International Journal of Cardiology 201 (2015) 96–101

Contents lists available at ScienceDirect

International Journal of Cardiology

journal homepage: www.elsevier.com/locate/ijcard

Validity of in-hospital mortality data among patients with acute myocardial infarction or stroke in National Health Insurance Research Database in Taiwan

Ching-Lan Cheng [a,b], Hsu-Chih Chien [a], Cheng-Han Lee [c], Swu-Jane Lin [d], Yea-Huei Kao Yang [a,b,*]

[a] School of Pharmacy and Institute of Clinical Pharmacy and Pharmaceutical Sciences, College of Medicine, National Cheng Kung University, Tainan, Taiwan
[b] Health Outcome Research Center, National Cheng Kung University, Tainan, Taiwan
[c] Department of Internal Medicine, National Cheng Kung University Hospital, Tainan, Taiwan
[d] University of Illinois at Chicago, Chicago, IL, USA

ARTICLE INFO

Article history:
Received 9 July 2015
Accepted 29 July 2015
Available online 1 August 2015

Keywords:
Acute myocardial infarction
Stroke
In-hospital mortality
NHIRD
Taiwan
Validation

ABSTRACT

Objective: The aim of this study was to determine the validity of in-hospital mortality records in the National Health Insurance Research Database (NHIRD) by cross-comparing with death records from the electronic medical records (EMR) of a medical center in southern Taiwan.

Methods: Data on patients admitted to the medical center for acute myocardial infarction (AMI) or stroke during the years 2005 to 2010 were extracted from the two databases and cross-linkages with patients' characteristics (birth date, gender, admission date, and discharge date). While the death record was available in the catastrophic illness registry data files (CIRD), we also estimated the insurance status and death record in the CIRD subset using confirmed death cases. Additionally, agreement in comorbidities between records from the two databases was evaluated.

Results: A total of 6197 cases were successfully linked, with a linkage rate of 96.56% of cases in the NHIRD when linked to those from the EMR. Among the linked population, 538 of 682 patients retrieved as expired in the NHIRD were also so recorded in the EMR. This yielded a positive predictive value of 0.79 when the EMR was used as the gold standard. Patients having death records in both the CIRD subset and the EMR totaled 364, which yielded a percentage positive agreement rate of 76%. The consistency in comorbidity diagnoses between the two databases was more than 90% among matched cases.

Conclusions: The accuracy of death records in the NHIRD was high, and appears to be a valid resource for population research in cardiovascular diseases.

1. Introduction

The reduction of the in-hospital mortality rate of acute myocardial infarction (AMI) or stroke is an important quality indicator for hospital performance or procedures, such as percutaneous coronary intervention or thrombolytic therapy, among these patients [1–4]. Population-based administration or claims database are two valuable resources to perform epidemiological studies and healthcare quality evaluation [5,6].

Understanding the underlying diseases (e.g. AMI or stroke) of mortality events has significant clinical, economic, and quality-of-life implications. Linkage to death certificate data is frequently used in epidemiology studies to confirm the cause of death for each individual. However, the accuracy of the certified cause of death is often a function of linkage accessibility, the experience of the researchers, and the specific health conditions [7–9]. Although cause of death from the death certificate might not be obtainable for some research, the primary discharge condition from hospitalization with a fatality is likely to be consistent with the cause of death and therefore could be considered a proxy indicator in epidemiological research [9].

The National Health Insurance Research Database (NHIRD) is a uniform population-based database, compiled from data from beneficiaries covered by the National Health Insurance (NHI) of Taiwan. Currently, the NHI covers over 99% of Taiwan's population, and the NHIRD has been widely used in epidemiology studies of cardiovascular disease (CVD) or drug-related cerebrovascular disorders [10,11]. Our previous studies have shown that the positive predict value (PPV) for patients with a primary diagnosis of acute myocardial infarction or ischemic stroke in the NHIRD was more than 90%, thus the database was deemed appropriate for research in this disease area [12,13]. However, mortality data in the NHIRD has not yet been validated because it could not be linked to the death certification records.

* Corresponding author at: University Road, School of Pharmacy and Institute of Clinical Pharmacy and Pharmaceutical Science, College of Medicine, National Cheng Kung University, Tainan 701, Taiwan.
E-mail address: yhkao@mail.ncku.edu.tw (Y.-H.K. Yang).

http://dx.doi.org/10.1016/j.ijcard.2015.07.075
0167-5273/

Thus, the objective of this study was to assess the accuracy of mortality data in patients with AMI or stroke in the NHIRD, by linking it with electronic medical records (EMR) from a medical center, using the latter as the gold standard.

2. Methods and materials

2.1. Data sources and record linkage

Datasets from the years 2005 through 2010 of the NHIRD and the EMR from one tertiary medical center in southern Taiwan were obtained for this validation study. This medical center is an academic healthcare center with approximately 1200 beds and an occupancy rate of about 78%. Data available in the EMR upon our request included patients' characteristics (sex, medical ID, date of birth), in-patient admission date, discharge date, diagnosis at discharge (up to 5), procedure (up to 5), status of discharge (including death), and date of death (if death at discharge). Data files available in the NHIRD for researchers were provided by the National Health Research Institute and included the beneficiary registration data and original claims datasets for in-patient and out-patient medical service, expenditures, and detailed prescription data. We extracted data of patients from both the NHIRD and the EMR databases who had primary diagnosis of AMI (ICD-9-CM code: 410xx), hemorrhagic stroke (ICD-9-CM code: 430xx, 431xx, 432xx), or ischemic stroke (ICD-9-CM code: 433xx, 434xx) during the years 2005 to 2010.

Due to our inability to directly match the NHIRD and the EMR by patient-specific identifiers, we used probabilistic record-linkage methods to identify the specific medical records without patient identifiers in previous studies [12]. Probabilistic record-linkage methods have also been used by other researchers to evaluate mortality data in databases without personal identifier [14,15]. In the current study, the date of birth, gender, admission date, and discharge date were used to match records from the two extracted datasets. We considered the EMR as the gold standard in validation studies; therefore, only those NHIRD records that could be matched to the EMR were included for analysis. If there were more than one patient within either NHIRD or EMR who shared the same attributes in all four matching variables, they were excluded. Due to the large number of records available in the NHIRD and EMR databases, we cannot rule out the possibility, however small, that more than one patient might share the same date of birth, gender, admission date, and discharge date. If a patient had multiple hospitalizations for AMI or stroke, only the last episode was included in the study (Fig. 1).

2.2. Validating in-hospital mortality

There are 11 codes for status (or disposition) at discharge in the NHIRD. Two of them in the in-patient dataset indicate either the patient expired during the hospitalization (coded as "4") or is in the acute terminal stage when discharged (coded as "A"). Beside these two codes, the code number "5" indicates that the patient is discharged against medical advice, which mostly happens in Taiwan when there is no more effective way to sustain life and the patient and/or the family wishes to bring patient home for comfort. Thus, this study was to evaluate whether the disposition codes of "4", "A", and "5" in the NHIRD could be indicative of mortality or terminal stage by using the mortality records in the EMR as the gold standard for verification.

Another possible source of identifying mortality status is the disenrollment record from the NHI registration files of beneficiaries. The next of kin or other responsible person needs to terminate an expired person's insurance from the NHI. The insurance information of a person in the registration files of the NHIRD includes: total sum insured, oneself /relative insured, enrollment date, and disenrollment date. [16]. Because NHI enrollment is mandatory for all residents, the three reasons that can allow disenrollment are death, missing more than 6 months, and immigration out of the country. Based on the disenrollment data, we hypothesize that a patient could be expired if he/she was withdrawn from NHI within 3 days of a major medical event, such as hospitalization due to AMI or stroke in the current study. In 1995, the NHI established the Catastrophic Illness Program (CIP) to provide additional financial assistance to patients with 30 specific medical conditions, such as cancer, hemodialysis, and stroke. The CIP also collects patient information in its Catastrophic Illness Registry Data file (CIRD), which includes data of gender, date of birth, application date for CIP, certificate date for CIP, and death. Thus, the third external source for mortality data from the NHIRD, after disposition status at discharge and disenrollment data, is the death information in the CIRD of the NHIRD. For patients with catastrophic disorder cards, both their death certificate and death date would be recorded in the CIRD of the NHIRD. We used three sources of mortality data together from the NHIRD to compare with data from the EMR to estimate the accuracy of the NHIRD data (Fig. 1).

2.3. Agreement in discharge diagnoses

There are up to five discharge diagnoses recorded in the NHIRD, including one primary and four secondary diagnoses. We considered the four secondary diagnoses as comorbidities. We further tabulated the comorbidity diagnosis at discharge in death and in survival cases. The consistency was evaluated of all discharge diagnoses between the EMR and the NHIRD records of the study population.

2.4. Statistical analysis

The matching rate was computed as the number of matched cases divided by the number of cases retrieved from the NHIRD. The validity of in-hospital death data in the NHIRD among the matched cases was estimated by calculating the positive predictive value (PPV), using the EMR as the gold standard. To investigate the reliability of using the disenrollment data of the NHIRD as a source of mortality status, we computed the proportion of disenrollment among patients who had expired as indicated in the EMR. For patients who were enrolled in the CIRD, we estimated the percentage of positive agreement (PPA) of mortality data in both the EMR and the catastrophic illness registry files of the NHIRD among patients with severe disorders. The PPA was calculated as the number of death records in both the EMR and the CIRD divided by the number of death records in either data file. To estimate the agreement in comorbidities at discharge between the two databases, we defined sensitivity as the probability that the diagnoses recorded in the EMR (denominator) were also recorded in the NHIRD (numerator). PPV is the conditional probability that diagnoses recorded in the NHIRD (denominator) were also recorded in the EMR (numerator).

All computations and 95% confidence intervals (CIs) for binominal proportions were performed with SAS version 9.3 (SAS Institute Inc., Cary, NC, USA). This study was reviewed and approved by the Institutional Review Board of the National Cheng Kung University Medical Center (A-ER-102-209).

3. Results

3.1. Study population

A total of 6418 cases (1579 AMI, 1444 hemorrhagic stroke, 3395 ischemic stroke) who met the inclusion criteria were extracted from the NHIRD, while 6414 cases (1610 AMI, 1420 hemorrhagic stroke and 3384 ischemic stroke) were extracted from the EMR. Linkage was achieved in 6197 cases, with a linkage rate of 96.56% (6197/6418), including 1500 patients with AMI, 1373 with hemorrhagic stroke, and 3324 with ischemic stroke. No cases shared the same attributes in all four matching variables in either database (Fig. 1). Among the linked cases, the mean age was 66.0 years ($\pm$14) and 63.0% were male.