ORIGINAL RESEARCH

Validation of ICD-9-CM and ICD-10-CM Diagnostic Codes for Identifying Patients with Out-of-Hospital Cardiac Arrest in a National Health Insurance Claims Database

Ming-Jen Tsai [1],*, Cheng-Han Tsai[2,3],*, Ru-Chiou Pan[4], Chi-Feng Hsu [1], Sheng-Feng Sung [5,6]

[1]Department of Emergency Medicine, Ditmanson Medical Foundation Chia-Yi Christian Hospital, Chiayi City, Taiwan; [2]Department of Emergency Medicine, Chiayi Branch, Taichung Veteran's General Hospital, Chiayi City, Taiwan; [3]Department of Information Management, Institute of Healthcare Information Management, National Chung Cheng University, Chiayi County, Taiwan; [4]Clinical Data Center, Department of Medical Research, Ditmanson Medical Foundation Chia-Yi Christian Hospital, Chiayi City, Taiwan; [5]Division of Neurology, Department of Internal Medicine, Ditmanson Medical Foundation Chia-Yi Christian Hospital, Chiayi City, Taiwan; [6]Department of Nursing, Min-Hwei Junior College of Health Care Management, Tainan, Taiwan

*These authors contributed equally to this work

Correspondence: Sheng-Feng Sung, Division of Neurology, Department of Internal Medicine, Ditmanson Medical Foundation Chia-Yi Christian Hospital, Chiayi City, Taiwan, Tel +886 5 276 5041 Ext 7283, Fax +886 5 278 4257, Email sfusng@cych.org.tw; richard.sfsung@gmail.com

Purpose: Taiwan's national health insurance (NHI) database is a valuable resource for large-scale epidemiological and long-term survival research for out-of-hospital cardiac arrest (OHCA). We developed and validated case definition algorithms for OHCA based on the International Classification of Diseases (ICD) diagnostic codes and billing codes for NHI reimbursement.

Patients and Methods: Claims data and medical records of all emergency department visits from 2010 to 2020 were retrieved from the hospital's research-based database. Death-related diagnostic codes and keywords were used to identify potential OHCA cases, which were ascertained by chart reviews. We tested the performance of the developed OHCA algorithms and validated them on an external dataset.

Results: The algorithm defining OHCA as any cardiac arrest (CA)-related ICD code in the first three diagnosis fields performed the best with a sensitivity of 89.5% (95% confidence interval [CI], 88.2–90.7%), a positive predictive value (PPV) of 90.6% (95% CI, 89.4–91.8%), and a kappa value of 0.900 (95% CI, 0.891–0.909). The second-best algorithm consists of any CA-related ICD code in any diagnosis field with a billing code for triage acuity level 1, achieving a sensitivity of 85.6% (95% CI, 84.1–87.0%), a PPV of 93.6% (95% CI, 92.5–94.5), and a kappa value of 0.894 (95% CI, 0.884–0.903). Both algorithms performed well in external validation. In subgroup analyses, the former algorithm performed the best in adult patients, outpatient claims, and during the ICD-9 era. The latter algorithm performed the best in the inpatient claims and during the ICD-10 era. The best algorithm for identifying pediatric OHCAs was any CA-related ICD code in the first three diagnosis fields with a billing code for triage acuity level 1.

Conclusion: Our results may serve as a reference for future OHCA studies using the Taiwan NHI database.

Keywords: administrative claims data, diagnosis, ICD-9-CM, ICD-10-CM, out-of-hospital cardiac arrest, validation

Introduction

Out-of-hospital cardiac arrest (OHCA) is a global public health issue that causes almost 3.8 million deaths each year.[1] The prevalence and outcomes of OHCA vary widely around the world.[2,3] The annual incidence of OHCA treated by emergency medical services (EMS) ranges from 30.0 to 97.1 per 100,000 people worldwide, whereas the percentage of survival to discharge varies from 3.1% to 20.4% according to reports from different regions of the world.[2]

Although several national-level OHCA registries have been used to study the epidemiology of OHCA, most studies still rely on regional-level registries or locally collected data. Nevertheless, data collection is typically costly and time-consuming.

According to the Utstein-style registry reporting templates for OHCA, the core reporting elements consist primarily of the patient status and interventions received during the prehospital and in-hospital stages from cardiac arrest to hospital discharge.[4,5] Data on survivors' long-term outcomes after discharge is typically absent. Hence, healthcare resource utilization by OHCA survivors following initial hospital discharge cannot be studied by using the Utstein-style OHCA registry. Besides, most OHCA registries only collect EMS-treated OHCAs,[2] and OHCA patients bypassing the EMS system may not be recorded. The incidence of OHCA could thus be underestimated by using such registries.

Moreover, no national OHCA registry currently exists in Taiwan; only local data are collected by each county or city's fire bureau-based EMS system. The data elements of these local registries also vary across regions. These limitations prevent researchers from investigating the impact of health interventions and resource utilization on the long-term outcomes of OHCA survivors and large-scale epidemiological studies of OHCA in Taiwan. In this regard, Taiwan's National Health Insurance (NHI) claims database may serve as an alternative data source for such studies if a validated algorithm to identify OHCA is available.

So far, OHCA has seldom been studied using Taiwan's NHI claims database. Wang et al used the NHI database to examine the prevalence and outcomes of OHCA in Taiwan between 2000 and 2012.[6] Hsu et al investigated the risk of OHCA among sepsis survivors in Taiwan between 2000 and 2013.[7] However, the validity of the International Classification of Diseases (ICD) diagnostic codes used to ascertain OHCA has not been adequately assessed in these studies. Besides, the coding system transition from ICD-9 to ICD-10 in Taiwan in 2016 may lead to inconsistent estimation of disease prevalence across the transition.[8]

To use Taiwan's NHI claims database to develop a national OHCA database for disease surveillance and research, we aimed to develop and validate case definition algorithms for OHCA in Taiwan's NHI claims database using data from two hospitals in Taiwan.

Materials and Methods
Study Settings

We conducted this retrospective study in two NHI-contracted hospitals in Taiwan. Almost all the hospitals in Taiwan are contracted with the NHI. All NHI-contracted hospitals must submit claims to the NHI Administration in a standard format. The dataset for developing case definition algorithms was obtained from the Ditmanson Medical Foundation Chia-Yi Christian Hospital, a 1000-bed private tertiary teaching hospital with approximately 220 OHCA admissions annually. The dataset for external validation of the algorithms was obtained from the Taichung Veterans General Hospital, Chiayi branch, a 650-bed public veterans' hospital with approximately 120 OHCA admissions each year. This hospital is ideal for external validation because it belongs to a large system of 15 veterans' hospitals spanning urban and rural areas of Taiwan. Currently, 13 of them, including the external validation hospital, use the same electronic medical record (EMR) system and health information system. The study protocol was approved independently by the institutional review boards of Ditmanson Medical Foundation Chia-Yi Christian Hospital (IRB2021080) and Taichung Veterans General Hospital (SE21370A) with a waiver of informed consent.

Taiwan's National Health Insurance Claims Database

Taiwan's NHI program, a compulsory single-payer healthcare system, was initiated in 1995 to provide affordable healthcare to all residents in Taiwan. It covers nearly all services needed for disease diagnosis and treatment, including inpatient care, outpatient care, laboratory testing, prescriptions, dental services, home care, and preventative services, but not including EMS. The NHI claims database comprehensively includes all the longitudinal claims records of nearly 99.8% of Taiwan's population (23.5 million), providing useful real-world data for studying disease prevalence, long-term prognosis, and healthcare policies.[9,10] NHI claims data consists of outpatient claims, inpatient claims, prescriptions dispensed at pharmacies, and registries for beneficiaries, medical facilities, and board-certified specialists. These datasets can be linked with encrypted personal identification numbers to provide patient-level information on demographic characteristics for research purposes.[9] The claims records of patients who are discharged directly from the emergency department (ED) including those who die in ED are included in the outpatient claims. In contrast, claims records of ED

patients who are eventually admitted to the intensive care unit or ward are integrated with the hospitalization claims records, which are included in the inpatient claims. The NHI claims database released for research allows a maximum of 3 diagnoses on outpatient claims and 5 diagnoses on inpatient claims. Diagnoses have been coded using the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) until the end of 2015. From 2016 onwards, diagnoses are coded using the International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM).

Data Sources and Participants

The data used for developing case definition algorithms were obtained from the Ditmanson Research Database, a deidentified research-based database that contains all claims data submitted to the NHI and EMRs. It presently holds clinical information of nearly 1.4 million patients who were treated at the Ditmanson Medical Foundation Chia-Yi Christian Hospital between January 2006 and February 2021, including 0.6 million inpatient and 21.5 million outpatient records. We extracted both claims data and medical records of all ED visits between January 2010 and December 2020 to develop case definition algorithms for OHCA. The data used for externally validating the developed algorithms were extracted from the hospital EMRs and claims data of all ED visits to Taichung Veterans General Hospital, Chiayi branch from March 2019 to December 2020.

Case Definition Algorithms

Table 1 lists the algorithms used to identify OHCA from the NHI claims data. The algorithms included the presence of ICD diagnostic codes related to cardiac arrest (CA) (Supplementary Table 1). Variations to the basic algorithm included: 1. addition of ICD diagnostic codes related to fatal ventricular arrhythmia (VA) (Supplementary Table 1), 2. the position of the ICD codes, that is, the primary diagnosis field, the first three diagnosis fields, or any diagnosis field, 3. presence of a billing code for triage acuity level 1 (00201B), and 4. presence of a billing code for cardiopulmonary resuscitation (CPR) (47029C).

Table 1 Case Definition Algorithms for Out-of-Hospital Cardiac Arrest

Algorithm	CA Codes	VA Codes	Primary Diagnosis Field	First Three Diagnosis Fields	Any Diagnosis Field	AND Billing Code for Triage Acuity Level 1	AND Billing Code for CPR
A	X		X				
B	X	X	X				
C	X			X			
D	X	X		X			
E	X				X		
F	X	X			X		
G	X		X			X	
H	X	X	X			X	
I	X			X		X	
J	X	X		X		X	
K	X				X	X	
L	X	X			X	X	
M	X		X				X
N	X	X	X				X
O	X			X			X
P	X	X		X			X
Q	X				X		X
R	X	X			X		X

Abbreviations: CA, cardiac arrest; CPR, cardiopulmonary resuscitation; VA, ventricular arrhythmia.

Ascertainment of Out-of-Hospital Cardiac Arrest

In Taiwan, the prehospital ambulance records of OHCA must be provided to ED physicians for medical record documentation on patient arrival at the ED. Therefore, the patient status and procedures performed before arrival, such as witnessed status, bystander CPR, use of the automated external defibrillator, initial cardiac rhythm, and location of cardiac arrest, etc., can be obtained from hospital medical records. In this study, we used two methods to inclusively identify all patients with OHCA, that is, the forward keyword search and the backward ICD code search (Figure 1). In the forward search, physicians' notes of all ED visits during the study period were screened for keywords related to OHCA (Supplementary Table 2) to identify potential OHCA cases. These keywords were curated by manually reviewing the physicians' notes of a random sample of 100 OHCA patients and were refined iteratively with four rounds of preliminary searches. In the backward search, ICD diagnostic codes related to CA or VA were used to identify potential OHCA cases from all inpatient and outpatient claims during the study period. The medical records of the identified potential OHCA cases were manually reviewed by two of the investigators (MJT and CFH) to determine if each patient was a true OHCA according to the Utstein criteria.[9] Discrepancies between the two investigators were resolved by consensus. The consensus diagnosis was considered the gold standard for all analyses (Figure 1).

Analysis and External Validation

The performance of each algorithm in identifying true OHCA cases was assessed using sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) according to the confusion matrix (Supplementary Figure 1). Their 95% confidence intervals (CI) were calculated using the Clopper-Pearson exact method. The main evaluation metric was the kappa coefficient, which was calculated to examine the agreement between algorithm-detected OHCA cases and true OHCA patients. The degree of agreement was interpreted as follows: "slight" (0.00–0.20), "fair" (0.21–0.40), "moderate" (0.41–0.60), "substantial" (0.61–0.80), and "perfect" (0.81–1.00).[11]

The top two algorithms with the highest kappa coefficients were externally validated on an independent dataset from Taichung Veterans General Hospital, Chiayi branch, where a similar process was undertaken to identify potential OHCA cases. One of the investigators (CHT) used the same criteria to determine if each patient was a true OHCA. In addition, we performed subgroup analysis stratified by ICD coding system (ICD-9-CM or ICD-10-CM), year of ED visits, age
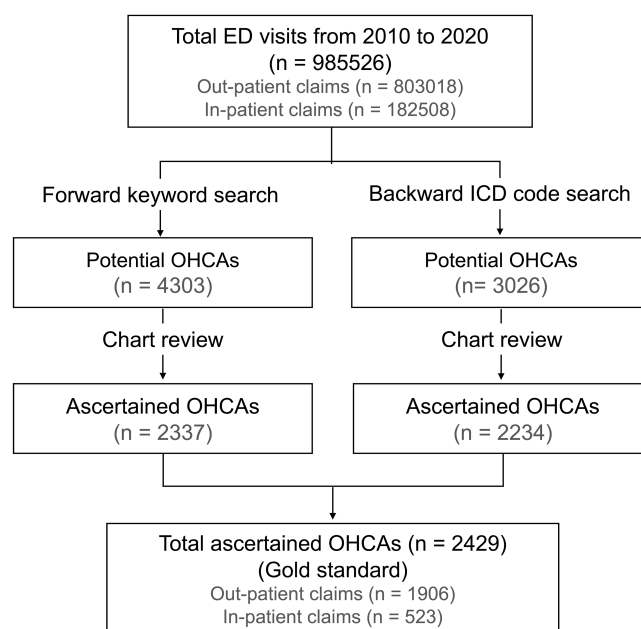

Figure 1 Flow chart outlining the true out-of-hospital cardiac arrest verification process.
Abbreviations: ED, emergency department; ICD, International Classification of Diseases diagnostic codes; OHCA, out-of-hospital cardiac arrest.

724

DovePress

Clinical Epidemiology 2022:14

(pediatric [<18 years] or adult [≥18 years]), and type of claims (outpatient or inpatient claims). Data analyses were performed using Stata 17.0 (StataCorp, College Station, Texas). Statistical significance was set at two-tailed p <0.05.

Results

The development cohort consisted of 985,526 ED visits, including 803,018 visits from the outpatient claims and 182,508 visits from the inpatient claims. The forward keyword search identified 4303 potential OHCA cases, of whom 2337 were adjudicated to be true OHCAs by manual chart review according to Utstein criteria (Figure 1). The backward ICD code search identified 3026 potential OHCA cases, and 2234 of them were adjudicated as true OHCAs by chart review. In total, 2429 OHCA cases were ascertained during the study period.

Results for the validity of the best two and all algorithms to identify OHCA are shown in Table 2 and Supplementary Table 3, respectively. The most excellent algorithm was algorithm C, which defines OHCA as any CA-related ICD code (ICD-9-CM codes: 798, 798.1, 798.2, 798.9, 799.9, 427.5; ICD-10-CM codes: R99, I46, I46.2, I46.8, I46.9) in the first three diagnosis fields, yielding a sensitivity, specificity, PPV, NPV, and kappa of 89.5%, 100%, 90.6%, 100%, and 0.900, respectively. The second-best algorithm was algorithm K, which defines OHCA as any CA-related ICD code in any diagnosis field with a billing code for triage acuity level 1. Its sensitivity, specificity, PPV, NPV, and kappa were 85.6%, 100%, 93.6%, 100%, and 0.894, respectively. In general, adding the billing code for triage acuity level 1 to the algorithms slightly increased the PPV at the expense of a reduction in the sensitivity (Supplementary Table 3, algorithms G to L vs A to F). However, adding the billing code for CPR largely decreased the sensitivity and kappa (Supplementary Table 3, algorithms M to R). The reasons for the false-negative and false-positive identification of OHCA by algorithms C and K were summarized in Table 3. The most common reason for false negatives for both algorithms C and K was being coded for underlying diseases without properly coded with CA-related ICD codes. Another major reason for false negatives for algorithm K was without a billing code for triage acuity level 1. On the other hand, in-hospital cardiac arrest (IHCA) occurring during the hospitalization or those happening at the ED were the main causes of false positives for both algorithms C and K.

Algorithms C and K were externally validated on the validation cohort, which consisted of 30,849 ED visits. The sensitivity, specificity, PPV, NPV and kappa of algorithm C were 98.8%, 99.9%, 93.1%, 100%, and 0.958, respectively, while algorithm K achieved 98.8%, 100%, 94.8%, 100%, and 0.968, respectively. Both algorithms performed well in identifying OHCA in the external validation cohort.

In the subgroup analysis stratified by the ICD coding system (Table 4 and Supplementary Table 4), the best algorithm during the ICD-9 era was algorithm C, yielding a sensitivity, specificity, PPV, NPV and kappa of 90.8%, 100%, 92.3%, 100%, and 0.915, respectively. During the ICD-10 era, algorithm K had the highest kappa value (0.898). Its sensitivity, specificity, PPV and NPV were 88.4%, 100%, 91.2%, 100%, respectively. We compared the sensitivity and PPV of algorithms C and K by year to assess the impact of coding system transition. As shown in Figure 2, algorithms C and K had an apparent drop in the sensitivity (74.0% and 73.2%, respectively) in the year 2016, indicating decreased coding accuracy for OHCA just after the transition of the coding system.

Table 2 Performance of the Best Two Algorithms for All Patients (n = 985,526)

		Ascertained OHCA		Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)	Kappa (95% CI)
		Positive	Negative					
Alg C	**Positive**	2174	225	89.5 (88.2–90.7)	100.0 (100.0–100.0)	90.6 (89.4–91.8)	100.0 (100.0–100.0)	0.900 (0.891–0.909)
	Negative	255	982,872					
Alg K	**Positive**	2079	143	85.6 (84.1–87.0)	100.0 (100.0–100.0)	93.6 (92.5–94.5)	100.0 (100.0–100.0)	0.894 (0.884–0.903)
	Negative	350	982,954					

Abbreviations: Alg, algorithm; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

Table 3 Reasons for False-Negative and False-Positive Identification of OHCA by Algorithms C and K

Reason	n
False negatives for algorithm C	**255**
Coded for underlying diseases instead of CA	121
Miscoded as unspecified illness or symptoms (R69, R68.89 or 799.89)	59
CA codes listed in the 4th or 5th diagnosis fields	49
VA codes listed in the 4th and 5th diagnosis fields	11
VA codes listed in the first three diagnosis fields	15
False positives for algorithm C	**225**
IHCA occurring during the hospitalization	130
IHCA occurring at the emergency department	72
Miscoded as CA or VA	23
False negatives for algorithm K	**350**
Coded for underlying diseases instead of CA	144
Without a billing code for triage acuity level 1	144
Miscoding as unspecified illness or symptoms codes (R69, R68.89 or 799.89)	52
VA codes listed in the 4th and 5th diagnosis fields	10
False positives for algorithm K	**143**
IHCA occurring during the hospitalization	81
IHCA occurring at the emergency department	61
Miscoded as CA or VA	1

Abbreviations: CA, cardiac arrest; IHCA, in-hospital cardiac arrest; OHCA, out-of-hospital cardiac arrest; VA, ventricular arrhythmia.

Table 4 Performance of the Best Algorithms Among Different Subgroups Stratified by ICD Coding System, Age, and Type of Claims

Subgroup	Alg	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)	Kappa (95% CI)
ICD-9-CM (2010–2015) (n = 561,331)	C	90.8 (89.2–92.3)	100.0 (100.0–100.0)	92.3 (90.7–93.6)	100.0 (100.0–100.0)	0.915 (0.904–0.926)
ICD-10-CM (2016–2020) (n = 424,195)	K	88.4 (86.3–90.2)	100.0 (100.0–100.0)	91.2 (89.3–92.9)	100.0 (100.0–100.0)	0.898 (0.884–0.911)
Adult patients (≥18 y/o) (n = 731,993)	C	89.6 (88.3–90.8)	100.0 (100.0–100.0)	91.0 (89.7–92.1)	100.0 (100.0–100.0)	0.903 (0.894–0.912)
Pediatric patients (<18 y/o) (n = 253,533)	I	74.5 (61.0–85.3)	100.0 (100.0–100.0)	100.0 (91.4–100.0)	100.0 (100.0–100.0)	0.854 (0.779–0.930)
Outpatient claims (n = 803,018)	C	90.8 (89.4–92.1)	100.0 (100.0–100.0)	96.4 (95.4–97.2)	100.0 (100.0–100.0)	0.935 (0.927–0.943)
Inpatient claims (n = 182,508)	K	89.1 (86.1–91.6)	99.9 (99.9–100.0)	80.9 (77.5–84.0)	100.0 (100.0–100.0)	0.848 (0.825–0.870)

Abbreviations: Alg, algorithm; CI, confidence interval; ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification; ICD-10-CM, International Classification of Diseases, Tenth Revision, Clinical Modification; PPV, positive predictive value; NPV, negative predictive value.
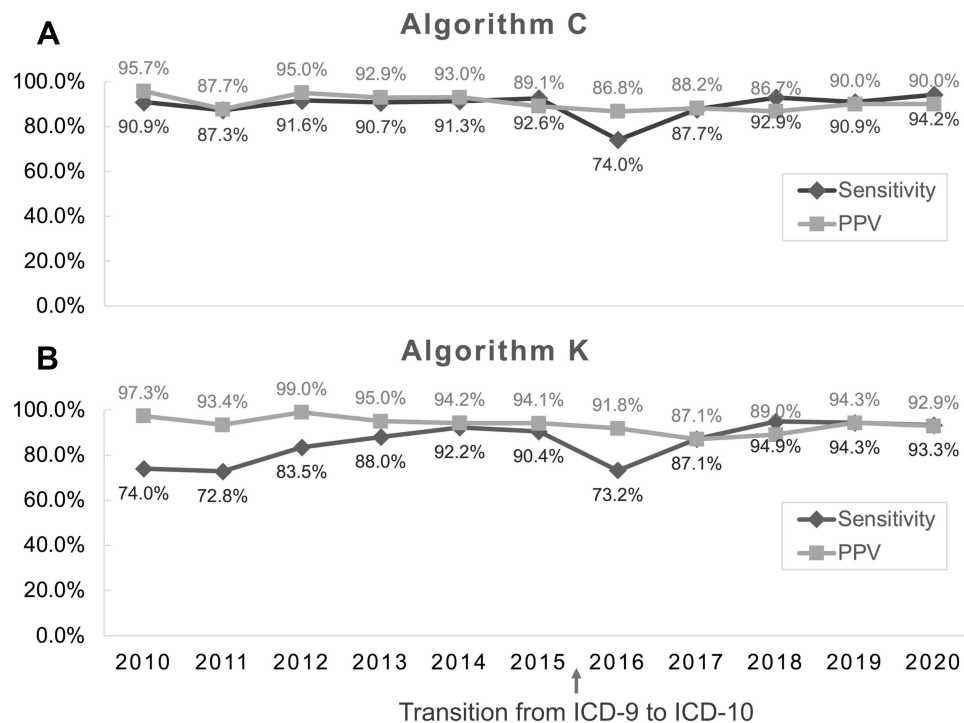
Figure 2 The sensitivity and positive predictive value of algorithms C (**A**) and K (**B**) to identify OHCA from 2010 to 2020.
Abbreviations: ICD, International Classification of Diseases diagnostic codes; PPV, positive predictive value.

In the subgroup analysis stratified by age (Table 4 and Supplementary Table 5). Algorithm C and K had kappa values of 0.903 and 0.895, making them the top two algorithms in adult patients. However, these algorithms performed lower in pediatric patients, with kappa values of 0.800 and 0.845, respectively. The best algorithm for identifying pediatric OHCAs was algorithm I, which attained a sensitivity, PPV, and kappa of 74.5%, 100% and 0.854, respectively. In the subgroup analysis stratified by type of claims (Table 4 and Supplementary Table 5), the algorithms generally performed lower in the inpatient claims than in the outpatient claims. Algorithm C is still the best algorithm to identify OHCAs in the outpatient claims with a kappa value of 0.935. However, the kappa value of algorithm K (0.848) was superior to algorithm C (0.786) in identifying OHCAs in the inpatient claims. In other words, in addition to using CA-related diagnostic codes, the billing code for triage acuity level 1 should be added to identify OHCA cases in the inpatient claims.

Discussion

This study assessed the validity of OHCA case definitions based on ICD diagnostic codes and billing codes in Taiwan's NHI claims database. Overall, the algorithm that defines OHCA as any CA-related ICD code in the first three diagnosis fields performed the best, followed by the algorithm that defines OHCA as any CA-related ICD code in any diagnosis field with a billing code for triage acuity level 1. Both algorithms performed very well in the external validation cohort. In subgroup analyses, the former algorithm also performed the best in adult patients, the outpatient claims, and during the ICD-9 era. The latter algorithm had the highest performance in the inpatient claims and during the ICD-10 era. On the other hand, the best algorithm for identifying pediatric OHCAs was the one that defines OHCA as any CA-related ICD code in the first three diagnosis fields with a billing code for triage acuity level 1.

The main reason for the false-positive identification of OHCA by using the two best-performing algorithms (C and K) was the misclassification of cases of IHCA that occurred either during the hospitalization or at the ED (Table 3). Given the lack of specific ICD codes, it is always a challenge to differentiate IHCA from OHCA.[12] Nonetheless, the number of false positives can be reduced by adding the billing code for triage acuity level 1 to the algorithm and limiting CA-related ICD codes to the primary diagnosis field. For example, algorithm G, which defines OHCA as any CA-related codes in the primary diagnosis field with a billing code for triage acuity level 1, yielded a PPV as high as 98.0% (95% CI 97.2–98.6%)

(Supplementary Table 3). Such an algorithm is well suited for studies where a cohort exclusively consisting of OHCA patients is needed, despite risking more false negatives and decreased sensitivity.

On the other hand, the false-negative identification of OHCA commonly occurred because of the failure to code OHCA visits with proper CA-related ICD codes or the placement of CA-related codes outside the first three diagnosis fields (Table 3). Stress experienced during resuscitation is likely to impact the decision-making and performance of ED physicians,[13] leading to either erroneous clinical diagnosis or incorrect or inadequate coding of the diagnosis. This is one drawback of using ICD-based algorithms to retrieve OHCA cases from claim databases. However, by incorporating more diagnosis fields or expanding OHCA-related ICD codes to include VA codes, such as algorithm F, the number of false negatives can be reduced with improved sensitivity (Supplementary Table 3).

Notably, the coding system transition harms the validity of OHCA algorithms. Taiwan's NHI changed the coding system from ICD-9-CM to ICD-10-CM in 2016. The unfamiliarity with the new coding system could increase coding errors, as shown by an apparent drop in algorithm sensitivity in 2016 (Figure 2). Previous studies have also demonstrated that coding system transition may lead to inconsistent estimation of disease prevalence in some disease conditions.[8,14] Therefore, researchers should be cautioned about the use of ICD code-based algorithms in the year immediately after coding system transition.

Furthermore, the validity of OHCA algorithms may vary across conditions such as different age groups and types of claims. For example, algorithms using ICD codes alone, such as algorithm C, performed better in adults than in children (Supplementary Table 5). The algorithm performance could be improved by adding the billing code for triage acuity level 1, as shown by algorithm I. Meanwhile, algorithms using ICD codes alone generally performed worse in the inpatient claims than in the outpatient claims (Supplementary Table 5). Similar findings have been shown in previous studies. The reason may be that algorithms using CA-related ICD codes alone is likely to identify IHCA rather than OHCA when inpatients claims are used as the data source.[15,16] In this case, an algorithm that combines CA-related ICD codes with a billing code for triage acuity level 1 (algorithm K) will be more appropriate. Considering the above findings, researchers interested in using Taiwan's NHI database for OHCA research may refer to our findings to select the optimal algorithm that best suits their research questions.

Up until now, only limited studies have examined the validity of ICD codes for the diagnosis of OHCA in health claims databases or national registries. In the US, Hennessy et al found that the first-listed CA and VA-related ICD-9-CM diagnostic codes (427.1, 427.4, 427.41, 427.42, 427.5, 798, 798.1, 798.2) had an overall PPV of 85.3% in identifying OHCA and VA in inpatient and ED claims of Medicaid and Medicare between 1999 and 2002.[15] Shelton et al reported a single-center validation of an ICD-9 code (427.5) for identifying ED patients with OHCA in US administrative databases between 2007 and 2012, with a sensitivity, specificity, PPV, and kappa of 86.5%, 99.4%, 92.0%, and 0.895, respectively.[17] Gray et al used single-hospital EMRs and the Canadian Resuscitation Outcome Consortium Database to validate the ICD-10-CM codes including a diagnostic code of cardiac arrest (I46), or a CPR intervention code combined with codes for sudden infant death syndrome (R95, R99), drowning (W65, W67, W69, W74, T75.1), and asphyxiation (T71, R09.0, R09.2). A sensitivity and PPV of 87.3% and 81.4% were reported for identifying pediatric OHCA in Canadian administrative data.[18] Nevertheless, validation studies on OHCA case definitions using ICD codes are scarce in other countries.[19]

Several population-based OHCA studies have been conducted using administrative databases in the US[20–24] and outside the US[7,25,26] based on the above-mentioned validated OHCA-related diagnostic codes. However, because coding practices and reimbursement policies may differ across healthcare systems, it should not be taken for granted that the validity of case definitions established in one database can be extrapolated to another database. So far, OHCA has seldom been investigated using Taiwan's NHI database. One possible reason may be the lack of adequate assessment of the validity of OHCA case definitions. In Wang's study that investigated the prevalence and outcomes of OHCA using Taiwan's NHI database,[6] OHCA was defined as ED patients coded with one of the death-related ICD-9-CM codes (798, 798.1, 798.2 and 798.9) as well as a billing code for CPR. However, as shown in this study (Supplementary Table 3), the sensitivity of identifying OHCA was as low as 50% to 60% when a billing code for CPR was added to the algorithm. In addition, the ICD-9-CM codes used in Wang's study were not as comprehensive as ours. Similar issues also arise in Hsu's study that investigated the risk of OHCA among sepsis survivors in Taiwan, where the first-listed ICD-9-CM codes of 427.5 and 798 as well as a billing code for CPR were used to define OHCA.[7] Without adequate data on the validity of OHCA diagnosis, their study findings would be harder to interpret.

National-level health insurance databases, which contain the longitudinal healthcare data of the whole population, provide useful real-world data for research investigating epidemiology, long-term outcomes, and healthcare utilization, and contribute to healthcare policy-making.[9] Nevertheless, before using such claims databases for research of a given disease, the validity of disease diagnosis should be examined. To this end, we believe that this study will provide a critical and meaningful reference for future OHCA research using Taiwan's NHI claims database.

This study has limitations worth noting. First, the algorithms were developed using data from only one hospital, just like some previous studies.[17,18] It is thus unclear whether the study findings can be generalized to the whole population covered by Taiwan's NHI. However, at least, the developed algorithms performed well on the dataset from another hospital. Second, the OHCA algorithms can only be applied to studies using Taiwan's NHI database for the time being. Nevertheless, future studies may investigate these algorithms using databases from other healthcare systems with similar coding practices. Third, claims data only includes patients who have had contact with the healthcare system. OHCAs not transported to the hospital, such as patients whose resuscitation was terminated and declared dead at the scene or those with a do-not-resuscitate directive, will not be identified from the claims database. Fourth, the case definition algorithms developed in this study cannot differentiate traumatic from non-traumatic OHCAs. Further studies are needed to investigate whether and how these two types of OHCAs can be distinguished using diagnostic or billing codes available in Taiwan's NHI database.

Conclusion

Validation of case definitions is required before they can be applied in epidemiological studies using claims databases. Taiwan's NHI claims database, one of the largest comprehensive administrative claims databases in the world, currently lacks a validated case definition for OHCA. This study developed and assessed various case definition algorithms for OHCA based on ICD diagnostic codes and billing codes for NHI reimbursement. The results of this validation study may aid future researchers in applying appropriate case definitions for OHCA that best suit their research objectives.

Acknowledgments

Disclosure

The authors report no conflicts of interest in this work.

References

1. Kudenchuk PJ, Sandroni C, Drinhaus HR, et al. Breakthrough in cardiac arrest: reports from the 4th Paris International Conference. *Ann Intensive Care*. 2015;5(1):22. doi:10.1186/s13613-015-0064-x

2. Kiguchi T, Okubo M, Nishiyama C, et al. Out-of-hospital cardiac arrest across the World: first report from the International Liaison Committee on Resuscitation (ILCOR). *Resuscitation*. 2020;152:39–49. doi:10.1016/j.resuscitation.2020.02.044

3. Berdowski J, Berg RA, Tijssen JG, Koster RW. Global incidences of out-of-hospital cardiac arrest and survival rates: systematic review of 67 prospective studies. *Resuscitation*. 2010;81(11):1479–1487. doi:10.1016/j.resuscitation.2010.08.006

4. Perkins GD, Jacobs IG, Nadkarni VM, et al. Cardiac arrest and cardiopulmonary resuscitation outcome reports: update of the utstein resuscitation registry templates for out-of-hospital cardiac arrest: a statement for healthcare professionals from a task force of the International Liaison Committee on Resuscitation (American Heart Association, European Resuscitation Council, Australian and New Zealand Council on Resuscitation, Heart and Stroke Foundation of Canada, InterAmerican Heart Foundation, Resuscitation Council of Southern Africa, Resuscitation Council of Asia); and the American Heart Association Emergency Cardiovascular Care Committee and the Council on Cardiopulmonary, Critical Care, Perioperative and Resuscitation. *Circulation*. 2015;132(13):1286–1300. doi:10.1161/CIR.0000000000000144

5. Jacobs I, Nadkarni V, Bahr J, et al. Cardiac arrest and cardiopulmonary resuscitation outcome reports: update and simplification of the Utstein templates for resuscitation registries: a statement for healthcare professionals from a task force of the International Liaison Committee on Resuscitation (American Heart Association, European Resuscitation Council, Australian Resuscitation Council, New Zealand Resuscitation Council, Heart and Stroke Foundation of Canada, InterAmerican Heart Foundation, Resuscitation Councils of Southern Africa). *Circulation*. 2004;110(21):3385–3397. doi:10.1161/01.CIR.0000147236.85306.15

6. Wang CY, Wang JY, Teng NC, et al. The secular trends in the incidence rate and outcomes of out-of-hospital cardiac arrest in Taiwan–a nationwide population-based study. *PLoS One*. 2015;10(4):e0122675. doi:10.1371/journal.pone.0122675

7. Hsu WT, Sherrod CF, Tehrani B, et al. Risk of out-of-hospital cardiac arrest among sepsis survivors in Taiwan: analysis of a nationwide population-based cohort. *BMJ Open*. 2021;11(9):e051502. doi:10.1136/bmjopen-2021-051502

8. Hsu MC, Wang CC, Huang LY, Lin CY, Lin FJ, Toh S. Effect of ICD-9-CM to ICD-10-CM coding system transition on identification of common conditions: an interrupted time series analysis. *Pharmacoepidemiol Drug Saf*. 2021;30(12):1653–1674. doi:10.1002/pds.5330

9. 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

10. National Health Insurance Administration. NHI profile. Universal health coverage in Taiwan. Avaiable from: https://www.nhi.gov.tw/English/Content_List.aspx?n=4D7051840BF42F52&topn=ED4A30E51A609E49. Accessed February 24, 2022.

11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174. doi:10.2307/2529310

12. Khera R, Spertus JA, Starks MA, et al. Administrative codes for capturing in-hospital cardiac arrest. *JAMA Cardiol*. 2017;2(11):1275–1277. doi:10.1001/jamacardio.2017.2904

13. Groombridge CJ, Maini A, Ayton D, et al. Emergency physicians' experience of stress during resuscitation and strategies for mitigating the effects of stress on performance. *Emerg Med J*. 2021;38(1):14. doi:10.1136/emermed-2020-209432

14. Metcalfe A, Sheikh M, Hetherington E. Impact of the ICD-9-CM to ICD-10-CM transition on the incidence of severe maternal morbidity among delivery hospitalizations in the United States. *Am J Obstet Gynecol*. 2021;225(4):422e421–422 e411. doi:10.1016/j.ajog.2021.03.036

15. Hennessy S, Leonard CE, Freeman CP, et al. Validation of diagnostic codes for outpatient-originating sudden cardiac death and ventricular arrhythmia in Medicaid and medicare claims data. *Pharmacoepidemiol Drug Saf*. 2010;19(6):555–562. doi:10.1002/pds.1869

16. De Bruin ML, van Hemel NM, Leufkens HG, Hoes AW. Hospital discharge diagnoses of ventricular arrhythmias and cardiac arrest were useful for epidemiologic research. *J Clin Epidemiol*. 2005;58(12):1325–1329. doi:10.1016/j.jclinepi.2005.04.009

17. Shelton SK, Chukwulebe SB, Gaieski DF, Abella BS, Carr BG, Perman SM. Validation of an ICD code for accurately identifying emergency department patients who suffer an out-of-hospital cardiac arrest. *Resuscitation*. 2018;125:8–11. doi:10.1016/j.resuscitation.2018.01.021

18. Gray K, Cameron S, McKenzie K, Miller M, Odoardi N, Tijssen JA. Validation of ICD-10 codes for the identification of paediatric out-of-hospital cardiac arrest patients. *Resuscitation*. 2021;171:73–79. doi:10.1016/j.resuscitation.2021.12.017

19. Wittwer MR, Ruknuddeen MI, Thorrowgood M, Zeitz C, Beltrame JF, Arstall MA. Overcoming challenges of establishing a hospital-based out-of-hospital cardiac arrest registry: accuracy of case identification using administrative data and clinical registries. *Resusc Plus*. 2021;6:100136. doi:10.1016/j.resplu.2021.100136

20. Dhopeshwarkar N, Brensinger CM, Bilker WB, et al. Risk of sudden cardiac arrest and ventricular arrhythmia with sulfonylureas: an experience with conceptual replication in two independent populations. *Sci Rep*. 2020;10(1):10070. doi:10.1038/s41598-020-66668-5

21. Fung KW, Baye F, Kapusnik-Uner J, McDonald CJ. Using medicare data to assess the proarrhythmic risk of non-cardiac treatment drugs that prolong the QT interval in older adults: an observational cohort study. *Drugs Real World Outcomes*. 2021;8(2):173–185. doi:10.1007/s40801-021-00230-1

22. Khan MZ, Sulaiman S, Agrawal P, et al. Targeted temperature management in cardiac arrest patients with a non-shockable rhythm: a national perspective. *Am Heart J*. 2020;225:129–137. doi:10.1016/j.ahj.2020.04.023

23. Patel H, Calip GS, DiDomenico RJ, Schumock GT, Suda KJ, Lee TA. Comparison of cardiac events associated with azithromycin vs amoxicillin. *JAMA Netw Open*. 2020;3(9):e2016864. doi:10.1001/jamanetworkopen.2020.16864

24. Perman SM, Siry BJ, Ginde AA, et al. Sex differences in "do not attempt resuscitation" orders after out-of-hospital cardiac arrest and the relationship to critical hospital interventions. *Clin Ther*. 2019;41(6):1029–1037. doi:10.1016/j.clinthera.2019.03.017

25. Kim JS, Seo DW, Kim YJ, et al. Prolonged length of stay in the emergency department and increased risk of in-hospital cardiac arrest: a nationwide population-based study in South Korea, 2016–2017. *J Clin Med*. 2020;9(7):2284. doi:10.3390/jcm9072284

26. Cowan A, Garg AX, McArthur E, Muanda Tsobo F, Weir MA. Cardiovascular safety of metoclopramide compared to domperidone: a population-based cohort study. *J Can Assoc Gastroenterol*. 2021;4(5):e110–e119. doi:10.1093/jcag/gwaa041