Toudaji MA, Beyzavi Z, Niknam L. Epidemiology of stroke in Shiraz, Iran. Iran J Neurol. 2015;14(3):158-63.
7. Hosseini AA, Sobhani-Rad D, Ghandehari K, Benamer HTS. Frequency and clinical patterns of stroke in Iran - Systematic and critical review. BMC Neurol. 2010;10:72. doi: 10.1186/1471-2377-10-72.
8. Ghandehari K. Epidemiology of stroke in Iran. Galen Med J. 2016;5(S1):3-9.
9. Habibi-koolaee M, Shahmoradi L, Niakan Kalhori SR, Ghannadan H, Younesi E. Prevalence of stroke risk factors and their distribution based on stroke subtypes in Gorgan: A retrospective hospital-based study-2015-2016. Neurol Res Int. 2018;2018:2709654. doi: 10.1155/2018/2709654.
10. Lichtman JH, Leifheit-Limson EC, Jones SB, Watanabe E, Bernheim SM, Phipps MS, et al. Predictors of hospital readmission after stroke: a systematic review. Stroke. 2010;41(11):2525-33. doi: 10.1161/STROKEAHA.110.599159.
11. Nouh AM, McCormick L, Modak J, Fortunato G, Staff I. High mortality among 30-day readmission after stroke: Predictors and etiologies of readmission. Front Neurol. 2017;8:632. doi: 10.3389/fneur.2017.00632.
12. Ingeman A, Andersen G, Hundborg HH, Svendsen ML, Johnsen SP. In-hospital medical complications, length of stay, and mortality among stroke unit patients. Stroke. 2011;42(11):3214-8. doi: 10.1161/STROKEAHA.110.610881.
13. Lingsma HF, Bottle A, Middleton S, Kievit J, Steyerberg EW, Marang-van de Mheen PJ. Evaluation of hospital outcomes: the relation between length-of-stay, readmission, and mortality in a large international administrative database. BMC Health Serv Res. 2018;18(1):116. doi: 10.1186/s12913-018-2916-1.
14. Schieb LJ, Mobley LR, George M, Casper M. Tracking stroke hospitalization clusters over time and associations with county-level socioeconomic and healthcare characteristics. Stroke. 2013;44(1):146-52. doi: 10.1161/STROKEAHA.112.669705.
15. Lachkhem Y, Minvielle, E, Rican S. Geographic variations of stroke hospitalization across France: A diachronic cluster analysis. Stroke Res Treat. 2018;2018:1897569. doi: 10.1155/2018/1897569.
16. Zakiei A, Kiani N, Morovati F, Komasi S. Classification of various types of disability and determining their predictive causes in western Iran. Clin Epidemiol Global Health. 2018. In Press. doi: 10.1016/j.cegh.2018.11.003.
17. Khazaie H, Najafi F, Hamzeh B, Chehri A, Rahimi-Movaghar A, Amin-Esmaeili M, et al. Cluster analysis of psychiatric profile, its correlates, and using mental health services among the young people aged 15–34: findings from the first phase of Iranian youth cohort in Ravansar. Soc Psychiatry Psychiatr Epidemiol. 2018;53(12):1339-1348. doi: 10.1007/s00127-018-1580-4.
18. van Rheenen S, Watson TW, Alexander S, Hill MD. An analysis of spatial clustering of stroke types, in-hospital mortality, and reported risk factors in Alberta, Canada, using geographic information systems. Can J Neurol Sci. 2015;42(5):299-309. doi: 10.1017/cjn.2015.241.
19. Zhou Y, Tian Y, Zhong C, Batu B, Xu T, Li H, et al. Clustering of cardiovascular risk factors and stroke: a prospective cohort study in Inner Mongolia. Neurol Res. 2016;38(11):988-993. doi: 10.1080/01616412.2016.1243610.
20. Bos MJ, Koudstaal PJ, Hofman A, Ikram MA. Modifiable etiological factors and the burden of stroke from the Rotterdam Study: A population-based cohort study. PLoS Med. 2014;11(4):e1001634. doi: 10.1371/journal.pmed.1001634.
21. Boehme AK, Esenwa C, Elkind MSV. Stroke risk factors, genetics, and prevention. Circ Res. 2017;120(3):472-495. doi: 10.1161/CIRCRESAHA.116.308398.
22. Owolabi MO, Sarfo F, Akinyemi R, Gebregziabher M, Akpa O, Akpalu A, et al. Dominant modifiable risk factors for stroke in Ghana and Nigeria (SIREN): a case-control study. Lancet Glob Health. 2018;6(4):e436-e446. doi: 10.1016/S2214-109X(18)30002-0.
23. Akbarpour S, Khalili D, Zeraati H, Mansournia MA, Ramezankhani A, Fotouhi A. Healthy lifestyle behaviors and control of hypertension among adult hypertensive patients. Sci Rep. 2018;8(1):8508. doi: 10.1038/s41598-018-26823-5.
24. Forouhi NG, Misra A, Mohan V, Taylor R, Yancy W. Dietary and nutritional approaches for prevention and management of type 2 diabetes. BMJ. 2018;361:k2234. doi: 10.1136/bmj.k2234.
25. Zamani-Alavijeh F, Araban M, Koohestani HR, Karimy M. The effectiveness of stress management training on blood glucose control in patients with type 2 diabetes. Diabetol Metab Syndr. 2018;10:39. doi: 10.1186/s13098-018-0342-5. eCollection 2018.
26. Zhou P, Hughes AK, Grady SC, Fang L. Physical activity and chronic diseases among older people in a mid-size city in China: a longitudinal investigation of bipolar effects. BMC Public Health. 2018;18(1):486. doi: 10.1186/s12889-018-5408-7.
27. Lager KE, Mistri AK, Khunti K, Haunton VJ, Sett AK, Wilson AD. Interventions for improving modifiable risk factor control in the secondary prevention of stroke. Cochrane Database Syst Rev. 2018;5:CD009103. doi: 10.1002/14651858.CD009103.pub3.
28. Bridgwood B, Lager KE, Mistri AK, Khunti K, Wilson AD, Modi P. Interventions for improving modifiable risk factor control in the secondary prevention of stroke. Cochrane Database Syst Rev. 2018;5:CD009103. doi: 10.1002/14651858.

CD009103.pub3.
29. Salinas J, Schwamm LH. Behavioral interventions for stroke prevention: The need for a new conceptual model. Stroke. 2017;48(6):1706-1714. doi: 10.1161/STROKEAHA.117.015909.
30. Tian T, Jin G, Yu C, Lv J, Guo Y, Bian Z, et al. Family history and stroke risk in China: Evidence from a large cohort study. J Stroke. 2017;19(2):188-195. doi: 10.5853/jos.2016.01270.
31. Chung JW, Kim BJ, Han MK, Kang K, Park JM, Park SS, et al. Family history and risk of recurrent stroke. Stroke. 2016;47(8):1990-6. doi: 10.1161/STROKEAHA.116.013148.
32. Kulshreshtha A, Vaccarino V, Goyal A, McClellan W, Nahab F, Howard VJ, et al. Family history of stroke and cardiovascular health in a national cohort. J Stroke Cerebrovasc Dis. 2015;24(2):447-54. doi: 10.1016/j.jstrokecerebrovasdis.2014.09.017.
33. Kelly-Hayes M. Influence of age and health behaviors on stroke risk: lessons from longitudinal studies. J Am Geriatr Soc. 2010;58 Suppl 2:S325-8. doi: 10.1111/j.1532-5415.2010.02915.x.
34. Soriano-Tárraga C, Giralt-Steinhauer E, Mola-Caminal M, Ois A, Rodríguez-Campello A, Cuadrado-Godia E, et al. Biological age is a predictor of mortality in ischemic stroke. Sci Rep. 2018;8(1):4148. doi: 10.1038/s41598-018-22579-0.
35. Marzona I, Proietti M, Farcomeni A, Romiti GF, Romanazzi I, Raparelli V, et al. Sex differences in stroke and major adverse clinical events in patients with atrial fibrillation: A systematic review and meta-analysis of 993,600 patients. Int J Cardiol. 2018;269:182-191. doi: 10.1016/j.ijcard.2018.07.044.
36. Yu C, An Z, Zhao W, Wang W, Gao C, Liu S, et al. Sex differences in stroke subtypes, severity, risk factors, and outcomes among elderly patients with acute ischemic stroke. Front Aging Neurosci. 2015;7:174. doi: 10.3389/fnagi.2015.00174.
37. Qiao Q, Hong Y, Zhao W, Zhou G, Liu Q, Ning X, et al. Sex differences in outcomes and associated factors among stroke patients with small artery occlusion in China. Biol Sex Differ. 2018;9(1):35. doi: 10.1186/s13293-018-0194-6.
38. Meirhaeghe A, Cottel D, Cousin B, Dumont MP, Marécaux N, Amouyel P, et al. Sex differences in stroke attack, incidence, and mortality rates in northern France. J Stroke Cerebrovasc Dis. 2018;27(5):1368-1374. doi: 10.1016/j.jstrokecerebrovasdis.2017.12.023.
39. Bravo-Alegria J, McCullough LD, Liu F. Sex differences in stroke across the lifespan: The role of T lymphocytes. Neurochem Int. 2017;107:127-137.doi:10.1016/j.neuint.2017.01.009.
40. Ahangar AA, Saadat P, Heidari B, Taheri ST, Alijanpour S. Sex difference in types and distribution of risk factors in ischemic and hemorrhagic stroke. Int J Stroke. 2018;13(1):83-86. doi: 10.1177/1747493017724626.
41. Pennlert J, Eriksson M, Carlberg B, Wiklund PG. Long-term risk and predictors of recurrent stroke beyond the acute phase. Stroke. 2014;45(6):1839-41. doi: 10.1161/STROKEAHA.114.005060.
42. Arntz RM, van Alebeek ME, Synhaeve NE, van Pamelen J, Maaijwee NA, Schoonderwaldt H, et al. The very long-term risk and predictors of recurrent ischaemic events after a stroke at a young age: The FUTURE study. Eur Stroke J. 2016;1(4):337-345. doi: 10.1177/2396987316673440.
43. Greenland S, Schwartzbaum JA, Finkle WD. Problems due to small samples and sparse data in conditional logistic regression analysis. Am J Epidemiol. 2000;151(5):531-9.
44. Greenland S, Mansournia MA, Altman DG. Sparse data bias: a problem hiding in plain sight. BMJ. 2016;352:i1981. doi: 10.1136/bmj.i1981.