Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Multifactorial Analysis of Aphasia Severity and Language Recovery Following Acute Ischemic Stroke: A Cross-Sectional and Longitudinal Observational Study

Guoqiang Cao
First Affiliated Hospital of Soochow University
Jing Li
First Affiliated Hospital of Soochow University
Xinyu Li
First Affiliated Hospital of Soochow University
Yanyan Zhou
First Affiliated Hospital of Soochow University
Xiaoli Shan
First Affiliated Hospital of Soochow University
Hui Dai
First Affiliated Hospital of Soochow University
Mengyi Guo
First Affiliated Hospital of Soochow University
Quanquan Zhang
zhangquanquan@suda.edu.cn
First Affiliated Hospital of Soochow University

Research Article

Keywords: Post-stroke aphasia, Language recovery, Reperfusion therapy, NIHSS score, Risk factors

Posted Date: November 1st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11182132/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Post-stroke aphasia (PSA) is a serious complication of cerebral infarction and may substantially affect patients' communication and quality of life. Early identification of factors associated with aphasia severity and language recovery may facilitate timely clinical management. This study aimed to investigate the severity and specific language dimensions of acute-phase PSA and to identify factors associated with spontaneous language recovery.

Methods

This observational study employed a cross-sectional analysis combined with longitudinal cohort follow-up to examine 102 patients with acute-phase PSA treated at the First Affiliated Hospital of Soochow University from October 2020 to March 2026. Aphasia was assessed using the Western Aphasia Battery (WAB). Baseline clinical data were collected, including sex, age, risk factors, National Institutes of Health Stroke Scale (NIHSS) scores, infarct location, and reperfusion therapy (thrombolysis and/or thrombectomy). The study analyzed factors influencing aphasia severity, various dimensions of language function, and spontaneous language recovery.

Results

In the cross-sectional study, we found that reperfusion therapy, including thrombolysis and/or thrombectomy, was associated with lower odds of severe aphasia (P = 0.028), while higher NIHSS scores (P = 0.027) and infarct location involving frontal lobe (P = 0.015) were associated with severe aphasia. Further analysis of individual language domains showed that higher NIHSS scores were associated with lower scores in information content (P = 0.001), fluency (P < 0.001), auditory comprehension (P < 0.001), repetition (P = 0.001), and naming (P = 0.011); reperfusion therapy was associated with better naming performance (P = 0.018); frontal lobe involvement was associated with lower information content scores (P = 0.043) and naming scores (P = 0.029); and temporal lobe involvement was associated with lower auditory comprehension scores (P = 0.046). In the longitudinal cohort follow-up study, we found that higher diastolic blood pressure (DBP) at onset (P = 0.029) and a history of diabetes (P = 0.034) were associated with lower recovery ratio.

Conclusions

The severity and recovery of acute-phase PSA may be influenced by multiple factors, including overall stroke severity, reperfusion therapy, infarct location, and underlying comorbidities.

Background

As one of the serious complications of cerebral infarction, PSA affects approximately one-third of stroke patients [1, 2]. PSA is more common among patients with infarcts in the territory of the left middle cerebral artery (MCA), which is related to the fact that language centers are located in the left cerebral cortex for most right-handed individuals and some left-handed individuals [3, 4]. Due to impaired language comprehension and expression, patients with PSA may experience difficulties in social interaction, emotional well-being, and quality of life, as well as communication difficulties that may negatively affect family relationships and hinder return to work [5, 6]. Although most patients recover spontaneously within six months after stroke, 30% to 40% continue to suffer from persistent language impairments [7]. Therefore, early identification of language deficits and associated risk factors is crucial for timely and targeted interventions.

Previous studies have identified several factors associated with PSA. However, important gaps remain. First, numerous studies have shown that multiple factors, including imaging characteristics such as infarct location and volume [8–13], NIHSS scores at onset [4, 14, 15], blood glucose levels [16, 17], educational background [14, 18, 19], and white matter fiber integrity [8, 20, 21], may be associated with PSA type, severity, and recovery. However, the relationship between certain demographic factors and PSA remains controversial. Kernbach et al. [22] and Ali et al. [23] suggested that younger age (< 55 years) and female sex may influence PSA severity and prognosis, whereas Zhou et al. [24] and Wang et al. [25] found no significant association between sex or age and PSA. Second, most prior research has adopted cross-sectional designs, resulting in relatively limited longitudinal studies on language recovery. Moreover, study populations have predominantly focused on subacute or chronic phases, leaving insufficient understanding of PSA injury characteristics and contributing factors during the acute phase. Additionally, although the aphasia quotient (AQ), calculated using the WAB in common clinical research, provides a clear reflection of PSA severity, it is derived from a composite of multiple language dimensions, which may obscure the effects of specific factors on individual language functions.

Against this background, our study focused on patients with acute-phase PSA and employed an observational study design combining cross-sectional analysis with longitudinal cohort follow-up, analyzing a total of 102 patients in the acute phase of PSA. First, through cross-sectional analysis, we examined the associations between demographic characteristics, laboratory tests, imaging findings, and other factors with the severity of acute-phase PSA and specific language function domains, aiming to identify key influencing factors of acute-phase PSA. Second, via cohort follow-up, we assessed how baseline factors affect PSA recovery and screen for predictive indicators of PSA outcome. Through this research design, the study aims to explore the factors influencing both acute-phase damage and recovery in PSA, providing insights for early identification of high-risk patients with severe PSA and timely intervention. Given that some patients face limitations due to economic constraints, time availability, or limited access to rehabilitation resources, making systematic language therapy unfeasible, this study primarily observes the spontaneous language recovery process of language function in PSA patients.

Methods

1.1 Study Population

This observational study included a cross-sectional analysis and a longitudinal cohort follow-up of 102 patients with acute-phase PSA admitted for treatment at the First Affiliated Hospital of Soochow University between October 2020 and March 2026. Acute-phase PSA was defined as aphasia assessed within 7 days after stroke onset.

Inclusion criteria: ▯ Diagnosed with ischemic stroke according to the diagnostic criteria of Chinese guidelines for diagnosis and treatment of acute ischemic stroke 2018[26], confirmed by brain magnetic resonance imaging (MRI), with the infarct lesion showing a high signal on diffusion-weighted imaging (DWI) sequence and a low signal on apparent diffusion coefficient (ADC) sequence; ▯ A WAB AQ score < 93.8, diagnosed with aphasia [27, 28]; ▯ Ischemic stroke with infarct primarily located in the left MCA territory, allowing for small infarcts in other regions, and no prior history of hemorrhagic or ischemic stroke; ▯ Time from stroke onset to completion of the first WAB assessment ≤ 7 days; ▯ Conscious and cooperative during examination; ▯ Right-handed as determined by the Edinburgh Handedness Inventory (EHI)[29]; ▯ Age ≥ 18 years; ▯ Native Chinese speaker with a monolingual language background; ▯ Normal language function and intelligence before stroke onset, and no history of dysarthria or psychiatric disorders; ▯ At least primary school education.

Exclusion criteria: ▯ AQ score ≥ 93.8; ▯ Presence of other conditions that may cause language dysfunction, such as intracranial tumors or infections; ▯ Prior history of dysarthria; ▯ Concurrent symptoms that prevent cooperation with assessment, including psychiatric disorders, visual impairments, non-aphasic auditory comprehension deficits, and dementia; ▯ Expected survival less than one year; ▯ Patients who received language rehabilitation during follow-up were excluded from the longitudinal cohort analysis. Rehabilitation interventions included speech and language therapy (SLT), melodic intonation therapy (MIT), and non-invasive brain stimulation.

It should be noted that due to lockdowns implemented during the Coronavirus Disease 2019 (COVID-19) pandemic and other objective factors, the patient enrollment period in this study was relatively long. However, key assessment tools such as updated stroke management guidelines and the WAB assessment did not undergo major revisions during the study period, so the extended time span did not affect the conclusions of this study.

1.2 Language Function Testing Methods

The WAB assessment was conducted jointly by two senior neurologists who had undergone standardized training, covering five domains: information content, fluency, auditory comprehension, repetition, and naming. Based on the WAB, the type of aphasia was classified and the AQ score was calculated. The AQ score is obtained by summing the scores across these five dimensions (each scored out of 10), multiplying the total by 2, with a maximum possible score of 100.

1.3 Data Collection

All patients underwent the initial WAB assessment within 48 hours of hospital admission (T1). For emergency patients, assessment was conducted within 48 hours of hospital admission after completion of initial diagnostic and therapeutic procedures—including blood tests, non-contrast head computed tomography (CT), CT angiography (CTA)/CT perfusion (CTP), and reperfusion therapy—and transfer to the ward. Clinical data collected at T1 included aphasia assessment, sex, age, vascular risk factors, NIHSS score, trial of ORG 10172 in acute stroke treatment (TOAST) classification, infarct location, and reperfusion treatment. Vascular risk factors included hypertension, diabetes, atrial fibrillation, dyslipidemia, hyperhomocysteinemia, and history of smoking or alcohol use.

Face-to-face follow-up assessments were conducted at 2 to 6 months after onset [median, 3.8 months; interquartile range (IQR), 3.18 to 5.78] (T2), during which two neurologists, both uniformly trained, jointly performed the WAB assessment and calculated the AQ score.

1.4 Study Design and Statistical Methods

At T1, patients' clinical information was collected to analyze the factors influencing the severity of acute-phase PSA and various dimensions of language function. At T2, a secondary assessment was conducted, combining baseline data to analyze the factors affecting PSA recovery. The study flow is shown in Fig. 1. Statistical analysis was performed using SPSS 27.0 software, with a significance level set at $\alpha = 0.05$.

1.4.1 Cross-sectional Analysis

Based on the AQ scores at T1 and using an AQ cutoff of 60, as reported by Cherney et al. [30] and Wang et al. [31], the 102 PSA patients were divided into a mild aphasia group (AQ ≥ 60, n = 33) and a severe aphasia group (AQ < 60, n = 69).Aphasia severity was used as the dependent variable to explore its influencing factors. For continuous independent variables distributed normally, Levene's test was used to assess the homogeneity of variances, and the assumption of equal variances was satisfied. Independent samples t-tests were conducted for univariate analysis. For non-normally distributed variables, the Mann-Whitney U test was applied. For categorical variables, chi-square tests were used to compare differences in distributions. Independent variables with $P < 0.05$ in univariate analysis were included in binary logistic regression analysis.

Further, using the scores on each dimension of language function as continuous dependent variables, independent variables with $P < 0.05$ in univariate analysis were included in linear regression to examine their influencing factors.

1.4.2 Cohort Follow-up Analysis

The 102 patients were scheduled for follow-up at 2–6 months after stroke onset. 65 were lost to follow-up for various reasons (see Fig. 1), and 37 completed the follow-up assessment. Baseline characteristics comparing those who were followed up versus those who were not are presented in the supplementary table. Due to a follow-up completion rate of only 36.3% and a significant difference in fasting blood glucose between the two groups (P = 0.023), the possibility of selection bias due to non-random loss to follow-up cannot be completely excluded.

According to Lazar et al. [32], the recovery ratio was calculated as (T2AQ – T1AQ)/(100 – T1AQ). By correcting for the maximum recoverable space at baseline, i.e., "100 – T1AQ," the recovery ratio to some extent reduces the influence of baseline severity on recovery. A higher ratio indicates more favorable recovery. The recovery ratio was used as the dependent variable in longitudinal analysis. Given the limited sample size of the follow-up cohort, the normality of continuous variables was not assumed, and Spearman rank correlation was used for univariate analysis. Binary variables were analyzed with the Mann-Whitney U test. Categorical variables with more than two groups were analyzed using the Kruskal-Wallis H test. Independent variables with P < 0.05 in univariate analysis were included in linear regression models for further analysis.

Results

2.1 Cross-sectional Analysis of Factors Associated with Aphasia Severity

2.1.1 Univariate Analysis of Aphasia Severity

The basic characteristics of all patients and the results of univariate analysis are shown in Table 1. Univariate analysis indicated that patients in the severe aphasia group had significantly higher NIHSS scores at admission compared to those in the mild aphasia group (10.0 vs 4.0, P < 0.001); the proportion of patients receiving thrombolysis and/or thrombectomy was significantly lower in the severe aphasia group than in the mild aphasia group (33.3% vs 54.5%, P = 0.041); and the severe aphasia group had significantly higher proportions of infarction involving the frontal (81.2% vs 45.5%, P < 0.001), parietal (73.9% vs 45.5%, P = 0.005), and temporal (73.9% vs 45.5%, P = 0.005) lobes, as well as multiple infarcts (79.7% vs 60.6%, P = 0.041).

Table 1
General Clinical Characteristics of Patients with PSA

Variable	Mild aphasia (n = 33)	Severe aphasia (n = 69)	P value
Demographic characteristics and vital signs			
Sex			0.254
Male, n (%)	23 (69.7)	40 (58.0)	
Female, n (%)	10 (30.3)	29 (42.0)	
Age, years	61.333 ± 16.188	62.000 ± 13.489	0.827
Body mass index (BMI), kg/m^2	24.490 ± 3.530	23.340 ± 2.800	0.078
SBP, mmHg	145.390 ± 20.230	149.720 ± 23.010	0.358
DBP, mmHg	88.000 (75.000 to 93.000)	89.000 (77.500 to 97.500)	0.219
Medical history and lifestyle factors			
Hypertension, n (%)	18 (54.5)	43 (62.3)	0.454
Diabetes, n (%)	7 (21.2)	20 (29.0)	0.405
Atrial fibrillation, n (%)	6 (18.2)	7 (10.1)	0.255
Smoking history, n (%)	9 (27.3)	20 (29.0)	0.858
Alcohol consumption history, n (%)	8 (24.2)	16 (23.2)	0.907
Laboratory tests			
Liver function			
Alanine aminotransferase, U/L	15.200 (10.750 to 19.450)	13.200 (10.200 to 17.950)	0.632
Aspartate aminotransferase, U/L	20.800 (15.150 to 24.650)	20.500 (14.900 to 26.700)	0.617
Renal function			
Urea, mmol/L	5.100 (3.900 to 7.150)	6.000 (4.680 to 7.900)	0.115
Creatinine, μmol/L	67.000 (52.350 to 78.300)	67.300 (59.250 to 83.550)	0.515
Uric acid, μmol/L	333.900 (243.900 to 392.750)	323.800 (245.600 to 378.500)	0.889
Lipid levels			
Total cholesterol, mmol/L	4.470 (3.900 to 5.220)	4.390 (3.560 to 5.300)	0.930
Triglycerides, mmol/L	1.310 (0.970 to 1.700)	1.260 (0.940 to 1.590)	0.652
High-density lipoprotein cholesterol (HDL-C), mmol/L	0.970 (0.860 to 1.130)	0.980 (0.800 to 1.190)	0.838
Low-density lipoprotein cholesterol (LDL-C), mmol/L	2.790 (2.230 to 3.240)	2.640 (1.910 to 3.450)	0.729
Glycemic parameters			
Fasting blood glucose, mmol/L	5.340 (4.720 to 6.400)	5.830 (5.150 to 7.660)	0.063
Glycated hemoglobin (HbA1c), %	5.800 (5.400 to 6.200)	5.900 (5.500 to 6.950)	0.201
Complete blood count			
White blood cells, ×10^9/L	7.990 (6.060 to 9.110)	7.770 (6.330 to 9.900)	0.775
Platelets, ×10^9/L	196.000 (169.500 to 237.000)	213.000 (177.500 to 250.500)	0.399
Blood coagulation			
D-dimer, mg/L	0.590 (0.230 to 1.040)	0.630 (0.320 to 1.820)	0.290
International normalized ratio (INR)	1.010 (0.980 to 1.050)	1.010 (0.980 to 1.090)	0.407
Homocysteine, μmol/L	10.00 (8.750 to 14.470)	10.600 (8.580 to 15.190)	0.791
Clinical features of stroke			
TOAST classification			0.217
Large-artery atherosclerosis, n (%)	18 (54.5)	51 (73.9)	
Cardioembolism, n (%)	7 (21.2)	10 (14.5)	

Variable	Mild aphasia (n = 33)	Severe aphasia (n = 69)	P value
Small-vessel occlusion, n (%)	3 (9.1)	1 (1.4)	
Stroke of other determined etiology, n (%)	4 (12.1)	5 (7.2)	
Stroke of undetermined etiology, n (%)	1 (3.0)	2 (2.9)	
Thrombolysis and/or thrombectomy			0.041
Yes, n (%)	18 (54.5)	23 (33.3)	
No, n (%)	15 (45.5)	46 (66.7)	
Length of hospital stay, days	10.000 (9.000 to 13.500)	12.000 (9.000 to 14.000)	0.232
Clinical scores			
Admission NIHSS score	4.000 (2.500 to 10.500)	10.000 (6.000 to 15.500)	< 0.001
Discharge NIHSS score	3.000 (2.000 to 6.500)	9.000 (4.500 to 14.000)	< 0.001
Aphasia assessment			
Aphasia types			
Global aphasia, n (%)	0 (0.0)	43 (62.3)	
Anomic aphasia, n (%)	26 (78.8)	0 (0.0)	
Broca's aphasia, n (%)	2 (6.1)	19 (27.5)	
Conduction aphasia, n (%)	2 (6.1)	2 (2.9)	
Transcortical motor aphasia, n (%)	3 (9.1)	1 (1.4)	
Transcortical mixed aphasia, n (%)	0 (0.0)	4 (5.8)	
Aphasia assessment			
AQ	81.100 (74.600 to 88.800)	6.800 (0.700 to 19.650)	
WAB language subtest scores			
Information content	8.000 (8.000 to 9.000)	0.000 (0.000 to 1.000)	
Fluency	7.000 (5.000 to 8.500)	0.000 (0.000 to 1.000)	
Auditory comprehension	9.450 (9.100 to 9.830)	2.200 (0.200 to 6.100)	
Repetition	8.600 (8.000 to 9.450)	0.000 (0.000 to 0.400)	
Naming	7.800 (6.700 to 8.650)	0.000 (0.000 to 0.000)	
Imaging characteristics			
Number of infarct lesions			0.041
Multiple infarcts, n (%)	20 (60.6)	55 (79.7)	
Single infarct, n (%)	13 (39.4)	14 (20.3)	
Infarction locations			
Frontal lobe infarction, n (%)	15 (45.5)	56 (81.2)	< 0.001
Parietal lobe infarction, n (%)	15 (45.5)	51 (73.9)	0.005
Temporal lobe infarction, n (%)	15 (45.5)	51 (73.9)	0.005
Insular infarction, n (%)	10 (30.3)	22 (31.9)	0.872
Occipital lobe infarction, n (%)	5 (15.2)	16 (23.2)	0.348
Thalamic infarction, n (%)	3 (9.1)	0 (0.0)	0.011
Basal ganglia and periventricular infarction, n (%)	21 (63.6)	34 (49.3)	0.173
Centrum semiovale infarction, n (%)	5 (15.2)	11 (15.9)	0.918
Infarction in other regions, n (%)	1 (3.0)	7 (10.1)	0.211

Variable	Mild aphasia (n = 33)	Severe aphasia (n = 69)	P value
Note:			
1. The normality of continuous variables was assessed using the Shapiro-Wilk test. Continuous variables are presented as mean ± standard deviation for normally distributed data and as median (IQR) for non-normally distributed data. Categorical variables are presented as n (%). When the distribution of a continuous variable differed between the two groups, median (IQR) was used for both groups.			
2. For patients receiving thrombolytic and/or thrombectomy therapy in the emergency setting, NIHSS scores were assessed jointly by two experienced neurologists immediately after admission to the emergency department and before thrombolytic and/or thrombectomy therapy.			
3. A history of hypertension or diabetes was defined as a previous diagnosis made by a physician, with corresponding pharmacological treatment.			
4. Fasting blood glucose and other laboratory parameters were measured using venous blood samples obtained from the antecubital vein after an 8-hour fast (from 22:00 on the day of admission to 06:00 the following day).			
5. Blood pressure was measured in a quiet ward environment after patients had refrained from smoking and caffeine, emptied their bladders, and rested in a seated position for 30 minutes. A single blood pressure measurement was obtained rather than a 24-hour ambulatory blood pressure measurement.			
6. Smoking history was defined as smoking ≥ 10 cigarettes/day for at least 6 months, with alcohol consumption history defined as alcohol intake ≥ 30 g/day for at least 6 months[13].			
7. Infarct locations were determined by two senior neurologists based on the imaging findings and the official radiology reports. In cases of disagreement, consensus was reached through discussion with radiologists at our center.			
8. Other infarct areas include periventricular, corpus callosum, hypothalamus, cerebral peduncle, corona radiata, cerebellum, and other regions.			

Although the involvement of the thalamus in infarction showed a significant effect on the severity of aphasia in the chi-square test (P = 0.011), all three patients with thalamic involvement were in the mild aphasia group, making it impossible to obtain stable logistic regression parameters; thus, it was excluded from the regression model to avoid bias.

2.1.2 Binary Logistic Regression Analysis of Aphasia Severity

Using variables selected by univariate analysis (see Section 2.1.1) as covariates and severity as the dependent variable, a logistic regression analysis was performed using the enter method, with results shown in Table 2. The regression model was statistically significant (Omnibus test P < 0.001). The Hosmer-Lemeshow test did not indicate poor fit (P = 0.468). The prediction accuracy rate was 75.5%. The results showed that receiving reperfusion therapy with thrombolysis and/or thrombectomy was associated with lower odds of severe aphasia after stroke [odds ratio (OR) = 0.310; 95% confidence interval (CI): 0.109 to 0.879; P = 0.028]. In contrast, a higher NIHSS score at admission (OR = 1.106; 95% CI: 1.011 to 1.209; P = 0.027) and infarct involvement of the frontal lobe (OR = 4.016; 95% CI: 1.314 to 12.271; P = 0.015) were associated with the occurrence of severe aphasia after stroke.

Table 2
Binary Logistic Regression Analysis

Variable	B	SE	Wald χ2	P value	OR	95% CI
Constant	-1.035	2.350	0.194	0.660	0.355	
Admission NIHSS score	0.101	0.046	4.867	0.027	1.106	1.011 to 1.209
Thrombolysis and/or thrombectomy	-1.172	0.532	4.853	0.028	0.310	0.109 to 0.879
Multiple infarcts	-0.067	0.674	0.010	0.921	0.935	0.250 to 3.504
Frontal lobe infarction	1.390	0.570	5.950	0.015	4.016	1.314 to 12.271
Parietal lobe infarction	0.473	0.585	0.654	0.419	1.604	0.510 to 5.044
Temporal lobe infarction	0.683	0.582	1.380	0.240	1.980	0.633 to 6.191

2.2 Multivariable Linear Regression Analysis of Language Function

Based on the univariate analysis of factors influencing the severity of aphasia, we further examined the effects of these factors on different dimensions of language function. We constructed five separate linear regression models, using the scores in five dimensions—information content, fluency, auditory comprehension, repetition, and naming—as dependent variables, and the variables with P < 0.05 in univariate analysis (see Section 2.1.1) as covariates. The parameters of these models are shown in Table 3 − 1, indicating that each model is statistically significant. The analysis results are presented in Table 3 − 2, indicating that higher NIHSS scores at admission were associated with lower scores across all language function dimensions (all B values < -0.1, all P values < 0.05). Patients who received reperfusion treatment with thrombolysis and/or thrombectomy had higher naming scores (B = 1.556; 95% CI: 0.270 to 2.843; P = 0.018). Patients with infarcts involving the frontal lobe had lower scores in the information content dimension (B = -1.604; 95% CI: -3.155 to -0.053; P = 0.043) and the naming dimension (B = -1.682; 95% CI: -3.188 to -0.177; P = 0.029). And those with infarcts affecting the temporal lobe had lower auditory comprehension scores (B = -1.569; 95% CI: -3.112 to -0.025; P = 0.046).

Table 3
− 1. Summary of Linear Regression Models for WAB Language Function Scores

Information content	R2	Adjusted R2	F	P value
	0.306	**0.246**	**5.125**	**< 0.001**
Fluency	0.246	0.181	3.795	< 0.001
Auditory comprehension	0.374	0.320	6.937	< 0.001
Repetition	0.237	0.172	3.618	0.001
Naming	0.288	0.277	4.713	< 0.001

Table 3
− 2. Coefficients of Linear Regression Models for WAB Language Function Scores

	Information content			Fluency			Auditory comprehension			Repetition			Naming		
	B	**95% CI**	**P value**	**B**	**95% CI**	**P value**	**B**	**95% CI**	**P value**	**B**	**95% CI**	**P value**	**B**	**95% CI**	**P valu**
Constant	7.513	1.576 to 13.451	0.014	4.024	-1.427 to 9.475	0.146	8.019	1.996 to 14.041	0.010	7.024	0.125 to 13.924	0.046	7.332	1.568 to 13.095	0.01
Admission NIHSS score	-0.188	-0.298 to -0.078	0.001	-0.190	-0.291 to -0.089	< 0.001	-0.227	-0.338 to -0.115	< 0.001	-0.220	-0.348 to -0.093	0.001	-0.139	-0.246 to -0.033	0.01
Reperfusion	1.121	-0.205 to 2.446	0.097	0.768	-0.448 to 1.985	0.213	1.339	-0.005 to 2.684	0.051	0.348	-1.193 to 1.888	0.655	1.556	0.270 to 2.843	0.01
Multiple infarcts	-0.577	-2.339 to 1.185	0.517	-0.338	-1.956 to 1.280	0.679	-0.845	-2.632 to 0.942	0.350	-0.646	-2.694 to 1.401	0.532	-0.751	-2.462 to 0.959	0.38
Frontal lobe infarction	-1.604	-3.155 to -0.053	0.043	-1.124	-2.548 to 0.300	0.120	-1.544	-3.117 to 0.029	0.054	-1.008	-2.810 to 0.794	0.269	-1.682	-3.188 to -0.177	0.02
Parietal lobe infarction	-0.422	-2.017 to 1.173	0.601	-0.009	-1.473 to 1.456	0.991	-0.313	-1.931 to 1.304	0.701	-0.589	-2.442 to 1.265	0.530	-0.155	-1.703 to 1.394	0.84
Temporal lobe infarction	-1.313	-2.835 to 0.209	0.090	-0.587	-1.984 to 0.811	0.407	-1.569	-3.112 to -0.025	0.046	-1.069	-2.837 to 0.700	0.233	-1.321	-2.798 to 0.157	0.07

2.3 Characteristics and Results of Cohort Follow-up Analysis Affecting Recovery Ratio

2.3.1 Univariate Analysis of Recovery Ratio

The aphasia characteristics of 37 patients who completed follow-up at T1 and T2 are shown in Table 4. Univariate analysis was performed with recovery ratio as the dependent variable. The baseline characteristics of the followed-up patients and the results of univariate analysis are shown in Table 5. The results indicate that higher baseline DBP (P = 0.013), admission NIHSS score (P = 0.015), urea (P = 0.019) and creatinine (P = 0.038) levels, as well as a history of diabetes (P = 0.015) and parietal lobe involvement (P = 0.031), were associated with lower recovery ratio; whereas higher baseline scores in the comprehension dimension were associated with higher recovery ratio (P = 0.041).

Table 4
Changes in Aphasia Measures from Baseline to Follow-up

Aphasia types	Baseline (T1)	Follow-up (T2)	Change (Δ)
No aphasia, n (%)	0 (0.0)	5 (13.5)	+ 5 (13.5)
Global aphasia, n (%)	14 (37.8)	6 (16.2)	-8 (21.6)
Anomic aphasia, n (%)	8 (21.6)	16 (43.2)	+ 8 (21.6)
Broca's aphasia, n (%)	7 (18.9)	4 (10.8)	-3 (8.1)
Conduction aphasia, n (%)	2 (5.4)	2 (5.4)	0 (0.0)
Transcortical motor aphasia, n (%)	2 (5.4)	2 (5.4)	0 (0.0)
Transcortical sensory aphasia, n (%)	0 (0.0)	2 (5.4)	+ 2 (5.4)
Transcortical mixed aphasia, n (%)	4 (10.8)	0 (0.0)	-4 (10.8)
Aphasia assessment			
AQ	19.700 (2.700 to 71.750)	76.200 (43.200 to 92.650)	+ 22.700 (11.700 to 49.000)
Aphasia severity classification			
No aphasia, n (%)	0 (0.0)	5 (13.5)	+ 5 (13.5)
Mild aphasia, n (%)	10 (27.0)	18 (48.6)	+ 8 (21.6)
Severe aphasia, n (%)	27 (73.0)	14 (37.8)	-13 (35.1)
WAB language subtest scores			
Information content	1.000 (0.000 to 6.500)	8.000 (4.500 to 9.000)	+ 2.000 (1.000 to 6.000)
Fluency	1.000 (0.000 to 6.000)	7.000 (4.00 to 9.000)	+ 2.000 (1.000 to 6.000)
Auditory comprehension	4.100 (0.750 to 9.325)	9.300 (5.425 to 9.900)	+ 0.950 (0.150 to 4.500)
Repetition	0.200 (0.000 to 8.300)	8.400 (3.800 to 9.400)	+ 1.400 (0.500 to 5.800)
Naming	0.000 (0.000 to 5.150)	7.300 (1.750 to 9.100)	+ 2.100 (0.950 to 6.300)

Note: The change (Δ) represents the median of the individual differences between T2 and T1 values and was not calculated by subtracting the median at T1 from the median at T2.

Table 5
Univariate Analysis of the Follow-up Group

Variable	Follow-up group (n = 37)	Test statistic	P value
Demographic characteristics and vital signs			
Sex		-0.188	0.851
Male, n (%)	27 (73.0)		
Female, n (%)	10 (27.0)		
Age, years	65.000 (49.500 to 77.000)	-0.274	0.101
BMI, kg/m^2	23.294 ± 2.555	0.255	0.127
SBP, mmHg	147.541 ± 21.518	-0.242	0.149
DBP, mmHg	85.622 ± 14.978	-0.404	0.013
Medical history and lifestyle factors			
Hypertension, n (%)	22 (59.5)	-1.748	0.080
Diabetes, n (%)	7 (18.9)	-2.443	0.015
Atrial fibrillation, n (%)	6 (16.2)	-1.896	0.058
Smoking history, n (%)	13 (35.1)	-1.050	0.294
Alcohol consumption history, n (%)	8 (21.6)	-0.406	0.685
Laboratory tests			
Liver function			
Alanine aminotransferase, U/L	14.800 (9.100 to 19.450)	0.107	0.527
Aspartate aminotransferase, U/L	20.800 (14.900 to 27.450)	-0.164	0.333
Renal function			
Urea, mmol/L	5.100 (4.550 to 7.250)	-0.385	0.019
Creatinine, μmol/L	72.449 ± 16.690	-0.343	0.038
Uric acid, μmol/L	335.257 ± 92.278	0.109	0.522
Lipid levels			
Total cholesterol, mmol/L	4.661 ± 1.153	0.091	0.592
Triglycerides, mmol/L	1.130 (0.850 to 1.505)	0.017	0.919
HDL-C, mmol/L	1.040 ± 0.227	-0.040	0.814
LDL-C, mmol/L	2.963 ± 1.137	0.031	0.854
Glycemic parameters			
Fasting blood glucose, mmol/L	5.370 (4.695 to 6.070)	-0.155	0.359
HbA1c, %	5.900 (5.450 to 6.500)	-0.103	0.542
Complete blood count			
White blood cells, $\times 10^9/L$	7.352 ± 2.091	-0.050	0.768
Platelets, $\times 10^9/L$	208.000 (175.000 to 273.000)	0.294	0.078
Blood coagulation			
D-dimer, mg/L	0.640 (0.270 to 1.710)	-0.141	0.406
INR	1.010 (0.980 to 1.090)	-0.313	0.059
Homocysteine, μmol/L	10.600 (8.950 to 16.200)	0.104	0.541
Clinical features of stroke			
TOAST classification		7.382	0.117

Note: For continuous variables, report the Spearman correlation coefficient (ρ) in the "Test Statistic" column; for dichotomous variables, report the Mann-Whitney U test statistic (Z); and for categorical variables with more than two groups, report the Kruskal-Wallis H statistic.

Variable	Follow-up group (n = 37)	Test statistic	P value
Large-artery atherosclerosis, n (%)	21 (56.8)		
Cardioembolism, n (%)	8 (21.6)		
Small-vessel occlusion, n (%)	1 (2.7)		
Stroke of other determined etiology, n (%)	5 (13.5)		
Stroke of undetermined etiology, n (%)	2 (5.4)		
Thrombolysis and/or thrombectomy		-0.736	0.462
Yes, n (%)	14 (37.8)		
No, n (%)	23 (62.2)		
Length of hospital stay, days	11.000 (9.000 to 14.000)	-0.233	0.165
Clinical scores			
Admission NIHSS score	8.000 (5.000 to 16.000)	-0.398	0.015
Discharge NIHSS score	5.000 (2.500 to 13.000)	-0.492	0.002
Follow-up interval, months	3.800 (3.184 to 5.784)	0.049	0.775
Aphasia assessment			
Aphasia types		7.459	0.189
Global aphasia, n (%)	14 (37.8)		
Anomic aphasia, n (%)	8 (21.6)		
Broca’s aphasia, n (%)	7 (18.9)		
Conduction aphasia, n (%)	2 (5.4)		
Transcortical motor aphasia, n (%)	2 (5.4)		
Transcortical mixed aphasia, n (%)	4 (10.8)		
Aphasia assessment			
AQ	19.700 (2.700 to 71.750)	0.177	0.294
Aphasia severity classification		-0.479	0.632
Mild, n (%)	10 (27.0)		
Severe, n (%)	27 (73.0)		
WAB language subtest scores			
Information content	1.000 (0.000 to 6.500)	0.106	0.534
Fluency	1.000 (0.000 to 6.000)	0.052	0.760
Auditory comprehension	4.100 (0.750 to 9.325)	0.337	0.041
Repetition	0.200 (0.000 to 8.300)	0.059	0.727
Naming	0.000 (0.000 to 5.150)	0.096	0.573
Recovery ratio, %	51.100 (27.800 to 79.750)		
Imaging characteristics			
Number of infarct lesions		-1.841	0.066
Multiple infarcts, n (%)	28 (75.7)		
Single infarct, n (%)	9 (24.3)		
Infarction locations			
Frontal lobe infarction, n (%)	26 (70.3)	-0.299	0.765
Parietal lobe infarction, n (%)	24 (64.9)	-2.163	0.031
Temporal lobe infarction, n (%)	25 (67.6)	-1.055	0.292

Note: For continuous variables, report the Spearman correlation coefficient (ρ) in the "Test Statistic" column; for dichotomous variables, report the Mann-Whitney U test statistic (Z); and for categorical variables with more than two groups, report the Kruskal-Wallis H statistic.

Variable	Follow-up group (n = 37)	Test statistic	P value
Insular infarction, n (%)	12 (32.4)	-1.736	0.083
Occipital lobe infarction, n (%)	6 (16.2)	-1.092	0.275
Thalamic infarction, n (%)	1 (2.7)	-0.562	0.574
Basal ganglia and periventricular infarction, n (%)	18 (48.6)	-0.486	0.627
Centrum semiovale infarction, n (%)	6 (16.2)	-1.277	0.201
Infarction in other regions, n (%)	3 (8.1)	-0.723	0.469
Note: For continuous variables, report the Spearman correlation coefficient (ρ) in the "Test Statistic" column; for dichotomous variables, report the Mann-Whitney U test statistic (Z); and for categorical variables with more than two groups, report the Kruskal-Wallis H statistic.			

2.3.2 Multivariable Linear Regression Analysis of Recovery Ratio

A linear regression analysis was performed with recovery ratio as the dependent variable and variables from univariate analysis with $P < 0.05$ (see Section 2.3.1) as independent variables. The model was statistically significant (adjusted $R^2 = 0.396$, $F = 4.377$, $P = 0.002$). The results (Table 6) indicate that higher baseline DBP (B = -0.006; 95% CI: -0.011 to -0.001; P = 0.029) and a history of diabetes (B = -0.206; 95% CI: -0.394 to -0.017; P = 0.034) are associated with lower recovery ratios.

Table 6
Linear Regression Analysis of Language Recovery Ratio

Constant	B	SE	β	t	P value	95% CI	VIF
	1.278	**0.263**		**4.852**	**< 0.001**	**0.740 to 1.817**	
Auditory comprehension	0.012	0.011	0.179	1.140	0.264	-0.010 to 0.035	1.473
DBP	-0.006	0.002	-0.311	-2.305	0.029	-0.011 to -0.001	1.089
Admission NIHSS score	-0.006	0.008	-0.133	-0.783	0.440	-0.022 to 0.010	1.724
Urea	-0.031	0.025	-0.222	-1.250	0.221	-0.081 to 0.020	1.880
Creatinine	0.000	0.003	-0.012	-0.072	0.943	-0.006 to 0.005	1.565
Diabetes	-0.206	0.092	-0.297	-2.227	0.034	-0.394 to -0.017	1.061
Parietal lobe infarction	-0.066	0.085	-0.117	-0.783	0.440	-0.240 to 0.107	1.327
Supplementary Table. Comparison of Baseline Characteristics Between Follow-up and Non-follow-up Groups							

Discussion

As one of the common disabling impairments following stroke, PSA can significantly affect patients' emotional well-being, quality of life, social interactions, and return to work. Therefore, early identification of patients at high risk of severe PSA and timely intervention are of great significance for aphasia recovery and improvement in patients' quality of life. This study analyzed clinical data from 102 patients with acute-phase PSA, examining influencing factors associated with overall severity of aphasia, specific language dimensions, and spontaneous language recovery. The findings revealed that higher NIHSS scores and frontal lobe involvement were associated with severe aphasia, while reperfusion treatment with thrombolysis and/or thrombectomy was associated with lower odds of severe aphasia. Analysis of individual language dimensions showed that higher NIHSS scores correlated with multidimensional deficits; frontal lobe involvement was primarily associated with reduced performance in information content and naming; and temporal lobe involvement was related to impaired auditory comprehension. Longitudinal analysis further indicated that elevated DBP and a history of diabetes were associated with lower recovery ratio.

3.1 Identifying factors associated with the severity of acute-phase PSA and different linguistic dimensions

3.1.1 Baseline Severity and Language Function Impairment

Our study found that higher NIHSS scores were associated with severe aphasia and reduced performance across various language dimensions, consistent with previous research findings. Agrawal et al. [4] reported that admission NIHSS scores were negatively correlated with AQ at 2 weeks and 3 months post-stroke, as well as with scores in individual language domains at 3 months, indicating that severe neurological deficits at baseline are often accompanied by significant language impairments. Haldin et al. [15] found that for every one-point increase in NIHSS score, the likelihood of mild or no aphasia decreased by 10%, which aligns with our results (OR = 1.106), further supporting the NIHSS score as a key clinical indicator of PSA severity. The association between NIHSS scores and PSA may be related to infarct volume. Previous studies have shown that higher NIHSS scores typically indicate more severe neurological deficits and larger infarct volumes[33, 34]. An expanded infarction area within the left MCA territory not only damages language-related cortex regions, such as Broca's and Wernicke's areas, but may also affect language-related white matter tracts, such as the arcuate fasciculus (AF) and inferior longitudinal

fasciculus (ILF), resulting in disruption of the overall language network structure rather than being limited to a specific dimension. Therefore, patients with high NIHSS scores exhibit not only more severe aphasia, but also deficits across multiple dimensions of language function, including spontaneous speech, auditory comprehension, repetition, and naming.

It should be noted that the NIHSS scale has a total score of 42, with language function accounting for 3 points—only 7.14% of the total. Although this proportion is small, there is some content overlap between the language assessment in NIHSS and the WAB, suggesting that their association may partly be influenced by the composition of the scales. Furthermore, our study found that NIHSS scores are associated with multiple language dimensions, while different infarct locations show specific associations with distinct language domains. This finding suggests that NIHSS may primarily reflect overall post-stroke language impairment, whereas lesion location may further determine the specific pattern of language deficits. Therefore, combining NIHSS scores with imaging characteristics could be more effective than relying solely on NIHSS for early identification of severe PSA patients and for assessing their specific language impairments.

3.1.2 Association between reperfusion therapy and aphasia

Our study found that reperfusion treatment with thrombolysis and/or thrombectomy is associated with lower odds of severe aphasia, consistent with previous research [35–38]. During the hyperacute phase of ischemic stroke, reperfusion therapy restores blood supply to the ischemic penumbra, salvaging brain tissue that has not yet developed irreversible infarction and thereby reducing damage to language-related brain regions [39, 40]. Ahmed et al. [36] observed that improvements in language function after reperfusion were greater than those in motor function. Menichelli et al. [38] found that the proportion of patients recovering from aphasia significantly increased following thrombolytic therapy, indicating that thrombolytic therapy is one of the predictive factors for aphasia recovery. These findings suggest that timely restoration of perfusion may have a protective effect on language function during the acute-phase PSA. After stroke onset, some language-related cortical regions may remain viable before irreversible infarction occurs. Prompt reperfusion can restore their blood flow and neural function, thus mitigating acute language deficits [36]. However, reperfusion therapy has limited efficacy in reversing damage to already infarcted deep white matter fibers. Therefore, the protective effect of reperfusion therapy on PSA may depend on treatment timing, the extent of ischemic tissue, and the degree of damage to the language network.

Our study also found that reperfusion therapy was associated with higher naming scores, which may be related to the anatomy of the language network. Hillis et al. [37, 41] studied patients with naming deficits after stroke using magnetic resonance perfusion-weighted imaging (PWI) on MRI and found that reperfusion in the left BA22, 37, 44, and 45 regions was linked to naming recovery and could predict early improvement in naming. Fridriksson et al. [42] used lesion-symptom mapping techniques and noted that there is no single specific lesion site for naming aphasia. Instead, naming function relies on a widespread cortical network. The frontal, temporal, and parietal cortices collectively participate in language comprehension, lexical retrieval, and speech production [42–44]. Therefore, damage to any of these nodes could lead to naming deficits. Reperfusion therapy may promote naming recovery by restoring blood flow to language-related cortical areas within the ischemic penumbra, thereby improving semantic processing and lexical access. However, as an observational cross-sectional design, the study cannot establish a direct protective effect of reperfusion therapy on naming function. Large-scale experimental studies are needed to further validate these findings in the future.

3.1.3 The Impact of Different Brain Regions on Language Function

Our study found that frontal lobe involvement is associated with severe aphasia, and showed the strongest association with severe aphasia among the infarct locations included in the logistic regression model. As a key component of the frontal language cortex, Broca's area participates in multiple processes during language production, including syntactic processing, semantic selection, motor programming, and lexical retrieval. Together with the middle frontal gyrus (MFG), superior temporal gyrus (STG), and middle temporal gyrus (MTG), it forms a frontotemporal functional network that plays a crucial role in language function [37, 45]. Similarly, the dual-stream model emphasizes the coordinated role of the frontal and temporal lobes, proposing that language processing primarily involves two pathways: a ventral stream primarily involved in speech comprehension and lexical processing, and a dorsal stream involved in speech production and expression. In this model, the inferior frontal gyrus (IFG) is regarded as a critical shared node between the two pathways [42–44, 46]. These findings suggest that frontal lobe damage may not only impair multiple stages of language production but also disrupt network connectivity, thereby increasing the risk of severe aphasia.

Further analysis revealed that infarction involving the frontal lobe is primarily associated with decreased scores in information content and naming across various dimensions of language function. Ding et al. [47] quantified the dimension of information content using the number and production speed of correct information units (CIUs), finding that information content was mainly associated with dorsal white matter pathways such as the AF and superior longitudinal fasciculus (SLF), as well as certain frontal gray matter regions. When the left frontal lobe is affected, core language abilities such as syntactic structure, grammatical processing, and lexical retrieval are impaired, resulting in symptoms like short utterances, incomplete sentence structures, and difficulty finding words, ultimately leading to a significant reduction in the informational content of speech. Naming function also relies on the frontotemporal language network. Thus, frontal lobe involvement may result in diminished naming ability.

Although the frontal and temporal lobes together form a specific language network and play an important role in language function, our logistic regression analysis did not find a significant association between temporal lobe involvement and severe aphasia, which may be related to the composition of the AQ. In the WAB, AQ is composed of information content, fluency, auditory comprehension, repetition, and naming, reflecting overall language function. When a particular brain region primarily affects only one dimension, its overall impact may be diminished due to relatively preserved performance in other dimensions. In our study, frontal lobe involvement was associated with a decrease of 1.604 points (95% CI = -3.155 to -0.053) in the mean information content score and 1.682 points (95% CI = -3.188 to -0.177) in the mean naming score, indicating a broader impact and thus a greater likelihood of severe aphasia. In contrast, temporal lobe involvement led only to a reduction of 1.569 points (95% CI = -3.112 to -0.025) in the average auditory comprehension score, which is less extensive both in terms of the number of affected functional domains and the magnitude of impact on specific dimensions compared to

frontal lobe involvement, suggesting a more limited effect on overall severity. It should be noted that these explanations are based on theoretical mechanisms, and the association between temporal lobe involvement and aphasia severity still requires more direct statistical validation.

Our study found that temporal lobe involvement was not significantly associated with overall aphasia severity, but it is associated with decreased performance in auditory comprehension, consistent with the traditional view that the left temporal lobe is the core cortical region for language comprehension. Previous research has shown that the extent and range of temporal lobe damage are closely related to deficits in auditory comprehension [37, 47–51]. Salvas et al. [49] found that comprehension difficulties with atypical sentences—those deviating from the conventional subject-verb-object structure—are primarily associated with the supramarginal gyrus (SMG), STG, and MTG, while impairments in understanding sequential commands were associated with damage to the STG, MTG, and insula. They proposed that the MTG plays a crucial role in recovery of auditory comprehension. Similarly, Lwi et al. [50] observed that single-word comprehension, yes/no question comprehension, and sequence command comprehension were all associated with the left MTG, particularly its posterior portion. Biondo et al. [51] further indicated that language comprehension not only depends on cortical regions such as the posterior and middle parts of the STG, MTG, and superior temporal sulcus (STS), but also closely involves white matter tracts including the ILF, inferior fronto-occipital fasciculus (IFOF), middle longitudinal fasciculus (MdLF), and uncinate fasciculus (UF). Therefore, auditory comprehension relies on a language network composed of temporal lobe cortical regions and frontotemporal white matter tracts. The findings of our study further support the critical role of the temporal lobe in auditory comprehension.

3.2 Exploring the association between baseline factors and language recovery ratio: cardiovascular and metabolic factors

During cohort follow-up, our study found that elevated DBP at onset was associated with a reduced ratio of language function recovery. Previous research has primarily focused on the relationship between acute-phase blood pressure (BP) and BP variability with overall stroke outcomes, while studies examining the link between acute-phase BP and PSA recovery have been limited. Prior evidence indicates that chronic hypertension can lead to endothelial dysfunction, hyaline degeneration, and fibrinoid necrosis, damaging small vessels and the blood-brain barrier, thereby promoting cerebral small vessel disease; at the same time, prolonged hypoperfusion may result in damage to deep white matter and axonal demyelination, impairing the remodeling capacity of white matter tracts such as the AF and SLF, thus reducing language recovery [52–54]. Additionally, hypertension induces microglia to release inflammatory factors, which in turn activate astrocytes, ultimately leading to neuroinflammation and neurotoxic effects; it also reduces the number and density of neuronal dendritic spines, diminishing neural plasticity [55–57]. Therefore, long-term elevated BP may impair language recovery in PSA patients through mechanisms involving small-vessel brain injury, white matter damage, and decreased neural plasticity. However, our study examined the association between a single measurement of DBP during the acute phase and language recovery, rather than the effects of chronic high diastolic pressure, so these mechanisms cannot directly explain the current findings. Moreover, few previous studies have examined the relationship between acute-phase DBP and PSA prognosis, and direct evidence supporting our findings remains limited. Thus, these findings need to be further validated in large-scale, prospective studies with standardized BP monitoring.

Meanwhile, our study found only a correlation between DBP and recovery ratio, with no significant association observed for systolic blood pressure (SBP). To avoid missing potential confounding factors, we included SBP (P = 0.149 in univariate analysis) in the linear regression model as an additional covariate, and no significant association was observed between SBP and recovery ratio (P = 0.793). Several factors may account for this finding. First, acute-phase BP is influenced by pathological, physiological, and psychological factors. Therefore, a single measurement may not accurately reflect a patient's usual BP level [58–60]. Second, through 24-hour continuous BP monitoring after admission, Geng et al. [61] demonstrated that SBP between 4:00 AM and 8:00 AM was associated with patient outcomes. In contrast, BP in our study was primarily measured between 10:00 AM and 4:00 PM. This difference in timing may explain why we did not observe an association between SBP and recovery. Additionally, the follow-up cohort included only 37 patients, resulting in a small sample size that cannot rule out the possibility of a true association between SBP and language recovery that went undetected—i.e., a type II error.

Our study also found that a history of diabetes was associated with poorer recovery ratio. Although there are currently few studies directly addressing the relationship between diabetes and PSA prognosis, existing research has shown that diabetes can reduce the benefits of speech therapy in patients. Roth et al. [17] found that diabetes may diminish the effectiveness of speech therapy in chronic PSA patients. Through further analysis using diffusion tensor imaging (DTI) on MRI, they revealed that diabetes may modify the relationship between structural network integrity and treatment outcomes—specifically, structural network integrity predicted treatment response in non-diabetic individuals but not in individuals with diabetes. This suggests that microstructural damage to white matter fibers caused by diabetes may impair the reorganization of language networks. Previous DTI-based studies have demonstrated reduced fractional anisotropy (FA) in white matter tracts among diabetic patients, indicating damage to language-related fiber bundles such as the ILF, IFOF, and UF [62, 63], which could hinder the remodeling of language networks. Additionally, pathophysiological changes associated with diabetes—including insulin resistance, endothelial dysfunction, and neuroinflammation—may further damage cerebral small vessels and compromise the blood-brain barrier, thereby affecting neural plasticity [64]. Therefore, diabetes may impair language recovery through mechanisms involving white matter microstructural damage, cerebral small vessel disease, and reduced neural plasticity. However, our study did not include imaging markers such as DTI, limiting the ability to confirm a causal link between diabetes-related white matter injury and language recovery, and the underlying mechanisms require further investigation.

3.3 Research Significance and Clinical Implications

This study has preliminarily established a clinical framework for the early identification and stratified management of acute-phase PSA by integrating multiple factors, including disease severity, lesion location, reperfusion therapy, and systemic vascular-metabolic conditions. Admission NIHSS scores and frontal lobe involvement may serve as simple indicators for identifying patients at high risk of severe aphasia. These indicators may facilitate early language screening and support the timely allocation of speech rehabilitation resources. Furthermore, the associations between different lesion locations and specific language domains provide imaging-based evidence for targeted language assessment and individualized rehabilitation training. Moreover, reperfusion

therapy is associated with lower odds of severe aphasia and higher naming scores, suggesting that timely restoration of cerebral perfusion—when clinically indicated—may be associated with better language outcomes. However, it should be noted that our study is observational and cannot establish causality between reperfusion therapy and language outcomes. Longitudinal analysis further suggests that acute-phase DBP and history of diabetes are associated with language recovery, indicating that long-term recovery in PSA may be closely associated with overall vascular and metabolic status. Therefore, clinical management of PSA should not be limited to language therapy alone but should integrate early language assessment, lesion characterization, reperfusion treatment, and comprehensive management of vascular and metabolic risk factors such as BP and blood glucose control, and ongoing rehabilitation follow-up. For patients unable to receive systematic language therapy due to economic, time, or resource constraints, our findings show that some individuals still achieve notable spontaneous language recovery within 2 to 6 months, with a median recovery ratio of 51.1% (IQR: 27.80 to 79.75), offering valuable insights for prognosis communication and rehabilitation planning. Overall, these results support a shift in acute-phase PSA management from mere identification of language deficits toward risk stratification, functional dimension evaluation, and comprehensive care, laying the foundation for future multicenter, prospective studies and precision-based rehabilitation interventions tailored to individual patient characteristics.

3.4 Limitations

Several limitations should be acknowledged. First, this was a single-center study with a relatively small sample size, particularly in the follow-up cohort, which may have limited the statistical power of the longitudinal analysis. The follow-up completion rate was 36.3%, and some patients were lost to follow-up for reasons potentially related to their recovery status, suggesting possible non-random attrition. In addition, some categorical variables showed complete separation and were excluded from the regression models, potentially leading to the omission of certain relevant information. Therefore, the longitudinal findings should be interpreted cautiously. Second, the absence of advanced neuroimaging, including DTI and detailed perfusion or functional imaging, limited our ability to characterize the structural and functional connectivity of the language network. Furthermore, language function was not assessed immediately before and after reperfusion therapy because of clinical workflow constraints, precluding evaluation of its immediate effects on language function. Third, only a single blood pressure measurement was available, preventing assessment of acute-phase blood pressure dynamics and variability. In addition, intravenous thrombolysis and mechanical thrombectomy were analyzed as a single reperfusion category because of the limited sample size, precluding comparisons between different reperfusion strategies. Finally, the strict study criteria and uneven distribution of aphasia subtypes may have limited the generalizability of our findings. In particular, although our study did not selectively include or exclude specific aphasia types, no patients were classified as having sensory or transcortical sensory aphasia, because acute-phase aphasia may change dynamically over time[65, 66]. Therefore, a single assessment of aphasia type may not accurately reflect the patient's subsequent stable aphasia phenotype. Future multicenter, prospective studies with larger cohorts, standardized longitudinal assessments, and comprehensive neuroimaging are warranted to validate these findings.

Conclusions

This study, based on patients with post-stroke aphasia during the acute phase, employed an observational design combining cross-sectional analysis with cohort follow-up to systematically investigate factors associated with post-stroke aphasia severity and language recovery. The results indicate that the severity and recovery of aphasia after acute stroke may be influenced by multiple factors, including overall stroke severity, receipt of reperfusion therapy, lesion location, and comorbid conditions. These findings add to the existing evidence on factors associated with PSA severity and recovery, providing preliminary evidence for early risk assessment and intervention in patients with post-stroke aphasia, as well as a foundation and direction for future large-scale, multicenter, multimodal neuroimaging studies aimed at elucidating the mechanisms underlying language networks.

Abbreviations

ADC apparent diffusion coefficient

AF arcuate fasciculus

AQ aphasia quotient

BMI body mass index

BP blood pressure

CI confidence interval

CIUs correct information units

COVID-19 Coronavirus Disease 2019

CT computed tomography

CTA CT angiography

CTP CT perfusion

DBP diastolic blood pressure

DTI diffusion tensor imaging

DWI diffusion-weighted imaging

EHI Edinburgh Handedness Inventory

FA fractional anisotropy

HbA1c glycated hemoglobin

HDL-C high-density lipoprotein cholesterol

IFG inferior frontal gyrus

IFOF inferior fronto-occipital fasciculus

ILF inferior longitudinal fasciculus

INR International normalized ratio

IQR interquartile range

LDL-C low-density lipoprotein cholesterol

MCA middle cerebral artery

MdLF middle longitudinal fasciculus

MFG middle frontal gyrus

MIT melodic intonation therapy

MRI magnetic resonance imaging

MTG middle temporal gyrus

NIHSS National Institutes of Health Stroke Scale

OR odds ratio

PSA post-stroke aphasia

PWI perfusion-weighted imaging

SBP systolic blood pressure

SLF superior longitudinal fasciculus

SLT speech and language therapy

SMG supramarginal gyrus

STG superior temporal gyrus

STS superior temporal sulcus

TOAST trial of ORG 10172 in acute stroke treatment

UF uncinate fasciculus

WAB Western Aphasia Battery

Declarations

Ethics approval and consent to participate

The study was approved by the Medical Ethics Committee of the First Affiliated Hospital of Soochow University [Approval Number: (2026) Ethics Approval No. 1197] and was conducted in accordance with the Declaration of Helsinki. The requirement for informed consent was waived by the Medical Ethics Committee of the First Affiliated Hospital of Soochow University because of the non-interventional nature of the study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the funding from the Natural Science Foundation of Jiangsu Province (BK20240367), the Pilot Project for Basic Research of Suzhou City (Project Number: SSD2025050) and the Horizontal Project of Soochow University (Project Number: H240792).

Author Contribution

CGQ and LJ contributed equally to the study design, manuscript writing, and revision. LXY, ZYY, SXL, and DH participated in data collection, data organization, and result analysis. GMY and ZQQ jointly supervised the study, performed quality control, and critically revised the manuscript. All authors read and approved the final manuscript.

Acknowledgement

The authors thank all patients included in this study, as well as the healthcare personnel involved in patient identification, data collection, and follow-up.

Data Availability

The datasets used during the current study are available from the corresponding author on reasonable request.

References

1. Vitti E, Hillis AE. Treatment of post-stroke aphasia: A narrative review for stroke neurologists. Int J Stroke. 2021;16(9):1002–8.
2. Berthier ML. Poststroke Aphasia: Epidemiology, Pathophysiology and Treatment. Drugs Aging. 2005;22(2):163–82.
3. Szaflarski JP, Binder JR, Possing ET, McKiernan KA, Ward BD, Hammeke TA. Language lateralization in left-handed and ambidextrous people: fMRI data. Neurology. 2002;59:238–44.
4. Agrawal M, Singh VK, Verma A, Pathak A, Kumar A, Joshi D, et al. Clinico-radiological factors associated with aphasia outcome in post stroke patients: A prospective follow up study from eastern part of India. J Clin Neurosci. 2024;123:130–6.
5. Liang H, Ren H, Yang X, Nie Z, Huang X. Investigation and analysis of factors influencing quality of life in patients with post-stroke aphasia. Chin J Mod Nurs. 2015(30):3688–90. (in Chinese).
6. Vogel-Eyny A, Jung Y-J, Hyun J, Siegle E, Haroun A, Nahar S, et al. Prevalence of depression in post-stroke aphasia: systematic review and meta-analysis. BMC Neurol. 2026;26(1):43.
7. Marebwa BK, Fridriksson J, Yourganov G, Feenaughty L, Rorden C, Bonilha L. Chronic post-stroke aphasia severity is determined by fragmentation of residual white matter networks. Sci Rep. 2017;7(1):8188.
8. Harvey DY, Parchure S, Hamilton RH. Factors predicting long-term recovery from post-stroke aphasia. Aphasiology. 2022;36(11):1351–72.
9. Levy DF, Entrup JL, Schneck SM, Onuscheck CF, Rahman M, Kasdan A, et al. Multivariate lesion symptom mapping for predicting trajectories of recovery from aphasia. Brain Commun. 2023;6(1):fcae024.
10. Kropp E, Varkanitsa M, Carvalho N, Falconer I, Billot A, Al-Dabbagh M, et al. Using unsupervised dimensionality reduction to identify lesion patterns predictive of post-stroke aphasia severity. Cortex. 2025;188:25–41.
11. Yu Q, Sun Y, Ju X, Ye T, Liu K. Prediction models of the aphasia severity after stroke by lesion load of cortical language areas and white matter tracts: An atlas-based study. Brain Res Bull. 2024;217:111074.
12. LI H. Analysis of the relationship between aphasia subtypes and lesion locations after cerebral infarction. China Foreign Med Treat. 2014;33(32):73–4. (in Chinese).
13. Yao J, Liu X, Chen H, Wang S, Zhang Y. Factors Associated with the Type of Post-aphasia: Implications from a Single-center Retrospective Study. Chin J Stroke. 2020;15(8):853–60. (in Chinese).
14. Kim KA, Lee JS, Chang WH, Kim DY, Shin Y-I, Kim S-Y, et al. Changes in Language Function and Recovery-Related Prognostic Factors in First-Ever Left Hemispheric Ischemic Stroke. Annals Rehabilitation Med. 2019;43(6):625–34.
15. Haldin C, Dai S, Piscicelli C, Marcon V, Lœvenbruck H, Pérennou D, et al. Multifactorial interplay on language recovery following left-hemispheric stroke: A retrospective study. J Neuropsychol. 2025;19(2):256–73.
16. Zhao J. Blood glucose level in patients with acute cerebral infarction and prognosis of clinical observation. China Mod Med. 2010;17(27):51–2. (in Chinese).
17. Roth R, Busby N, Wilmskoetter J, Schwen Blackett D, Gleichgerrcht E, Johnson L, et al. Diabetes, brain health, and treatment gains in post-stroke aphasia. Cereb Cortex. 2023;33(13):8557–64.

18. Roberts SM, Bruce R, Hope TMH, Geva S, Anderson S, Woodgate H, et al. The impact of pre-stroke formal education on language test performance in aphasic and non-aphasic stroke survivors. Aphasiology. 2025;39(11):1480–502.
19. Wang X. Study on Rehabilitation Training in Patients with Aphasia. J Math Med. 2016;29(11):1727–8. (in Chinese).
20. Kiran S, Thompson CK. Neuroplasticity of Language Networks in Aphasia: Advances, Updates, and Future Challenges. Front Neurol. 2019;10:295.
21. Braun EJ, Billot A, Meier EL, Pan Y, Parrish TB, Kurani AS, et al. White matter microstructural integrity pre- and post-treatment in individuals with chronic post-stroke aphasia. Brain Lang. 2022;232:105163.
22. Kernbach JM, Hartwigsen G, Lim J-S, Bae H-J, Yu K-H, Schlaug G, et al. Bayesian stroke modeling details sex biases in the white matter substrates of aphasia. Commun Biology. 2023;6(1):354.
23. Ali M, VandenBerg K, Williams LJ, Williams LR, Abo M, Becker F, et al. Predictors of Poststroke Aphasia Recovery: A Systematic Review-Informed Individual Participant Data Meta-Analysis. Stroke. 2021;52(5):1778–87.
24. Zhou L, Ye J. Analysis on Aphasia after Stroke, Prognosis and Other Related Ingredients. Guide China Med. 2013;11(19):22–4. (in Chinese).
25. Wang Y, Sun C, Bai Y. Retrospective analysis of factors influencing the type and severity of aphasia in patients with post-stroke aphasia. Chin J Rehabilitation. 2024;39(6):323–8. (in Chinese).
26. Neurology CSo, Society CS. Chinese guidelines for diagnosis and treatment of acute ischemic stroke 2018. Chin J Neurol. 2018;51(9):666–82. (in Chinese).
27. Wang Y. Introduction to the Western Aphasia Battery (WAB) (Part). Chin J Rehabilitation Theory Pract. 1997;3(2):87–9. (in Chinese).
28. Wang Y. Introduction to the Western Aphasia Battery (WAB) (Part). Chin J Rehabilitation Theory Pract. 1997;3(3):135–40. (in Chinese).
29. R.C.Oldfield. The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia. 1971;9(1):97–113.
30. Cherney LR, Kaye RC, Van Vuuren S. Acquisition and Maintenance of Scripts in Aphasia: A Comparison of Two Cuing Conditions. Am J Speech-Language Pathol. 2014;23(2):S343–60.
31. Wang L, Li C, Liu L, Zhang W, Liu Y. The effect of venlafaxine on language function in patients with subcortical aphasia. J Psychiatr Res. 2025;186:172–91.
32. Lazar RM, Minzer B, Antoniello D, Festa JR, Krakauer JW, Marshall RS. Improvement in Aphasia Scores After Stroke Is Well Predicted by Initial Severity. Stroke. 2010;41(7):1485–8.
33. Helenius J, Henninger N. Leukoaraiosis Burden Significantly Modulates the Association Between Infarct Volume and National Institutes of Health Stroke Scale in Ischemic Stroke. Stroke. 2015;46(7):1857–63.
34. Weigel K, Gaser C, Brodoehl S, Wagner F, Jochmann E, Güllmar D, et al. Acute Stroke Severity Assessment: The Impact of Lesion Size and Functional Connectivity. Brain Sci. 2025;15(7):735.
35. Furlanis G, Ridolfi M, Polverino P, Menichelli A, Caruso P, Naccarato M, et al. Early Recovery of Aphasia through Thrombolysis: The Significance of Spontaneous Speech. J Stroke Cerebrovasc Dis. 2018;27(7):1937–48.
36. Ahmed Z, Yedavalli V, Gonzalez WS, Hillis AE. Relative improvement in language vs. motor functions with reperfusion therapies for large vessel occlusion. Sci Rep. 2025;15(1):6683.
37. Hillis AE, Kleinman JT, Newhart M, Heidler-Gary J, Gottesman R, Barker PB, et al. Restoring Cerebral Blood Flow Reveals Neural Regions Critical for Naming. J Neurosci. 2006;26(31):8069–73.
38. Menichelli A, Furlanis G, Sartori A, Ridolfi M, Naccarato M, Caruso P, et al. Thrombolysis' benefits on early post-stroke language recovery in aphasia patients. J Clin Neurosci. 2019;70:92–5.
39. Olivot J-M, Mlynash M, Thijs VN, Kemp S, Lansberg MG, Wechsler L, et al. Relationships Between Infarct Growth, Clinical Outcome, and Early Recanalization in Diffusion and Perfusion Imaging for Understanding Stroke Evolution (DEFUSE). Stroke. 2008;39(8):2257–63.
40. Mikulik R, Dusek L, Hill MD, Fulep E, Grotta JC, Ribo M, et al. Pattern of Response of National Institutes of Health Stroke Scale Components to Early Recanalization in the CLOTBUST Trial. Stroke. 2010;41(3):466–70.
41. Hillis AE, Kane A, Tuffiash E, Ulatowski JA, Barker PB, Beauchamp NJ, et al. Reperfusion of Specific Brain Regions by Raising Blood Pressure Restores Selective Language Functions in Subacute Stroke. Brain Lang. 2001;79(3):495–510.
42. Fridriksson J, Den Ouden D-B, Hillis AE, Hickok G, Rorden C, Basilakos A, et al. Anatomy of aphasia revisited. Brain. 2018;141(3):848–62.
43. Wang Y, Zhang J, Zhong S, Lou J, Tang Y, Han C, et al. Systematic review of the rehabilitation mechanisms of post-stroke subcortical aphasia based on multimodal neuroimaging. Chin J Rehabilitation Med. 2024;39(11):1724. (in Chinese).
44. Fan R, Li X, Huang X, Xu M, Cao Y, Zhang D, et al. Functional Connectivities in Right Hemisphere for Post-stroke Aphasia: Based on Dual Stream Model. Chin J Rehabilitation Theory Pract. 2020;26(5):572–8. (in Chinese).
45. Bulut T. Domain-general and domain-specific functional networks of Broca's area underlying language processing. Brain Behav. 2023;13(7):e3046.
46. Liuzzi AG, Bruffaerts R, Peeters R, Adamczuk K, Keuleers E, De Deyne S, et al. Cross-modal representation of spoken and written word meaning in left pars triangularis. NeuroImage. 2017;150:292–307.
47. Ding J, Middleton EL, Mirman D. Impaired discourse content in aphasia is associated with frontal white matter damage. Brain Commun. 2023;5(6):fcad310.
48. Liu Q-q, Li W-b, Song Y-w, Zhao Z-b, Yan Y-n, Yang Y-h, et al. Correlation between focal lesion sites and language deficits in the acute phase of post-stroke aphasia. Folia Neuropathol. 2022;60(1):60–8.

49. Salvas C, Paré J, Brambati SM, García AO, Brisebois A, Marcotte K. A Longitudinal Study of Auditory Comprehension in Poststroke Aphasia. Am J Speech-Language Pathol. 2025;34:2212.
50. Lwi SJ, Herron TJ, Curran BC, Ivanova MV, Schendel K, Dronkers NF, et al. Auditory Comprehension Deficits in Post-stroke Aphasia: Neurologic and Demographic Correlates of Outcome and Recovery. Front Neurol. 2021;12:680248.
51. Biondo N, Ivanova MV, Pracar AL, Baldo J, Dronkers NF. Mapping sentence comprehension and syntactic complexity: evidence from 131 stroke survivors. Brain Commun. 2024;6(6):fcae379.
52. Hannan J, Busby N, Roth R, Wilmskoetter J, Newman-Norlund R, Rorden C, et al. Under pressure: the interplay of hypertension and white matter hyperintensities with cognition in chronic stroke aphasia. Brain Commun. 2024;6(3):fcae200.
53. Osman A, Kanaan A, Azeroual S, Abubakr M, Jwayyed J, Bhutta R. Hypertension and Cerebral Small Vessel Disease: A Review of the Pathophysiology, Progression, and Prevention. Cureus. 2025.
54. Hainsworth AH, Markus HS, Schneider JA. Cerebral Small Vessel Disease, Hypertension, and Vascular Contributions to Cognitive Impairment and Dementia. Hypertension. 2024;81(1):75–86.
55. Maïer B, Kubis N. Hypertension and Its Impact on Stroke Recovery: From a Vascular to a Parenchymal Overview. Neural Plast. 2019;2019:1–14.
56. Liddelow SA, Guttenplan KA, Clarke LE, Bennett FC, Bohlen CJ, Schirmer L, et al. Neurotoxic reactive astrocytes are induced by activated microglia. Nature. 2017;541(7638):481–7.
57. Flores G, Flores-Gómez GD, De Jesús Gomez-Villalobos M. Neuronal changes after chronic high blood pressure in animal models and its implication for vascular dementia: High Blood Pressure and Brain. Synapse. 2016;70(5):198–205.
58. Carlberg B, Asplund K, Hägg E. Factors influencing admission blood pressure levels in patients with acute stroke. Stroke. 1991;22(4):527–30.
59. Toyoda K, Okada Y, Fujimoto S, Hagiwara N, Nakachi K, Kitazono T, et al. Blood Pressure Changes During the Initial Week After Different Subtypes of Ischemic Stroke. Stroke. 2006;37(10):2637–9.
60. Fischer U, Cooney MT, Bull LM, Silver LE, Chalmers J, Anderson CS, et al. Acute post-stroke blood pressure relative to premorbid levels in intracerebral haemorrhage versus major ischaemic stroke: a population-based study. Lancet Neurol. 2014;13(4):374–84.
61. Geng X, Liu X, Li F, Wang J, Sun H, Feng A, et al. Blood pressure variability at different time periods within first 24 hours after admission and outcomes of acute ischemic stroke. J Clin Hypertens. 2020;22(2):194–204.
62. Zhou C, Li J, Dong M, Ping L, Lin H, Wang Y, et al. Altered White Matter Microstructures in Type 2 Diabetes Mellitus: A Coordinate-Based Meta-Analysis of Diffusion Tensor Imaging Studies. Front Endocrinol. 2021;12:658198.
63. Huang H, Ma X, Yue X, Kang S, Li Y, Rao Y, et al. White Matter Characteristics of Damage Along Fiber Tracts in Patients with Type 2 Diabetes Mellitus. Clin Neuroradiol. 2023;33(2):327–41.
64. Krinock MJ, Singhal NS. Diabetes, stroke, and neuroresilience: looking beyond hyperglycemia. Ann N Y Acad Sci. 2021;1495(1):78–98.
65. Zhang T, Li S, Bai Y, Chen Z, Deng J, Gong W, et al. Expert consensus on the clinical management of post-stroke aphasia. Chin J Rehabilitation Theory Pract. 2022;28(1):15–23. (in Chinese).
66. Wilson SM, Entrup JL, Schneck SM, Onuscheck CF, Levy DF, Rahman M, et al. Recovery from aphasia in the first year after stroke. Brain. 2023;146(3):1021–39.

Figures

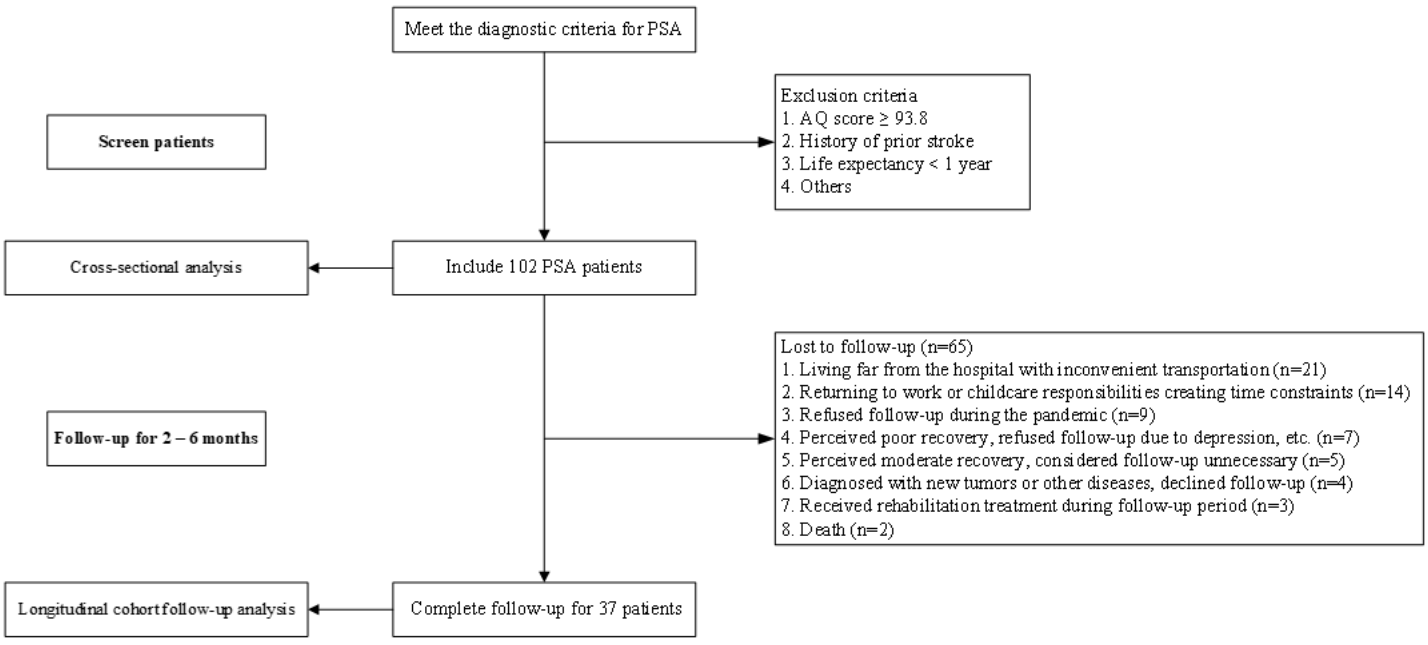


Figure 1

Flowchart of Patient Selection and Study Procedures

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable.docx