Positive Predictive Value of ICD-10 Codes for Cerebral Venous Sinus Thrombosis in Taiwan’s National Health Insurance Claims Database

Shu-Chen Liao¹,²
Shih-Chieh Shao³,⁴
Edward Chia-Cheng Lai⁴
Swu-Jane Lin⁵
Wei-I Huang⁶
Cheng-Yang Hsieh⁴,⁷

¹Department of Emergency Medicine, Keelung Chang Gung Memorial Hospital, Keelung, Taiwan; ²Chang Gung University College of Medicine, Taoyuan, Taiwan; ³Department of Pharmacy, Keelung Chang Gung Memorial Hospital, Keelung, Taiwan; ⁴School of Pharmacy, Institute of Clinical Pharmacy and Pharmaceutical Sciences, College of Medicine, National Cheng Kung University, Tainan, Taiwan; ⁵Department of Pharmacy Systems, Outcomes and Policy, College of Pharmacy, University of Illinois at Chicago, Chicago, IL, USA; ⁶Taiwan Drug Relief Foundation, Taipei, Taiwan; ⁷Department of Neurology, Tainan Sin Lau Hospital, Tainan, Taiwan

Objective: This study aims to determine the positive predictive value (PPV) of case definitions for cerebral venous sinus thrombosis (CVST) in Taiwan’s National Health Insurance claims database based on the International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM) diagnostic codes.

Study Design and Setting: Inpatient records with ICD-10-CM codes of G08, I629, I636, or I676 were retrieved from the claims data of all hospital branches of Chang Gung Medical Foundation. Manual review of the medical records and images was performed in order to ascertain the diagnosis. The PPV of various case definitions for CVST was estimated.

Results: Of the 380 hospitalizations, 166 and 214 were determined to be true-positive and false-positive episodes of acute CVST, respectively. The PPV of the ICD-10-CM codes of G08, I629, I636, and I676 was 88.2%, 2.0%, 100.0%, and 91.3%, respectively. The PPV generally increased when acute CVST was defined as a primary diagnosis or as ICD-10-CM codes plus anticoagulant use. Miscoding in other conditions, tentative diagnosis, and remote episode of CVST were determined as the main reasons for false-positive diagnosis of acute CVST.

Conclusion: This study determined the PPV of ICD-10-CM codes for identifying CVST, which may offer a reference for future claims-based research.

Keywords: hospitalization claims data, positive predictive value, diagnostic codes, ICD-10-CM, cerebral venous sinus thrombosis, Taiwan

Introduction

Massive administration of SARS-CoV2 vaccines has been identified as the mainstay in the control of the unprecedented coronavirus disease 2019 (Covid-19) pandemic. Cerebral venous sinus thrombosis (CVST) is a rare manifestation of cerebrovascular disease, accounting for 0.5%–1.0% of all stroke admissions.¹ However, in April 2021, three independent reports from Europe described a total of 39 cases of CVST after administration of ChAdOx1 nCoV-19 (AstraZeneca) vaccine containing adenovirus vectors.²³ Moreover, CVST events were also reported in the United States following vaccination with Ad26.COV2.S (Janssen), another adenovirus vector vaccine.⁴ An excess of about 1 CVST event per 40,000 vaccinations has been reported.⁵ On the contrary, no increased risks of thromboembolic events were recorded following administration of mRNA SARS-CoV2 vaccines such as BNT162b2 (Pfizer–BioNTech).⁶ Exploration of these results is needed in the Asian population, whose epidemiologic features of CVST at background are different from those of the Western population.⁷
Administrative claims data, derived from electronic data reported by healthcare providers and insurers for the purposes of reimbursement, provide an affordable and efficient way to monitor vaccine safety. For example, Taiwan’s National Health Insurance Research Database (NHIRD) has been utilized for research on influenza vaccination. Recently, researchers of Norway and Denmark also conducted a population-based cohort study on the risk of CVST after administration of ChAdOx1-S (AstraZeneca) vaccine. However, the diagnostic code they chose to identify cases of CVST might not be representative enough. This could lead to an underestimate of actual risk.

At the initiation of the nationwide SARS-CoV2 vaccination program, Taiwan Drug Relief Foundation, which operates the Taiwan National Adverse Drug Reaction Reporting Center, began to monitor the differences between the observed and the expected number of CVST for signal detection. Nevertheless, the diagnostic codes of CVST in Taiwan’s NHIRD have not been validated. It is unknown whether the degree of positive predictive value (PPV) of these diagnostic codes is sufficient to accurately estimate the excess risk of CVST. Hence, we carried out the current study to validate various case definitions of CVST based on the International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM) codes, using the National Health Insurance claims data from a large healthcare system in Taiwan.

**Materials and Methods**

**Study Settings**

This study was conducted in all hospital branches (Keelung, Linkou, Taipei, Taoyuan, Kaohsiung, Chiayi, Yunlin, and Fengshian) of the Chang Gung Medical Foundation (CGMF) in Taiwan. CGMF is the largest multi-institutional healthcare system in Taiwan, with more than 10% of inpatients hospitalized during this period. We chose I629 to identify possible cases of CVST because some of them might present with intracerebral hemorrhage, according to our prior study. If a patient had multiple hospitalization episodes with aforementioned discharge diagnoses during the study period, only the first hospitalization was retrieved for analysis.

**Ascertainment of Acute CVST**

We identified a CVST event according to the definition of the Taiwan Stroke Registry program, which is “an infarct or/and intracerebral hemorrhage on computed tomography, and an evidence of a cerebral sinus or venous occlusion on magnetic resonance image and magnetic resonance angiography or on conventional angiography”. Figure 1 shows the process of case ascertainment. In brief, physician 1, who is a specialist in emergency medicine with access to the CGMF electronic medical records, extracted the discharge notes and relevant image reports of each patient, and made the initial ascertainment of CVST. Thereafter, physician 2, a neurology specialist who was not involved in extracting data, also reviewed the extracted data, and made the second ascertainment of CVST independently. In case of disagreement, discussions were held between physicians 1 and 2 until consensus was reached for final ascertainment. If the two reviewers still failed to come to an agreement after discussion, then a third reviewer will be sought to adjudicate the cases.

For hospitalizations determined not to be true episodes of CVST, the reasons were documented and categorized into the following groups: 1) a tentative diagnosis of CVST, which was later excluded after clinical evaluation and imaging studies; 2) remote history of CVST; and 3) other diagnoses miscoded as CVST (e.g., congenital hypoplasia of the transverse sinus).
This study examined various case definitions for identifying cases of CVST by including individual or combination of ICD-10-CM codes in the 1) primary diagnosis; 2) primary or secondary diagnosis; 3) primary, secondary, or tertiary diagnosis; and 4) any field of the diagnosis. We further tested whether adding in-hospital prescriptions of any anticoagulants (i.e., heparin, low-molecular-weight heparin, warfarin, and direct oral anticoagulants) can improve the identification of CVST cases. The positive predictive value (PPV) was computed as the proportion of hospitalizations with confirmed CVST among all possible CVST hospitalizations, and the 95% confidence interval (CI) of PPV was estimated using the Clopper-Pearson exact method. Data analyses were performed using SAS Enterprise Guide 7.13 (SAS Institute, Inc., Cary, NC, USA).

**Results**

Between January 2017 and December 2020, we were able to identify 68, 201,15, and 103 hospitalizations from the hospitalization claims data with an ICD-10-CM code of G08, I629, I636, and I676 in any field of the discharge diagnoses, respectively. Seven hospitalizations were coded with more than one of these ICD-10-CM codes. Therefore, a total of 380 patients were reviewed. Physicians 1 and 2 reached agreement in 365 patients and were able to come to an agreement after discussion on cases that disagreed during the first-round of assessment (Figure 1). No third reviewer was needed for adjudication of disputed cases.

Table 1 shows the number of true-positive CVST hospitalizations and PPV under each diagnostic code. Using only the primary diagnosis, the PPV was >91% for G08, I636, and I676, but was extremely low for I629 (<2%). However, when expanding the case definition to the secondary, tertiary, or any field of diagnosis, the PPV of each ICD-10-CM code decreased, while the number of identified CVST cases increased.

For false-positive cases with ICD-10-CM code of G08 (n = 8), five were tentative diagnosis, one was remote CVST, and two were miscoding. For false-positive cases with I676 (n = 9), six were tentative diagnosis, while three were miscoding. All false-positive cases with I629 (n = 197) were miscoding.

Table 2 shows the PPV for combination use of ICD-10-CM codes. PPV for G08, I636, or I676 in any field of diagnosis was 90.7% (95% CI, 85.5%–94.5%). However, the PPV reduced by half when I629 was added to the case definitions. Adding in-hospital exposure of any anticoagulants improved the PPV to almost 100%, including case definitions using only I629 (Table 3) or I629 in combination with G08, I636, or I676 (Table 4). However, such improvements were all at the expense of a significant decrease in CVST case numbers.

**Discussion**

We assessed the PPV of various case definitions using ICD-10-CM diagnostic codes for CVST. The PPVs of G08, I636, and I676 were found to be good to excellent when they were used alone or in combination with other relevant codes in any field of diagnosis, while the PPV of I629 was poor. Restricting ICD-10-CM codes to the primary diagnosis or adding anticoagulants in case definitions improved the PPV to almost 100%, but at the expense of much fewer cases of identified CVST.

---

**Figure 1** Process of case ascertainment.

**Abbreviations:** CGMF, Chang Gung Medical Foundation; ICD-10-CM, International Classification of Diseases, 10th Revision, Clinical Modification.
In contrast to acute ischemic stroke\textsuperscript{15} and hemorrhagic stroke\textsuperscript{13}, there were few prior studies on the PPV of ICD-10-CM codes for CVST. In a recent validation study conducted in the UK, most of the ascertained CVST cases (n = 60) were coded with G08 (90%), rather than 1636 (1.7%) or 1676 (1.7%).\textsuperscript{10} In contrast, our ascertained CVST cases (n = 166) were more likely be coded with I676 (54.2), followed by G08 (41.0%) and I636 (9.0%) (The sum of percentages was greater than 100% due to coding with more than one ICD-10-CM code in 7 patients). Such a difference could be due to the different coding behaviors among different healthcare systems and databases. Additionally, Asian CVST patients were reported to have a higher frequency of underlying hypercoagulable state.\textsuperscript{7} Such an etiological difference might also explain why

| Case Definitions (ICD-10-CM Codes) | TP, n | FP, n | PPV, % (95% CI) |
|------------------------------------|-------|-------|----------------|
| G08: Intracranial and intraspinal phlebitis and thrombophlebitis | | | |
| G08 as primary diagnosis | 11 | 1 | 91.7 (61.5–99.8) |
| G08 as primary or secondary diagnosis | 33 | 4 | 89.2 (75.6–97.0) |
| G08 as primary, secondary, or tertiary diagnosis | 43 | 5 | 89.6 (77.3–96.5) |
| G08 in any field of diagnosis | 60 | 8 | 88.2 (78.1–94.8) |
| I629: Nontraumatic intracranial hemorrhage, unspecified | | | |
| I629 as primary diagnosis | 1 | 74 | 1.3 (0.0–7.2) |
| I629 as primary or secondary diagnosis | 4 | 125 | 3.1 (0.9–7.8) |
| I629 as primary, secondary, or tertiary diagnosis | 4 | 155 | 2.5 (0.7–6.3) |
| I629 in any field of diagnosis | 4 | 197 | 2.0 (0.5–5.0) |
| I636: Cerebral infarction due to cerebral venous thrombosis, non-pyogenic | | | |
| I636 as primary diagnosis | 11 | 0 | 100.0 (71.5–100.0) |
| I636 as primary or secondary diagnosis | 14 | 0 | 100.0 (76.8–100.0) |
| I636 as primary, secondary, or tertiary diagnosis | 14 | 0 | 100.0 (76.8–100.0) |
| I636 in any field of diagnosis | 15 | 0 | 100.0 (78.2–100.0) |
| I676: Non-pyogenic thrombosis of intracranial venous system | | | |
| I676 as primary diagnosis | 48 | 3 | 94.1 (83.8–98.8) |
| I676 as primary or secondary diagnosis | 70 | 7 | 90.9 (82.2–96.3) |
| I676 as primary, secondary, or tertiary diagnosis | 84 | 8 | 91.3 (83.6–96.2) |
| I676 in any field of diagnosis | 94 | 9 | 91.3 (84.1–95.9) |

**Table 2** Positive Predictive Value of Combined Case Definitions for Identifying Cases of Cerebral Venous Sinus Thrombosis

| Case Definitions (ICD-10-CM Codes) | TP, n | FP, n | PPV, % (95% CI) |
|------------------------------------|-------|-------|----------------|
| G08, I636, I676 | | | |
| G08 or I636 or I676 as primary diagnosis | 70 | 4 | 94.6 (86.7–98.5) |
| G08 or I636 or I676 as primary or secondary diagnosis | 115 | 11 | 91.3 (84.9–95.6) |
| G08 or I636 or I676 as primary, secondary, or tertiary diagnosis | 137 | 13 | 91.3 (85.6–95.3) |
| G08 or I676 or I676 in any field of diagnosis | 165 | 17 | 90.7 (85.5–94.5) |

**Note:** Seven cases have simultaneous coding with G08, I636, I676, or I629 in any field of diagnosis.

**Abbreviations:** TP, true positive; FP, false positive; ICD-10-CM, International Classification of Diseases, 10th Revision, Clinical Modification; PPV, positive predictive value; CI, confidence interval.
more of our CVST cases were coded with I676 (non-pyogenic thrombosis of intracranial venous system), rather than G08 (intracranial and intraspinal phlebitis and thrombophlebitis).

Based on the study results, we have the following suggestions when using Taiwan’s NHIRD to conduct research on CVST. First, if the researchers aim to create a cohort of patients with CVST, we recommend the case definitions of “G08 or I636 or I676 as primary diagnosis,” with or without anticoagulant use. These definitions will produce the highest PPVs (94.6% to 100%), ensuring that patients in this cohort are true cases of CVST. The trade-off is that such definitions might lead to a loss of true cases of CVST whose ICD-10-CM

Table 3 Positive Predictive Value of Individual Case Definition Plus Anticoagulant Use for Identifying Cases of Cerebral Venous Sinus Thrombosis

| Case Definitions (ICD-10-CM Codes) | TP, n | FP, n | PPV, % (95% CI) |
|------------------------------------|-------|-------|------------------|
| G08: Intracranial and intraspinal phlebitis and thrombophlebitis | | | |
| G08 as primary diagnosis | 4 | 0 | 100.0 (40.0–100.0) |
| G08 as primary or secondary diagnosis | 13 | 1 | 92.9 (66.1–99.8) |
| G08 as primary, secondary, or tertiary diagnosis | 18 | 1 | 94.7 (74.0–99.9) |
| G08 in any field of diagnosis | 24 | 1 | 96.0 (79.7–99.9) |
| I629: Nontraumatic intracranial hemorrhage, unspecified | | | |
| I629 as primary diagnosis | 1 | 0 | 100.0 (2.5–100.0) |
| I629 as primary or secondary diagnosis | 3 | 0 | 100.0 (29.2–100.0) |
| I629 as primary, secondary, or tertiary diagnosis | 3 | 0 | 100.0 (29.2–100.0) |
| I629 in any field of diagnosis | 3 | 0 | 100.0 (29.2–100.0) |
| I636: Cerebral infarction due to cerebral venous thrombosis, non-pyogenic | | | |
| I636 as primary diagnosis | 10 | 0 | 100.0 (69.2–100.0) |
| I636 as primary or secondary diagnosis | 12 | 0 | 100.0 (73.5–100.0) |
| I636 as primary, secondary, or tertiary diagnosis | 12 | 0 | 100.0 (73.5–100.0) |
| I636 in any field of diagnosis | 13 | 0 | 100.0 (75.3–100.0) |
| I676: Non-pyogenic thrombosis of intracranial venous system | | | |
| I676 as primary diagnosis | 39 | 0 | 100.0 (91.0–100.0) |
| I676 as primary or secondary diagnosis | 54 | 0 | 100.0 (93.4–100.0) |
| I676 as primary, secondary, or tertiary diagnosis | 60 | 0 | 100.0 (94.0–100.0) |
| I676 in any field of diagnosis | 64 | 0 | 100.0 (94.4–100.0) |

Abbreviations: TP, true positive; FP, false positive; ICD-10-CM, International Classification of Diseases, 10th Revision, Clinical Modification; PPV, positive predictive value; CI, confidence interval.

Table 4 Positive Predictive Value of Combined Case Definitions Plus Anticoagulant Use for Identifying Cases of Cerebral Venous Sinus Thrombosis

| Case Definitions (ICD-10-CM Codes) | TP, n | FP, n | PPV, % (95% CI) |
|------------------------------------|-------|-------|------------------|
| G08 or I636 or I676 | 53 | 0 | 100.0 (93.3–100.0) |
| G08 or I636 or I676 as primary or secondary diagnosis | 78 | 1 | 98.7 (93.2–100.0) |
| G08 or I636 or I676 as primary, secondary, or tertiary diagnosis | 88 | 1 | 98.9 (93.9–100.0) |
| G08 or I636 or I676 in any field of diagnosis | 99 | 1 | 99.0 (94.6–100.0) |
| G08, I636, I676, I629 | | | |
| G08 or I636 or I676 or I629 as primary diagnosis | 54 | 0 | 100.0 (93.4–100.0) |
| G08 or I636 or I676 or I629 as primary or secondary diagnosis | 79 | 1 | 98.8 (93.2–100.0) |
| G08 or I636 or I676 or I629 as primary, secondary, or tertiary diagnosis | 88 | 1 | 98.9 (93.9–100.0) |
| G08 or I636 or I676 or I629 in any field of diagnosis | 99 | 1 | 99.0 (94.6–100.0) |

Abbreviations: TP, true positive; FP, false positive; ICD-10-CM, International Classification of Diseases, 10th Revision, Clinical Modification; PPV, positive predictive value; CI, confidence interval.
codes are in the secondary or other fields of discharge diagnosis. The trade-off might be tolerable given the large sample size of Taiwan’s NHIRD. Second, if CVST itself is the main research outcome (e.g., the researchers aim to determine the incidence rate of CVST after vaccination), we recommend the case definition of “G08 or I636 or I676 in any field of diagnosis.” Although this definition is associated with a little lower PPV, it will help researchers to identify more true cases of CVST and to avoid misclassification bias.

Our results demonstrated the importance of a validation study on the criteria for case selection/definition before using administrative claims data to conduct research on vaccine safety. For example, a CVST case was defined as using administrative claims data to conduct research on study on the criteria for case selection/definition before codes for CVST. We did not know how many true specificity of those case definitions using ICD-10-CM negative rate, negative predictive value, sensitivity, and it would fail to identify 98.3% of true CVST cases and lead to a significant underestimate of risks.

Our study has several limitations. First, due to the design of the study, we could not determine the false-negative rate, negative predictive value, sensitivity, and specificity of those case definitions using ICD-10-CM codes for CVST. We did not know how many true CVST patients were miscoded as I61 (nontraumatic intracerebral hemorrhage) or I63 (cerebral infarction). Further studies using data linkage with stroke registry and claims data are needed to achieve this purpose. Second, we did not investigate the PPV of other similar codes such as O225 (cerebral venous thrombosis in pregnancy) and O873 (cerebral venous thrombosis in the puerperium). Currently, almost all pregnant women in Taiwan received mRNA-1273 SARS-CoV-2 vaccine (Moderna), which was considered to have lower risk of CVST. Third, 112 (56.9%) of the 197 patients with an ICD-10-CM diagnostic code I629 only received a plain computed tomography as their key image study modality. The diagnosis of CVST might be missed among these patients. Fourth, the CGMF healthcare system covers only a small fraction (10.2%) of the national inpatient population. However, it was a large healthcare system with academic medical centers and regional and district hospitals, and the results might be representative of the inpatient population in Taiwan.

**Conclusion**

The PPVs of ICD-10-CM codes of G08, I636, and I676 were generally high to identify cases of CVST, while that of I629 was not. Our results support the use of Taiwan’s NHIRD to study CVST.

**Funding**

This work was supported by the Tainan Sin Lau Hospital (grant number SLH109-09).

**Disclosure**

The authors report no conflicts of interest in this work.

**References**

1. Olivi L, Squitieri M, Cohen H, Cowley P, Werring DJ. Cerebral venous thrombosis: a practical guide. Pract Neurol. 2020;20(5):356–367. doi:10.1136/practneurol-2019-002415
2. Schultz NH, Sørvoll IH, Michelsen AE, et al. Thrombosis and thrombocytopenia after ChAdOx1 nCoV-19 vaccination. N Engl J Med. 2021;384(22):2124–2130. doi:10.1056/NEJMoa2104882
3. Scully M, Singh D, Lown B, et al. Pathologic antibodies to platelet factor 4 after ChAdOx1 nCoV-19 vaccination. N Engl J Med. 2021;384(23):2202–2211. doi:10.1056/NEJMoa2105385
4. See I, Su JR, Lale A, et al. US case reports of cerebral venous sinus thrombosis with thrombocytopenia after Ad26.COV2.S vaccination, March 2 to April 21, 2021. JAMA. 2021;325(24):2448–2456. doi:10.1001/jama.2021.7517
5. Pottegård A, Lund LC, Karlstad Ø, et al. Arterial events, venous thromboembolism, thrombocytopenia, and bleeding after vaccination with Oxford-AstraZeneca ChAdOx1-S in Denmark and Norway: population based cohort study. BMJ. 2021;373:n1114. doi:10.1136/bmj.n1114
6. Simpson CR, Shi T, Vasileiou E, et al. First-dose ChAdOx1 and BNT162b2 COVID-19 vaccines and thrombocytopenic, thromboembolic and hemorrhagic events in Scotland. Nat Med. 2021;27(7):1290–1297. doi:10.1038/s41591-021-01408-4
7. Wasay M, Kaul S, Menon B, et al. Asian study of cerebral venous sinus thrombosis. J Stroke Cerebrovasc Dis. 2019;28(10):104247. doi:10.1016/j.jstrokecerebrovasdis.2019.06.005
8. Duszynski KM, Stark JH, Cohet C, et al. Suitability of databases in the Asia-Pacific for collaborative monitoring of vaccine safety. Pharmacoepidemiol Drug Saf. 2021;30(7):843–857. doi:10.1002/pds.5214
9. Sung SF, Hsieh CY, Hu YH. Two decades of research using Taiwan’s National Health Insurance claims data: bibliometric and text mining analysis on PubMed. J Med Internet Res. 2020;22(6):e18457. doi:10.2196/jmir.21965
10. Handley JD, Emsley HC. Validation of ICD-10 codes shows intra-cranial venous thrombosis incidence to be higher than previously reported. Health Inf Manag. 2020;49(1):58–61. doi:10.1177/1833358318819105
11. Hsieh CY, Su CC, Shao SC, et al. Taiwan’s National Health Insurance Research Database: past and future. Clin Epidemiol. 2019;11:349–358. doi:10.2147/clep.S196293
12. Shao SC, Chan YY, Kao Yang YH, et al. The Chang Gung Research Database-a multi-institutional electronic medical records database for real-world epidemiological studies in Taiwan. Pharmacoepidemiol Drug Saf. 2019;28(5):593–600. doi:10.1002/pds.4713
13. Hsieh MT, Huang KC, Hsieh CY, Tsai TT, Chen LC, Sung SF. Validation of ICD-10-CM diagnosis codes for identification of patients with acute hemorrhagic stroke in a National Health Insurance claims database. Clin Epidemiol. 2021;13:43–51. doi:10.2147/clep.S288518
14. Hsieh FI, Lien LM, Chen ST, et al. Get With the Guidelines-Stroke performance indicators: surveillance of stroke care in the Taiwan Stroke Registry. *Circulation*. 2010;122(11):1116–1123. doi:10.1161/circulationaha.110.936526

15. Hsieh MT, Hsieh CY, Tsai TT, Wang YC, Sung SF. Performance of ICD-10-CM diagnosis codes for identifying acute ischemic stroke in a National Health Insurance Claims Database. *Clin Epidemiol*. 2020;12:1007–1013. doi:10.2147/clep.S273853