A Nationwide Population-Based Cohort Study of Epilepsy Incidence in Patients with Post-Stroke Aphasia

Hui-Lin Lin

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Public Health

University of Washington

2025

Committee:

Noel S. Weiss

Nicholas L. Smith

Amanda Phipps

Program Authorized to Offer Degree:

Department of Epidemiology

University of Washington

Abstract

A Nationwide Population-Based Cohort Study of Epilepsy

Incidence in Patients with Post-Stroke Aphasia

Hui-Lin Lin

Chair of the Supervisory Committee:

Noel S. Weiss
Department of Epidemiology

Post-stroke epilepsy (PSE) affects 2–20% of stroke survivors. Post-stroke aphasia is associated with a further increase in this risk. Identifying factors associated with developing PSE potentially could guide targeted interventions in future trials. We utilized Taiwan's National Health Insurance claims data to identify individuals aged 18 years or older who were hospitalized for a first stroke between 2003 and 2007, including those diagnosed with aphasia during admission or within 90 days after discharge. Beginning 91 days following hospital discharge, the incidence of epilepsy was compared between stroke patients with and without aphasia until December 31, 2020. Hazard ratios controlling for other risk factors for epilepsy were derived using Cox proportional hazards regression, and we calculated adjusted rate differences to quantify the absolute magnitude of the excess risk. With a median follow-up of 7.56 years for the aphasia group (n = 23,431) and 8.67 years for the non-aphasia group (n =

130,058), the incidence of post-stroke epilepsy was higher among patients with aphasia than in those without aphasia (13.02 vs. 4.56 per 1000 person-years, adjusted rate difference [aRD] 2.96, 95% CI 2.60–3.37 per 1,000 person-years). Excess risk was similar (aRD 2.38 and 3.11 per 1,000 person-years) in female and male participants, respectively, and somewhat higher in younger than in older adults (aRD 3.52, 2.42, and 2.88 per 1,000 person-years in the <45, 45–64, and ≥65 age groups, respectively). The excess rate of epilepsy associated with the development of aphasia was 3.52 and 2.64 per 1,000 person-years among those with hemorrhagic and ischemic stroke, respectively. The association was strongest in the first year of follow-up after stroke and declined thereafter. Patients with post-stroke aphasia are at increased risk of epilepsy irrespective of age, sex, and stroke type.

Introduction

Post-stroke epilepsy (PSE) is a severe and common adverse complication of a stroke, affecting 2–14 percent of ischemic stroke survivors and 10–20 percent of those with hemorrhagic stroke[1]. Identifying factors associated with developing PSE potentially could lead to the identification of persons at particularly high risk for selective inclusion in future treatment trials of stroke patients that are aimed at reducing the risk of epilepsy.

Aphasia is an impairment of language production, comprehension, reading, or writing, most often resulting from injury to dominant-hemisphere language networks (left perisylvian cortex) following stroke or traumatic brain injury[2,3]. Aphasia is present in approximately one-third of patients after ischemic stroke[4]. Prior evidence linking post-stroke aphasia(PSA) to subsequent epilepsy is limited, coming largely from hospital-based cohorts with small samples[5,6] and short follow-up periods[5,6]. Previous studies have evaluated post-stroke epilepsy risk overall without stratifying by aphasia status[7,8]. A large population-based study in young adults observed aphasia to be an independent risk factor for PSE (aHR, 1.77; 95% CI, 1.20–2.60)[9]. However, that study was restricted to 19–44-year-olds with ischemic stroke [9]. We conducted a nationwide, population-based retrospective cohort study using Taiwan's National Health Insurance claims to compare the risk of PSE between stroke patients with and without aphasia, stratified by age, sex, and stroke type.

Methods

Data source

Launched in 1995, Taiwan's National Health Insurance (NHI) has mandated government-funded healthcare for nearly all 23 million people, achieving 99.5 percent enrollment[10]. The health insurance database maintains a comprehensive record of each insured person's visit dates, operation and treatment codes, and prescription and medication details for various types of medical services, including outpatient, inpatient, dental, traditional Chinese medicine, long-term care, and preventive care. Diagnoses were coded using the National Health Insurance (NHI) A code system and the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) before 2016, and have been coded with ICD-10-CM since 2016. These data are securely stored at the Welfare Data Science Center of the Ministry of Health and Welfare of Taiwan and are maintained under strict confidentiality by the NHI Administration. By linking claims records to the National Cause of Death Registry, we obtain each patient's survival status as well as the exact date and underlying cause of death. All personal identifiers used for linking databases had been recorded before being released to users for privacy protection. This study was approved by the Research Ethics Committee of China Medical University and Hospital, Taiwan (CMUH111-REC2-155(CR-2)).

Study population

For the present investigation, patients aged 18 years or older admitted for a first stroke between 2003 and 2007 were identified from claims data using codes for ischemic strokes (ICD-9-CM 433.x) and hemorrhagic strokes (ICD-9-CM 430.x–432.x, 434.x, 436.x) (see

Supplementary Table 1). We excluded patients with a stroke diagnosis before 2003 or who died or withdrew from NHI within 90 days of discharge (Fig. 1). Patients with a prior aphasia diagnosis or who underwent speech therapy before their stroke admission, as well as those lacking continuous NHI coverage during the two years before hospitalization, were excluded. Patients were excluded if they had an epilepsy or seizure diagnosis in any inpatient or outpatient claim before, during, or within 90 days after stroke hospitalization, so as to ensure that epilepsy (outcome) occurred only after the development of aphasia (exposure) (Fig. 1) We identified from the claims data those patients with a new diagnosis of aphasia (ICD-9-CM code 438.1, primarily referring to speech and language deficits that developed following the stroke) and those who had received speech therapy during the first stroke hospitalization and within three months after discharge. Patients with stroke were divided into two groups based on the presence or absence of aphasia. In the full cohort of 153,489 stroke patients (Figure 1), 23,431 were classified into the aphasia group and 130,058 into the non-aphasia group. We also established two propensity score–matched cohorts of 22,913 patients each, one with aphasia and one without, for further analyses. (Fig. 1; Table 1).

Ascertainment of potential confounders

Potential confounders in our analysis included stroke severity, patient characteristics, and the presence of comorbidities at the time of the stroke. Prior research has identified stroke severity as a major determinant in the onset of PSE[9,11,12]. Our assessment of stroke severity was based

on the following: total length of hospital stay; ICU admission; and receipt of mechanical ventilation (intubation), airway suctioning, nasogastric tube placement, and/or a urinary catheter [13].

Patients were classified as having a comorbidity - diabetes mellitus[14], hypertension[14,15], hyperlipidemia[1,16], peripheral arterial disease[17], coronary heart disease[17], atrial fibrillation[18], malignancy[19], and chronic obstructive pulmonary disease[20] - if the corresponding diagnostic code (listed in Supplementary Table 1) appeared in at least two outpatient visits or one hospital admission during the two years preceding their stroke.

Study outcome and follow-up

Each of the two study cohorts was monitored from baseline until the first of the following events: development of PSE, death, disenrollment from the NHI program, or December 31, 2020.PSE cases were identified by the diagnosis of an ICD-9-CM (345.0–345.9) or ICD-10-CM (G40.0–G40.9) code (see Supplementary Table 1) recorded in at least two outpatient visits or one inpatient admission, with each diagnosis verified by the treating neurologist to ensure accuracy. Outcome status was ascertained through NHI claim records and linkage to the National Death Register for vital status.

Statistical analysis

We examined baseline demographics, stroke characteristics, and comorbidities between patients with and without PSA in both the unmatched cohort and the propensity score–

matched cohort. We quantified imbalance using the standardized mean difference, calculated as the absolute difference in the mean or proportion of each variable between groups divided by the pooled standard deviation for that variable. We then calculated incidence rates of post-stroke epilepsy for each cohort by sex, age group, and stroke type and summarized absolute differences as rate differences (RDs); adjusted RDs were estimated as $r_0 \times (aHR - 1)$, with r_0 denoting the non-aphasia rate in each stratum; 95% confidence intervals were derived using the corresponding lower and upper limits of the adjusted hazard ratio. We applied Cox regression to calculate hazard ratios and 95% confidence intervals for PSE, comparing patients with post-stroke aphasia to those without.

In a second analysis, we estimated each patient's probability of post-stroke aphasia (exposure) using a multivariable logistic regression that included age, sex, stroke type (ischemic vs. hemorrhagic), numbers of outpatient visits and hospitalizations in the two years before stroke, index length of stay, markers of stroke severity (ICU admission, and the receipt of mechanical ventilation, nasogastric tube, urinary catheter, airway suctioning), the presence of comorbidities, and calendar year of the index stroke. Patients with aphasia were then matched 1:1 to patients without aphasia via nearest-neighbor matching with a caliper of 0.2 standard deviations of the logit of the propensity score[21]. We re-estimated the association between aphasia and PSE (outcome) in this matched sample to assess the robustness of the findings.

In the unmatched cohort, we applied regression modeling that adjusted for relevant covariates to calculate the adjusted hazard ratio for PSE associated with poststroke aphasia. In the propensity score–matched cohort, we fitted Cox proportional hazards models with robust variance estimation to derive adjusted hazard ratios for post-stroke epilepsy and related outcomes in patients with aphasia[22,23]. In neither analysis did we adjust for time since stroke, given that the distribution of person-time following stroke was nearly identical for patients with and without aphasia. We estimated hazard ratios for PSE within discrete follow-up intervals (≤1 year, 1–2 years, 2–3 years, and >3 years after stroke) to allow risk estimates to vary over time. All data extraction, computation, and statistical analyses were carried out in SAS version 9.4 (SAS Institute Inc., Cary, NC, USA).

Results

Baseline Cohort Characteristics

Between 2003 and 2007, 205,448 patients were admitted for a first stroke. After applying exclusion criteria, 153,489 remained, including 23,431 (15.3%) patients with aphasia in the first 90 days after stroke and 130,058 without aphasia (84.7 percent) (Figure 1). Approximately 60 percent of participants were male and ≥65 years, and hypertension was common in both groups (88.3 percent vs. 84.9 percent) (Table 1). The aphasia group had experienced a higher proportion of hemorrhagic strokes (30.3 percent vs. 20.7 percent), longer hospital stays (median 16 days vs. 7 days), and greater use of intensive care, although their

average follow-up was slightly shorter. In our second analysis, after propensity score matching on 22,913 pairs, all baseline characteristics were well aligned between the two groups (Table 1).

Risk of Epilepsy Following Post-Stroke Aphasia

Over the 18-year observation period, in both analyses, the incidence of PSE was higher in patients with than without aphasia. The crude PSE rate among stroke survivors with aphasia was 13.02 per 1,000 person-years, compared with 4.56 per 1,000 person-years in those without aphasia. The adjusted hazard ratio was 1.65 (95% CI 1.57–1.74), and the adjusted rate difference was 2.96 per 1,000 person-years (95% CI 2.60–3.37). In the propensity score–matched sample(22,913 pairs), the corresponding PSE rates were 12.72 vs. 7.66 per 1,000 person-years. The adjusted hazard ratio was 1.69 (95% CI 1.58–1.81; Table 2), and the adjusted rate difference was 5.06 per 1000 person-years (95% CI 4.40-5.72).

Adjusted hazard ratios and rate differences were elevated across all subgroups in both the entire cohort and the propensity score–matched analysis (Table 2). Across all age strata (<45, 45–64, ≥65 years), patients with aphasia had higher absolute rates of PSE than those without aphasia. The adjusted excess epilepsy rates per 1,000 person-years were 3.52 (95% CI 2.06–5.22), 2.42 (95% CI 1.88–3.01), and 2.88 (95% CI 2.32–3.53), respectively. Aphasia was associated with elevated absolute risks of epilepsy, 2.38 (95% CI 1.81–2.98) in females vs. 3.11 (95% CI 2.57–3.66) in males per 1,000 person-years. Patients with aphasia had an

absolute excess risk of developing PSE after hemorrhagic stroke (3.52 per 1,000 person-years, 95% CI 2.68–4.52). The corresponding excess risk in persons with ischemic stroke was 2.64 per 1,000 person-years (95% CI 2.19–3.12). In the propensity score–matched analysis, the findings were similar (Table 2). The size of both the adjusted hazard ratio and adjusted rate difference declined with increasing time following the stroke (Table 3), but these remained elevated in all time intervals. (Table 3).

Discussion

In this nationwide cohort of 153,489 stroke patients, 15.3% experienced post-stroke aphasia (Table 1). After controlling for other risk factors, the presence of aphasia during the first three months after a stroke was associated with about a 65% increase in the incidence of epilepsy, corresponding to an absolute increase of about three per 1000 person-years. The association was present irrespective of age, sex, and type of stroke, and was particularly strong during the remainder of the first year following the stroke. The latter observation respectively is consistent with prior reports that early post-stroke changes may transiently increase seizure susceptibility[8,9,24] .

Post-stroke aphasia usually results from damage to the left perisylvian language cortex, predominantly supplied by the middle cerebral artery (MCA). Cortical involvement is a well-established risk factor for post-stroke epilepsy (PSE)[1,7,8,11,17,26–28], and lesions in the MCA

territory have been specifically associated with higher PSE risk[12,29,30]. The anatomical pattern may explain the higher PSE incidence in PSA patients.

Certain limitations of this study should be acknowledged. First, the patients in this cohort received care from multiple providers, and diagnostic information was obtained from the NHI database. To enhance the accuracy of aphasia identification, only the ICD-9-CM code 438.1 was used, which specifically denotes speech and language impairments resulting from cerebrovascular disease. PSE was defined by an ICD-9-CM (345.0–345.9) or ICD-10-CM (G40.0–G40.9) code recorded in at least two outpatient visits or one hospitalization, with each diagnosis confirmed by a board-certified neurologist. Ischemic and hemorrhagic stroke diagnoses were likewise verified by the treating neurologist for accuracy. Second, as our cohort included only patients who sought medical care, PSE may be underascertained in aphasic patients, potentially biasing results toward the null and underestimating the true incidence in PSA. Third, PSE incidence was assessed only after the first 90 days post-stroke, as patients with aphasia were enrolled within this period following receipt of speech therapy; thus, data on seizures occurring during the first 90 days were unavailable, which may represent the period of greatest risk.

Summary

PSA is an independent predictor of developing post-stroke epilepsy, conferring an excess risk across age, sex, and stroke type, with the largest excess in the early post-stroke period.

Conflict of Interest Statement

The authors have no conflicts of interest to disclose.

Acknowledgment

We are grateful to the Health Data Science Center, China Medical University Hospital, and Taoyuan General Hospital, Ministry of Health and Welfare for providing administrative, technical, and funding support.

Statement of Ethics

This study was approved by the Research Ethics Committee of China Medical University and Hospital, Taiwan (CMUH111-REC2-155(CR-2)). All personal identifiers used for data linkage were encrypted and replaced with surrogate numbers before release to ensure privacy protection. Written consent was therefore waived.

Data Availability Statement

Researchers must schedule appointments to perform data analyses at the Health Data Science Center designated by the Taiwan Ministry of Health and Welfare.

References

1. Pitkänen A, Roivainen R, Lukasiuk K. Development of epilepsy after ischaemic stroke. *Lancet Neurol*. 2016;15(2):185-197. doi:10.1016/S1474-4422(15)00248-3

2. Plowman E, Hentz B, Ellis C. Post-stroke aphasia prognosis: a review of patient-related and stroke-related factors. *J Eval Clin Pract*. 2012;18(3):689-694. doi:10.1111/j.1365-2753.2011.01650.x

3. Lin HL, Sung FC, Muo CH, Chen PC. Depression Risk in Post-Stroke Aphasia Patients: A Nationwide Population-Based Cohort Study. *Neuroepidemiology*. 2023;57(3):162-169. doi:10.1159/000530070

4. Grönberg A, Henriksson I, Stenman M, Lindgren AG. Incidence of Aphasia in Ischemic Stroke. *Neuroepidemiology*. 2022;56(3):174-182. doi:10.1159/000524206

5. Lamy C, Domigo V, Semah F, et al. Early and late seizures after cryptogenic ischemic stroke in young adults. *Neurology*. 2003;60(3):400-404. doi:10.1212/wnl.60.3.400

6. Bladin CF, Alexandrov AV, Bellavance A, et al. Seizures after stroke: a prospective multicenter study. *Arch Neurol*. 2000;57(11):1617-1622. doi:10.1001/archneur.57.11.1617

7. Phan J, Ramos M, Soares T, Parmar MS. Poststroke Seizure and Epilepsy: A Review of Incidence, Risk Factors, Diagnosis, Pathophysiology, and Pharmacological Therapies. *Oxid Med Cell Longev*. 2022;2022:7692215. doi:10.1155/2022/7692215

8. Ebbesen MQB, Dreier JW, Andersen G, Johnsen SP, Christensen J. Stroke and Risk of Epilepsy: A Danish Nationwide Register-Based Study. *Stroke*. 2024;55(4):972-982. doi:10.1161/STROKEAHA.123.044203

9. Do PT, Chen LY, Chan L, Hu CJ, Chien LN. Risk Factors for Postischemic Stroke Epilepsy in Young Adults: A Nationwide Population-Based Study in Taiwan. *Front Neurol*. 2022;13:880661. doi:10.3389/fneur.2022.880661

10. Hsing AW, Ioannidis JPA. Nationwide Population Science: Lessons From the Taiwan National Health Insurance Research Database. *JAMA Intern Med*. 2015;175(9):1527-1529. doi:10.1001/jamainternmed.2015.3540

11. Tanaka T, Ihara M, Fukuma K, et al. Pathophysiology, Diagnosis, Prognosis, and Prevention of Poststroke Epilepsy: Clinical and Research Implications. *Neurology*. 2024;102(11):e209450. doi:10.1212/WNL.0000000000209450

12. Tanaka T, Ihara M. Post-stroke epilepsy. *Neurochem Int*. 2017;107:219-228. doi:10.1016/j.neuint.2017.02.002

13. Sung SF, Hsieh CY, Kao Yang YH, et al. Developing a stroke severity index based on administrative data was feasible using data mining techniques. *J Clin Epidemiol*. 2015;68(11):1292-1300. doi:10.1016/j.jclinepi.2015.01.009

14. Johnson EL, Krauss GL, Lee AK, et al. Association Between Midlife Risk Factors and Late-Onset Epilepsy: Results From the Atherosclerosis Risk in Communities Study. *JAMA Neurol*. 2018;75(11):1375-1382. doi:10.1001/jamaneurol.2018.1935

15. Stefanidou M, Himali JJ, Devinsky O, et al. Vascular risk factors as predictors of epilepsy in older age: The Framingham Heart Study. *Epilepsia*. 2022;63(1):237-243. doi:10.1111/epi.17108

16. Galovic M, Ferreira-Atuesta C, Abraira L, et al. Seizures and Epilepsy After Stroke: Epidemiology, Biomarkers and Management. *Drugs Aging*. 2021;38(4):285-299. doi:10.1007/s40266-021-00837-7

17. Ferreira-Atuesta C, Döhler N, Erdélyi-Canavese B, et al. Seizures after Ischemic Stroke: A Matched Multicenter Study. *Ann Neurol*. 2021;90(5):808-820. doi:10.1002/ana.26212

18. Naylor J, Churilov L, Johnstone B, et al. The Association Between Atrial Fibrillation and Poststroke Seizures is Influenced by Ethnicity and Environmental Factors. *J Stroke Cerebrovasc Dis Off J Natl Stroke Assoc*. 2018;27(10):2755-2760. doi:10.1016/j.jstrokecerebrovasdis.2018.05.044

19. Brauchitsch S von, Strzelczyk A, Rosenow F, et al. High end-of-life incidence of seizures and status epilepticus in patients with primary and secondary brain tumors. *J Neurooncol*. 2022;160(2):277-284. doi:10.1007/s11060-022-04133-1

20. De Reuck J, Proot P, Van Maele G. Chronic obstructive pulmonary disease as a risk factor for stroke-related seizures. *Eur J Neurol*. 2007;14(9):989-992. doi:10.1111/j.1468-1331.2007.01829.x

21. Austin PC. A comparison of 12 algorithms for matching on the propensity score. *Stat Med*. 2014;33(6):1057-1069. doi:10.1002/sim.6004

22. Lekoubou A, Nguyen C, Kwon M, Nyalundja AD, Agrawal A. Post-stroke Everything. *Curr Neurol Neurosci Rep*. 2023;23(11):785-800. doi:10.1007/s11910-023-01308-9

23. Austin PC. The performance of different propensity score methods for estimating

marginal hazard ratios. *Stat Med*. 2013;32(16):2837-2849. doi:10.1002/sim.5705

24. Graham NSN, Crichton S, Koutroumanidis M, Wolfe CDA, Rudd AG. Incidence and associations of poststroke epilepsy: the prospective South London Stroke Register. *Stroke*. 2013;44(3):605-611. doi:10.1161/STROKEAHA.111.000220

25. Yu Y, Tuo M, Chen Z, Zhang J, Yang Y, Wang Y. Association between severity of stroke and poststroke epilepsy among patients with ischemic stroke: Mediating role of infection and early seizure. *Epilepsia*. 2025;66(6):2011-2021. doi:10.1111/epi.18352

26. Verburgt E, Fellah L, Ekker MS, et al. Risk of Poststroke Epilepsy Among Young Adults With Ischemic Stroke or Intracerebral Hemorrhage. *JAMA Neurol*. 2025;82(6):597-604. doi:10.1001/jamaneurol.2025.0465

27. Galovic M, Döhler N, Erdélyi-Canavese B, et al. Prediction of late seizures after ischaemic stroke with a novel prognostic model (the SeLECT score): a multivariable prediction model development and validation study. *Lancet Neurol*. 2018;17(2):143-152. doi:10.1016/S1474-4422(17)30404-0

28. Haapaniemi E, Strbian D, Rossi C, et al. The CAVE score for predicting late seizures after intracerebral hemorrhage. *Stroke*. 2014;45(7):1971-1976. doi:10.1161/STROKEAHA.114.004686

29. Hsieh CY, Su CC, Lai ECC, et al. Risk of Incident Epilepsy After a Middle Cerebral Artery Territory Infarction. *Front Neurol*. 2022;13:765969. doi:10.3389/fneur.2022.765969

30. Leung W, Wong YK, Ho WH, et al. H-SeLECT: A Novel MRI-based Risk Score for the Prediction of Post-stroke Epilepsy (S41.009). *Neurology*. 2024;102(7_supplement_1):5883. doi:10.1212/WNL.0000000000206093

Table 1. Demographic and clinical characteristics of stroke patients with and without aphasia

	Entire cohort of patients with stroke					Propensity score-matched groups				
	Aphasia N = 23431		Non-aphasia N = 130058			Aphasia N = 22913		Non-aphasia N = 22913		
Variable	n	%	n	%	SMD	n	%	n	%	SMD
Sex										
Male	14268	60.9	76192	58.6	0.047	13916	60.7	13783	60.2	0.012
Female	9163	39.1	53866	41.4	0.047	8997	39.3	9130	39.9	0.012
Age, years										
18-44	1453	6.20	6894	5.30	0.039	1390	6.07	1720	7.51	0.057
45-64	8231	35.1	41645	32.0	0.066	8016	35.0	7728	33.7	0.026
≥ 65	13747	58.7	81519	62.7	0.082	13507	59.0	13465	58.8	0.004
Mean (SD)	66.0	(13.1)	67.4	(13.2)	0.104	66.1	(13.0)	66.4	(14.4)	0.019
Stroke type										
Hemorrhagic stroke	7109	30.3	26909	20.7	0.223	6936	30.3	7022	30.7	0.008
Ischemic stroke	16322	69.7	103149	79.3	0.223	15977	69.7	15891	69.4	0.008
Length of hospital stay										
1–7 days	5338	22.8	70960	54.6	0.690	5337	23.3	6168	26.9	0.084
8-14	5363	22.9	33983	26.1	0.075	5359	23.4	5196	22.7	0.017
15-21	3532	15.1	11106	8.54	0.204	3514	15.3	3480	15.2	0.004
22-28	2789	11.9	5805	4.46	0.274	2737	12.0	2818	12.3	0.011
>28	6409	27.4	8204	6.31	0.586	5966	26.0	5251	22.9	0.073
Median (Q1, Q3)	16	(8, 29)	7	(4, 12)		16	(8, 29)	15	(7, 27)	
Median OPD visit (Q1,Q3)[#]	37	(21, 59)	41	(23, 67)		37	(21, 60)	37	(20, 60)	

Median IPD visit (Q1,Q3)[#]	2	(1, 3)	1	(1, 2)		2	(1, 3)	2	(1, 3)	
Comorbidity[‡]										
Diabetes	9278	39.6	51548	39.6	0.001	9096	39.7	9218	40.2	0.011
Hypertension	20679	88.3	110411	84.9	0.099	20203	88.2	20203	88.2	0.000
Hyperlipidemia	7366	31.4	44012	33.8	0.051	7193	31.4	7126	31.1	0.006
Peripheral Arterial Disease	1841	7.86	11522	8.86	0.036	1800	7.86	1978	8.63	0.028
Coronary Heart disease	7862	33.6	46025	35.4	0.039	7699	33.6	8129	35.5	0.039
Atrial fibrillation	3268	14.0	10585	8.14	0.186	3121	13.6	3151	13.8	0.004
Malignancy	983	4.20	5506	4.23	0.002	963	4.20	991	4.33	0.006
Chronic obstructive pulmonary disease	6155	26.3	36871	28.4	0.047	6072	26.5	6330	27.6	0.025
Utilization of intensive care during hospitalization for stroke										
Intensive care unit	10432	44.5	29289	22.5	0.479	10079	44.0	10396	45.4	0.028
Intubation	2964	12.7	7507	5.77	0.240	2933	12.8	3158	13.8	0.029
Airway suctioning	8948	38.2	16440	12.6	0.614	8526	37.2	8567	37.4	0.004
Nasogastric tube (NG tube)	13590	58.0	25509	19.6	0.857	13076	57.1	13332	58.2	0.023
Foley catheter	12502	53.4	31590	24.3	0.625	12030	52.5	12409	54.2	0.033
Follow–up years										
Mean (SD)	8.07	(5.52)	8.67	(5.65)	0.109	8.08	(5.52)	7.36	(5.78)	0.127
Median (Q1-Q3)	7.56	(2.83, 13.4)	8.67	(3.24, 14.0)		7.58	(2.86, 13.4)	6.14	(1.80, 13.3)	

Chi-square, †t-test, and ‡Wilcoxon rank test, SMD, standardized mean difference (≥0.1 indicates imbalance); SD, standard deviation; Q1, lower quartile; Q3, upper quartile, IPD/OPD, inpatient/outpatient department

＃Measured within two years prior to the stroke hospitalization

‡The presence of a diagnostic code in at least 2 outpatient claims or 1 inpatient claim within two years prior to the stroke hospitalization.

Table 2. Epilepsy risk in stroke patients with aphasia compared with stroke patients without aphasia stratified by age, sex and stroke type

Variable	Entire cohort of patients with stroke						Propensity score-matched groups					
	Aphasia		Non- aphasia		aRD(95%CI)‖	aHR(95% CI)‡‡	Aphasia		Non- aphasia		RD (95% CI)	aHR(95% CI)§
	Event no.	Rate†	Event no.	Rate†			Event no.	Rate†	Event no.	Rate†		
Overall	2460	13.02	5140	4.56	2.96 (2.60–3.37)	1.65 (1.57-1.74)	2354	12.72	1292	7.66	5.06 (4.40-5.72)	1.69 (1.58-1.81)
Sex												
Female	828	11.19	1920	4.03	2.38 (1.81–2.98)	1.59 (1.45-1.74)	799	10.99	436	6.66	4.32 (3.34-5.31)	1.68 (1.50-1.89)
Male	1632	14.19	3220	4.94	3.11 (2.57–3.66))	1.63 (1.52-1.74)	1555	13.84	856	8.29	5.55 (4.67-6.43)	1.69 (1.55-1.84)
Age												
< 45	289	18.21	530	6.07	3.52 (2.06–5.22)	1.58 (1.34-1.86)	273	17.95	198	9.97	7.98 (5.44-10.53)	1.77 (1.47-2.12)
45-64	1085	12.88	1963	4.18	2.42 (1.88–3.01)	1.58 (1.45-1.72)	1030	12.50	537	6.96	5.54 (4.58-6.51))	1.80 (0.63-2.00)
≥ 65	1086	12.22	2647	4.64	2.88 (2.32–3.53)	1.62 (1.50-1.76)	1051	12.01	557	7.77	4.24 (3.27-5.21)	1.62 (1.46-1.80)
Stroke type												
Hemorrhagic stroke	1106	17.68	1832	7.66	3.52 (2.68–4.52)	1.46 (1.35-1.59)	1059	17.35	631	11.45	5.90 (4.53-7.28)	1.57 (1.42-1.73)
Ischemic stroke	1354	10.71	3308	3.72	2.64 (2.19–3.12)	1.71 (1.59-1.84)	1295	10.44	661	5.82	4.62 (3.90-5.34)	1.81 (1.65-1.99)

Note: †per 1000 person-years; aRD(95% CI):adjusted Rate Difference (95% confidence interval); aHR (95% CI): adjusted hazard ratio (95% confidence interval).

‡Models adjusted for age, sex, stroke type, previous hospitalization, previous number of outpatient clinic visits, diabetes mellitus, hyperlipidemia, peripheral arterial disease, coronary heart disease, chronic obstructive pulmonary disease, length of hospital stay and use of intensive care during the index hospitalization for stroke (stay in the intensive care unit, intubation, airway suctioning, nasogastric tube, Foley catheter).

§Hazard ratios were estimated using Cox proportional hazards models in propensity score–matched pairs

‖ Adjusted RD (per 1,000 person-years) for both the overall and propensity score–matched analyses was derived from the corresponding aHR as $r_0 \times (aHR - 1)$, where r_0 is the non-aphasia incidence rate in the same stratum; 95% CIs were calculated using the lower and upper limits of the aHR. The aHRs were obtained from the models described in ‡.

Table 3. Epilepsy risk in stroke patients with aphasia compared with stroke patients without aphasia stratified by follow-up year

Variable	The entire cohort of patients with stroke						Propensity score-matched groups					
	Aphasia		Non- aphasia		aRD(95%CI)[‖]	aHR (95% CI)[‡]	Aphasia		Non- aphasia		RD (95% CI)	aHR (95% CI)[§]
	Event no.	Rate[†]	Event no.	Rate[†]			Event no.	Rate[†]	Event no.	Rate[†]		
<= 1 years	909	41.40	1315	10.73	8.70 (6.87–10.52)	1.81 (1.64-1.98)	867	40.36	456	22.09	18.26 (14.90-21.63)	1.83 (1.64-2.05)
1-2 years	410	20.78	763	6.83	4.71 (3.28–6.41)	1.69 (1.48-1.94)	390	20.18	214	12.02	8.16 (5.59-10.73)	1.68 (1.42-1.99)
2-3 years	231	12.78	467	4.52	3.16 (1.90–4.70)	1.70 (1.42-2.04)	224	12.65	113	7.08	5.57 (3.46-7.68)	1.79 (1.43-2.24)
> 3 years	910	7.04	2595	3.28	1.48 (1.08–1.91)	1.45 (1.33-1.58)	873	6.90	509	4.45	2.44 (1.84-3.04)	1.54 (1.38-1.72)

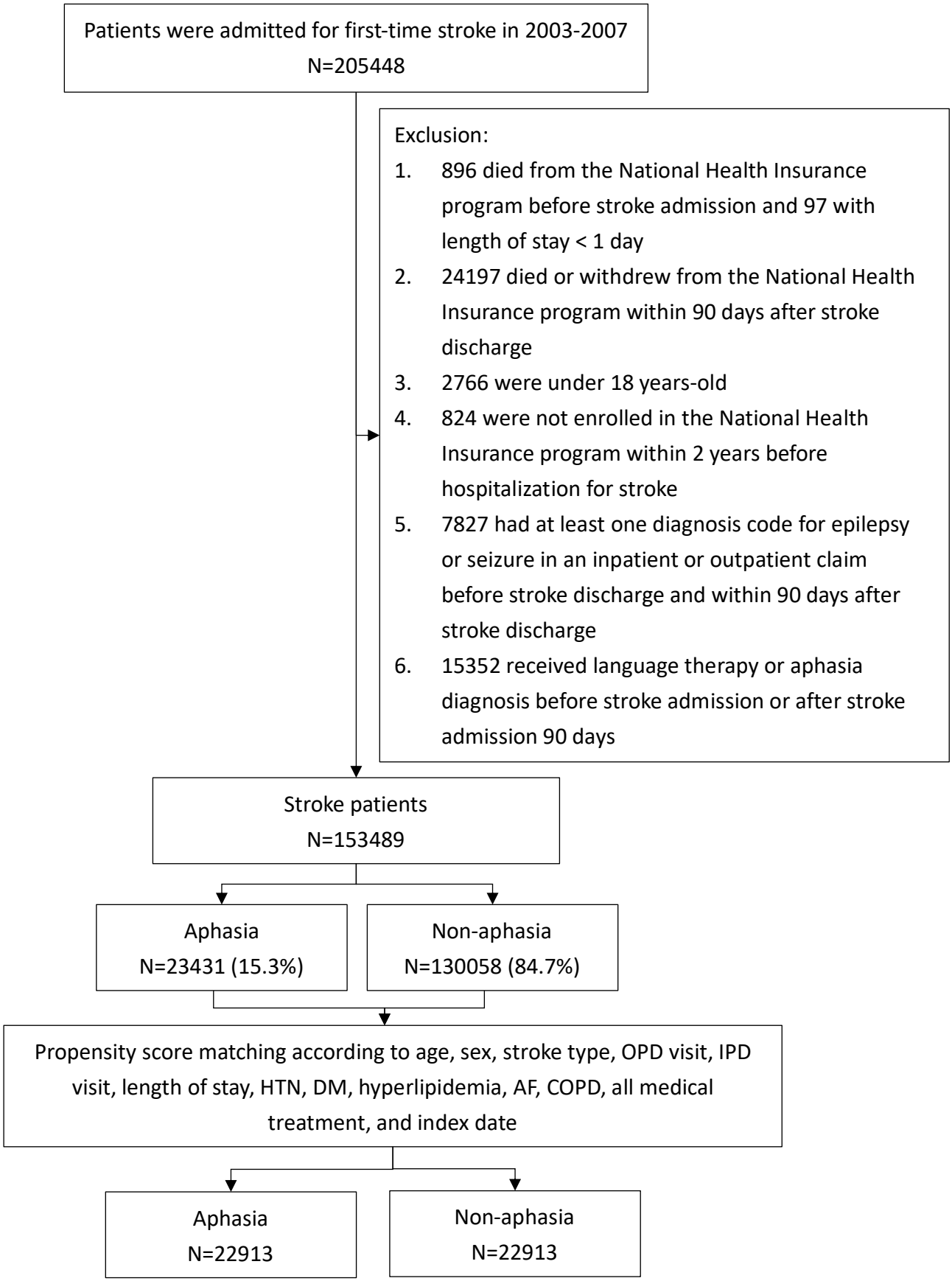
Note: [†]per 1000 person-years; aRD(95% CI): adjusted Rate Difference (95% confidence interval); aHR (95% CI): adjusted hazard ratio (95% confidence interval).

[‡]Models adjusted for age, sex, stroke type, previous hospitalization, previous number of outpatient clinic visits, diabetes mellitus, hyperlipidemia, peripheral arterial disease, coronary heart disease, chronic obstructive pulmonary disease, length of hospital stay and use of intensive care during the index hospitalization for stroke (stay in the intensive care unit, intubation, airway suctioning, nasogastric tube, Foley catheter).

[§]Hazard ratios were estimated using Cox proportional hazards models in propensity score–matched pairs

[‖] Adjusted RD (per 1,000 person-years) for both the overall and propensity score–matched analyses was derived from the corresponding aHR as $r_0 \times (aHR - 1)$, where r_0 is the non-aphasia incidence rate in the same stratum; 95% CIs were calculated using the lower and upper limits of the aHR. The aHRs were obtained from the models described in ‡.

Figure 1. Flow chart for study subjects



Supplement Table 1 Diagnosis codes

Name	ICD–9–CM code
Ischemic stroke	433.x, 434.x, or 436.x
Hemorrhagic stroke.	430.x–432.x
Aphasia	438.1
Epilepsy	345.0–345.9
Diabetes Mellitus	250
Hypertension	401–405
Hyperlipidemia	272
Peripheral Arterial Disease	440-449
Coronary heart disease	410-414
Atrial fibrillation	427.3
Malignancy	140-208
Chronic obstructive pulmonary disease	490-496, 500-505, 506.4

Abbreviations: ICD–9–CM, International Classification of Diseases, Ninth Revision, Clinical Modification

*The diagnosis of depression was identified using both ICD–9–CM and ICD–10–CM codes because, in Taiwan, the ICD–10–CM has been adopted since 2016. All other diseases were identified using ICD–9–CM only because they were baseline comorbidities in our analysis and were extracted before 2006 from the dataset.