

RESEARCH

Open Access



# Validation of diagnostic codes of major clinical outcomes in a National Health Insurance database

Jiesuck Park<sup>1†</sup>, Soonil Kwon<sup>1†</sup> , Eue-Keun Choi<sup>1\*</sup>, You-jung Choi<sup>1</sup>, Euijae Lee<sup>1</sup>, Wonseok Choe<sup>1</sup>, So-Ryoung Lee<sup>1</sup>, Myung-Jin Cha<sup>1</sup>, Woo-Hyun Lim<sup>2</sup> and Seil Oh<sup>1</sup>

## Abstract

**Background and objectives:** The Korean National Health Insurance Service (NHIS) database has been widely used for cardiovascular research. We validated the primary diagnostic codes of major clinical outcomes, including acute myocardial infarction (AMI), gastrointestinal bleeding (GIB), stroke, and intracranial hemorrhage (ICH) used for Korea NHIS claims.

**Subjects and methods:** From 2016 to 2017, 800 patients with primary diagnostic codes of AMI, GIB, stroke, or ICH at discharge were randomly selected from a single tertiary medical center in Korea (200 patients per each diagnosis). The positive predictive value (PPV), sensitivity, and specificity of the primary diagnostic codes were calculated using hospital medical record review as the gold standard. Further improvement in the diagnostic validity of the codes was assessed by combining clinical information such as duration of hospitalization, blood transfusion, brain imaging studies, or prescription records of antithrombotic agents.

**Results:** Among 200 patients with AMI as the primary discharge diagnosis, 184 patients were clinically confirmed (PPV of 92.0%). For GIB, 184 (92.0%) patients with the primary discharge diagnosis were verified to have true GIB events, showing PPV of 92%. For stroke, 181 (90.5%) patients were clinically confirmed with true stroke events. For ICH, 143 (71.5%) patients were verified to be true ICH events. In stroke and ICH, the PPV and specificity improved after combining with the hospitalization duration, imaging studies, and prescription of antithrombotic agents.

**Conclusions:** For major clinical outcomes in the NHIS database, the primary diagnostic codes showed favorable reliability. For stroke and ICH, considerations of relevant clinical information could improve the accuracy of diagnosis.

**Keywords:** Myocardial infarction, Gastrointestinal bleeding, Stroke, Intracranial hemorrhages, Validity

## Introduction

Nationwide population-based analysis is now widely used to evaluate the prevalence of a specific disease and to discover the relationship between new risk factors and clinical outcomes [1, 2]. Moreover, as the analysis targets the whole population, this approach could give a real-world evidence, especially in patients who were not included in well-designed randomized controlled

trials [3]. The National Health Insurance Service (NHIS) is a single-insurer system, which covers the whole population of Korea [4, 5]. The NHIS provides various medical information of the Korean population over the past decades, including inpatient and outpatient records of diagnostic codes, procedures, examinations including imaging studies, medication prescriptions, and demographic information of patients. Recently, public access to the NHIS database has been available, which enabled many researchers to apply the population-based analysis in various study fields including epidemiology, economics, environmental policy and industry and public health and medicine [6–10]. In a population-based analysis,

\*Correspondence: choiek17@snu.ac.kr

<sup>†</sup>Jiesuck Park and Soonil Kwon contributed equally to this work

<sup>1</sup> Department of Internal Medicine, Seoul National University Hospital, 101, Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea

Full list of author information is available at the end of the article



diagnostic codes of the disease based on the International Classification of Diseases (ICD) code is used to identify clinical outcomes. However, in Korea, as the record of the diagnosis is cumulatively stored in the claims database, it is hard to differentiate true clinical events from previous records based on diagnostic codes. In this regard, most of the population-based studies in Korea had focused on primary prevention of the disease by excluding patients with previous diagnosis of the study outcome [6–10]. Previously, several studies had attempted to evaluate the validity of primary diagnostic codes in nationwide claims based on its higher association with true clinical events [11, 12]. However, there is a paucity of reports assessing the validity of the primary diagnostic codes in the NHIS database in Korea. Therefore, we aimed to validate the primary diagnostic codes at discharge in patients with major clinical outcomes including acute myocardial infarction (AMI), gastrointestinal bleeding (GIB), stroke, and intracranial hemorrhage (ICH).

**Methods**

**Data source and study population**

From the database of a single tertiary medical center in Korea (Seoul National University Hospital), patients who had primary diagnostic codes of AMI, GIB, stroke, or ICH at discharge from 2016 to 2017 were randomly selected (200 patients for each diagnosis). The diagnostic codes of clinical outcomes were based on the International Classification of Diseases code, tenth version (ICD-10 code, Table 1). The study protocol of the current study was approved waived for informed consent by the Seoul National University Hospital Institutional Review Board (No. 1805-133-948).

**The validation of diagnosis for acute myocardial infarction and gastrointestinal bleeding**

Among the 200 patients with primary discharge diagnostic codes of AMI (ICD-10 code I21, I22) at discharge, true AMI events were identified based on the universal definition of AMI [13] by reviewing medical records including discharge records electrocardiograms, laboratory results such as cardiac enzymes, procedure records including

percutaneous coronary intervention or coronary artery bypass graft surgery and prescriptions of medications (antiplatelet agents, statin or beta-blockers).

Primary discharge diagnostic codes representing upper and lower GIB were selected (Table 1). Patients with a true GIB event were identified through reviewing discharge records, procedure records including gastroduodenoscopy or colonoscopy with endoscopic hemostasis or transcatheter arterial embolization, the initial serum hemoglobin level or the existence of blood transfusions of red blood cells. The true GIB events were defined as following: (1) definite visual detection of bleeding or clearly documented medical records of bleeding described by physicians; or (2) identification of bleeding sources by endoscopic or imaging studies, which were accordant to the previous study [14]. Moreover, the history of red blood cell (RBC) transfusion was used as the inclusion criteria to evaluate improvement in diagnostic accuracy.

**The validation of diagnosis for stroke and intracranial hemorrhage**

True cases of stroke were verified based on medical records, including discharge records, a duration of hospitalization, brain imaging studies including a computed tomography (CT) scan or a magnetic resonance imaging (MRI), and prescription records of antithrombotic agents. From the previous studies, the median duration of hospital stay was reported for 7.4 days (interquartile range 5.1–11.7) [15]. Therefore, the duration of hospitalization more than 3 days was analyzed whether it improves the accuracy of diagnosis. For antithrombotic agents, any prescription record of the medication at discharge was considered for inclusion criteria in patients with stroke diagnoses.

For the confirmation of patients with true ICH events, we reviewed medical records, including discharge records, a duration of hospitalization, brain CT or MRI, and prescription records of antithrombotic agents. To improve the accuracy, we added no prescriptions of antithrombotic agents at discharge as one of the diagnostic criteria of ICH.

**Table 1 Diagnostic codes (ICD-10) for clinical outcomes**

Clinical outcomes	Diagnostic codes
Myocardial infarction	I21, I22
Gastrointestinal bleeding	K22.6, K25.0, K25.2, K25.4, K25.6, K26.0, K26.2, K26.4, K26.6, K27.0, K27.2, K27.4, K27.6, K28.0, K28.2, K28.4, K28.6, K29.0, K62.5, K92.0, K92.1, K92.2
Ischemic stroke	I63, I64
Intracranial hemorrhage	I60, I61, I62

**Table 2 Validation of diagnostic codes for myocardial infarction**

	Total	True	False	PPV (%)
Hospitalization + primary diagnostic code	200 (100)	184 (92.0)	16 (8.0)	92.0

PPV positive predictive value

**Statistical analysis**

The two investigators (JSP and KSI) reviewed the medical records independently to discriminate true clinical outcomes. In the cases with discordance, agreement by discussion was used to make a final decision. Positive predictive value (PPV), sensitivity, and specificity were evaluated for each diagnostic code of the clinical outcomes. The PPV was calculated using the following formula: the numbers of patients with the true event of study outcome confirmed by medical record review/ the total numbers of patients with ICD-10 codes × 100 (%). The final diagnoses for false positive cases in each clinical outcome were summarized to overview the factors affecting the accuracy. Statistical analyses were performed using SPSS, version 21.0 (SPSS Inc., Windows, Chicago, IL, USA).

**Results**

**Acute myocardial infarction**

The validation result of AMI is given in Table 2. Among the 200 cases, 184 cases (92%) were identified to have true AMI events who had admitted to the hospital for. The rest of 16 (8.0%) cases were those who were admitted for

follow-up elective coronary angiography (*n* = 11, 5.5%), presented with angina (*n* = 2, 1.0%) or had previous history of AMI (*n* = 3, 1.5%) (Additional file 1: Table S1). The PPV of the diagnostic code was 92.0%.

**Gastrointestinal bleeding**

Among the 200 randomly selected cases, 184 patients (92%) were confirmed to be true GIB events (Table 3). The major proportion of false cases were those who had a previous history of GIB (*n* = 9, 4.5%) followed by those with no clinical evidence of GIB (*n* = 4, 2.0%) and bleeding event other than GI tract (3, 1.5%) (Additional file 1: Table S1). When RBC transfusion was combined with diagnostic code, 44 cases (22.0%), with 41 true events, were excluded, and there was no improvement in PPV (91.7%) with sensitivity and specificity of 77.7% and 18.8%, respectively.

**Stroke**

Various combinations of diagnostic criteria for stroke are given in Table 4. The combination of primary diagnostic code with brain imaging studies showed PPV and sensitivity of 92.2% and 91.2%, respectively. By combining prescription records of antithrombotic agents with primary diagnostic codes, PPV and sensitivity were 93.4% and 93.4%, respectively. Among various inclusion criteria, the highest PPV, sensitivity, and specificity were achieved by combining primary diagnostic code with hospitalization over 3 days and prescription records of antithrombotic therapy (PPV 94.3%, sensitivity 90.1%, and specificity 47.4%).

**Table 3 Validation of diagnostic codes for gastrointestinal bleeding**

GI bleeding	Total	True	False	PPV (%)	Sensitivity (%)	Specificity (%)
Hospitalization + primary diagnostic code	200 (100)	184 (92.0)	16 (8.0)	92.0		
Hospitalization + primary diagnostic code + RBC transfusion	156 (100)	143 (91.7)	13 (8.3)	91.7	77.7	18.8

PPV positive predictive value, RBC red blood cell

**Table 4 Validation of diagnostic codes for stroke**

Ischemic stroke	Total	True	False	PPV (%)	Sensitivity (%)	Specificity (%)
Hospitalization + primary diagnostic code	200 (100)	181 (90.5)	19 (9.5)	90.5	–	–
Hospitalization + primary diagnostic code + imaging diagnosis <sup>a</sup>	179 (100)	165 (92.2)	14 (7.8)	92.2	91.2	26.3
Hospitalization + primary diagnostic code + prescription <sup>b</sup>	181 (100)	169 (93.4)	12 (6.6)	93.4	93.4	36.8
Hospitalization (over 3 days) + primary diagnostic code + imaging diagnosis	170 (100)	157 (92.4)	13 (7.6)	92.4	86.7	31.6
Hospitalization (over 3 days) + primary diagnostic code + prescription	173 (100)	163 (94.2)	10 (5.8)	94.2	90.1	47.4

PPV positive predictive value

<sup>a</sup> Brain imaging studies including a computed tomography (CT) scan or a magnetic resonance imaging (MRI)

<sup>b</sup> At least ≥ 1 prescription of antithrombotic agents at discharge

### Intracranial hemorrhage

Table 5 shows the validation results of various inclusion criteria for ICH events. The combination of primary diagnostic code with hospitalization and brain imaging studies showed PPV, sensitivity, and specificity of 81.4%, 95.1%, and 45.6%, respectively. When combining the primary diagnostic code with hospitalization and prescription of antithrombotic therapy, we found that PPV, sensitivity, and specificity were 72.1%, 90.2%, and 12.3%, respectively. By considering the condition for the duration of hospitalization (over 3 days), PPV and specificity showed mild improvement in both criteria.

### Discussion

We have validated that the primary diagnostic codes at discharge of major clinical outcomes comprising AMI, GIB, stroke, and ICH were favorable and reliable, which could be used in nationwide population-based studies based on the NHIS database.

#### Validation of primary diagnostic code in identifying acute myocardial infarction

In our study results, a reliable diagnosis for AMI was shown with PPV 92% based on primary diagnosis codes at discharge. The PPV of AMI in our result is comparable to those previously reported, which ranges from 93 to 96.9%, though the diagnostic codes in these studies were based on ICD-9 [11, 12]. Cheng et al. had reviewed medical records of 281 patients with the primary diagnostic code of AMI and showed PPV of 93% [11]. Also, they compared multiple inclusion criteria by combining medical records of antithrombotic agents or procedures for coronary revascularization. The highest PPV (95%) was found in that comprising primary diagnosis for AMI and percutaneous coronary intervention, although it showed the lowest sensitivity (62%). In the current study, we did not include information of coronary revascularization, such as percutaneous coronary intervention or bypass surgery. As a substantial proportion of patients with AMI would not require coronary revascularization

[16], considering procedure information with diagnostic code in defining AMI may increase false negative rates. Petersen et al. had reported PPV of 96.9% of AMI among 4712 patients with primary diagnostic code for AMI at admission [12]. In his report, of 147 false positive cases, 54% ( $n=82$ ) of them were those only with a previous history of AMI, which were similar to our results. Previously, Kimm et al. had reported validation study for AMI using diagnostic codes based on ICD-10, which showed somewhat a lower diagnostic accuracy of 71.4% compared to our result [17]. In this study, the authors selected patients with any diagnostic code of I21 (not I22) from 1994 to 2007 and investigated medical information of the patients by survey with questionnaires from multiple study centers. Therefore, those with insufficient survey records would be excluded, and higher numbers of false positive cases could be included compared to previous studies using primary diagnostic code, which may result in the relatively lower rate of PPV. However, due to development in electronic medical record system and high sensitivity diagnostic modalities, the accuracy of the diagnostic codes for AMI has been improved recently, which is supported by the recent study from the same research group reporting improved PPV of 85.6% among 164 patients from 2009 to 2011 [18].

#### Validation in gastrointestinal bleeding

In our study results, the primary discharge diagnostic code showed PPV of 82% for true GIB event. Cunningham et al. had reported a validation of bleeding events among 186 warfarin users by using primary diagnostic code during hospitalization [19]. Of the 120 patients who had a primary diagnostic code for GI bleeding, PPV was 85.8%. And, 73.8% of those patients were definite bleeding events. Another study validated bleeding codes comprising intracranial, gastrointestinal, genitourinary, and respiratory bleeding from ICD-9 in 361 warfarin users reporting PPV of 91% [14]. In this study, patients who were hospitalized due to the bleeding event, accompanied with a decrease in hemoglobin level > 2 g/L or had

**Table 5 Validation of diagnostic codes for intracranial hemorrhage**

Intracranial hemorrhage	Total	True	False	PPV (%)	Sensitivity (%)	Specificity (%)
Hospitalization + primary diagnostic code	200 (100)	143 (71.5)	57 (28.5)	71.5	–	–
Hospitalization + primary diagnostic code + imaging diagnosis <sup>a</sup>	167 (100)	136 (81.4)	31 (18.6)	81.4	95.1	45.6
Hospitalization + primary diagnostic code + prescription <sup>b</sup>	179 (100)	129 (72.1)	50 (27.9)	72.1	90.2	12.3
Hospitalization (over 3 days) + primary diagnostic code + imaging diagnosis	153 (100)	127 (83.0)	26 (17.0)	83.0	88.8	54.4
Hospitalization (over 3 days) + primary diagnostic code + prescription	142 (100)	118 (83.1)	24 (16.9)	83.1	82.5	57.9

PPV positive predictive value

<sup>a</sup> Brain imaging studies including a computed tomography (CT) scan or a magnetic resonance imaging (MRI)

<sup>b</sup> No prescription of any antithrombotic agents at discharge

undergone  $\geq 2$  units of RBC transfusion were defined as those with major bleeding events, which showed PPV of 87%. Although hemoglobin level is not available in the NHIS database, history of RBC transfusion could be analyzed in those with GIB events. However, we found that combining information of RBC transfusion with primary diagnosis at discharge did not improve the PPV and resulted in lower sensitivity. Although RBC transfusion related to bleeding events was considered as a major or severe bleeding condition in the previous classification system [20], it could exclude patients with evident GIB events but without indication for blood transfusion. In our result, after combining RBC transfusion with primary diagnostic code and hospitalization, 44 patients were excluded whom 93.2% of them were true events. Also, as most of the false positive cases had RBC transfusion based on a low level of hemoglobin without clear gastrointestinal bleeding focus, information on RBC transfusion cannot discriminate these patients. However, considering that all patients with GIB in our study had hospital admission, the primary diagnostic code of GIB with hospitalization might be sufficient to identify patients with true major GIB.

#### Validation of diagnostic code in stroke and intracranial hemorrhage

Previous number of studies had validated the diagnostic codes for stroke, and ICH with a wide range of PPV depends on different criteria for outcome confirmation [21, 22]. Golomb et al. had validated ICD-9 codes of stroke among 663 children from 1999 to 2004 with outcome confirmation based on radiologic evidence of stroke. In this report, compared to diagnostic code in any position, those in the primary position showed better PPV (88% to 92% for code 436 and 37% to 53% for code 342). Park et al. had reported a validation study among 626 patients with a primary diagnosis for cerebrovascular disease (ICD-10, I60–I68) in medical claims from 1993 to 1997 [23]. The accuracy rate of the diagnostic code for stroke was 83.4% for I63 and 71.4% for I64–I68. Although the medical information of the patients was retrieved by trained record administrator using survey methods from study centers, 18.2% ( $n=114$ ) of patients were excluded for analysis due to incomplete or unavailable medical information, which may result in relatively lower accuracy than our result.

In the current study, we tried to improve the accuracy of the primary diagnostic code by combining relevant clinical information. In both stroke and ICH, a combination of primary discharge diagnostic code, hospitalization at least 3 days, and conditions for prescription of antithrombotic agents showed the highest PPV. However, the specificity was low ranged less than 60% regardless

of combination with brain imaging studies or conditions for prescription of antithrombotic agents (26.3% to 47.4% for stroke and 12.3% to 57.9% for ICH). As for the false positive cases in stroke event, most of them were those with a previous history of stroke or suspected for symptoms which show no evidence of stroke in clinical or any diagnostic modalities. Majority of false positive cases in ICH were those with previous history of ICH and those admitted for unrelated medical issues such as a routine follow-up examination for cerebral angiography. Previously, Park et al. had investigated factors associated with the inaccuracy of diagnostic codes in medical insurance claims [23]. Among them, the most frequent reason of the inaccuracy was to meet review criteria of medical insurance benefits ( $n=45$  of 85 cases, 52.9%) followed by incorrect coding habits ( $n=13$ , 15.3%) and diagnostic error ( $n=9$ , 10.6%), which would explain the relatively higher false positive rate in diagnostic codes for stroke and ICH.

#### Limitations

Several limitations of our study must be clarified. First, as for MI validation, we did not include the prescription records in validation compared to the previous report [11]. However, we concluded that it is sufficient to discriminate the true MI events based on primary diagnosis, and there could be false negative cases excluded by applying strict inclusion criteria. Second, as for GIB, laboratory data of hemoglobin level or hemodynamic condition of the patient such as systolic blood pressure or pulse rate could not be considered as inclusion criteria which were not available in the NHIS database. Third, the specificity of the inclusion criteria in GIB, stroke, and ICH was relatively low. For GIB and stroke, the primary diagnostic code demonstrated favorable performance in PPV (92.0% and 90.5% for GIB and stroke, respectively), and the number of false positive cases was small (16 and 19 for GIB and stroke, respectively). As most of those cases were included to the additional criteria (i.e., RBC transfusion, brain imaging studies, or prescription records of antithrombotic therapy) irrelevant to the true diagnosis, the specificity was low. For ICH, however, the primary diagnostic code showed moderate accuracy (PPV 71.5%). Therefore, the additional inclusion criteria improved both PPV and specificity. These results suggest that primary diagnostic code would be sufficient to identify true patients correctly, and supplementary clinical information could be considered in case higher PPV is pursued. Fourth, this study was performed in a tertiary medical center, and the definition was not validated in primary and secondary hospitals, so further study including primary and secondary hospital would be needed. Finally, this study could not evaluate the patients with false

negative. We only analyzed the patients with the primary diagnostic code of clinical outcomes, but it is also important to assess the patients without primary diagnostic code but actual clinical events.

## Conclusion

Primary discharge diagnostic code showed favorable reliability in AMI and GIB which showed high PPV in those with primary diagnostic code at discharge and hospitalization. Stroke and ICH showed the highest accuracy in a combination of primary diagnostic code, hospitalization, and prescription. Primary discharge diagnostic code of AMI, GIB, stroke, and ICH could be used for outcome study using the NHIS database.

## Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s42444-019-0005-0>.

**Additional file 1: Table S1.** The final diagnoses of false-positive cases.

## Abbreviations

AMI: acute myocardial infarction; CT: computed tomography; GIB: gastrointestinal bleeding; ICD: International Classification of Diseases; ICH: intracranial hemorrhage; MRI: magnetic resonance imaging; NHIS: the National Health Insurance Service; PPV: positive predictive value; RBC: red blood cell.

## Acknowledgements

None.

## Authors' contributions

JP and SK involved in writing draft and analyzing data. E-KC was involved in writing draft and creating concept of study. YC, EL, WC, S-RL, M-JC, W-HL and SO involved in data review and writing draft. All authors read and approved the final manuscript.

## Funding

This study was supported by Korean Heart Rhythm Society (KHRS 2016-3).

## Availability of data and materials

No agreement on data release.

## Ethics approval and consent to participate

The study protocol of the current study was approved and waived for informed consent by the Seoul National University Hospital Institutional Review Board (No. 1805-133-948).

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup> Department of Internal Medicine, Seoul National University Hospital, 101, Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea. <sup>2</sup> Department of Internal Medicine, Seoul National University Boramae Medical Center, Seoul, Republic of Korea.

Received: 5 April 2019 Accepted: 30 April 2019

Published online: 20 November 2019

## References

- Lee SR, Choi EK, Han KD, Cha MJ, Oh S. Trends in the incidence and prevalence of atrial fibrillation and estimated thromboembolic risk using the CHA2DS2-VASc score in the entire Korean population. *Int J Cardiol*. 2017;236:226–31.
- Lee S-R, Choi E-K, Han K-D, Cha M-J, Oh S. Prevalence of non-valvular atrial fibrillation based on geographical distribution and socioeconomic status in the entire Korean population. *Korean Circ J*. 2018a;48:622–34.
- Cha MJ, Choi EK, Han KD, Lee SR, Lim WH, Oh S, Lip GH. Effectiveness and safety of non-vitamin K antagonist oral anticoagulants in asian patients with atrial fibrillation. *Stroke*. 2017;48:3040–8.
- Song SO, Jung CH, Song YD, Park CY, Kwon HS, Cha BS, Park JY, Lee KU, Ko KS, Lee BW. Background and data configuration process of a nationwide population-based study using the Korean national health insurance system. *Diabetes Metab J*. 2014;38:395–403.
- Lee YH, Han K, Ko SH, Ko KS, Lee KU, Taskforce Team of Diabetes Fact Sheet of the Korean Diabetes A. Data analytic process of a nationwide population-based study using national health information database established by national health insurance service. *Diabetes Metab J*. 2016a;40:79–82.
- Lee SR, Choi EK, Rhee TM, Lee HJ, Lim WH, Kang SH, Han KD, Cha MJ, Cho Y, Oh IY, Oh S. Evaluation of the association between diabetic retinopathy and the incidence of atrial fibrillation: a nationwide population-based study. *Int J Cardiol*. 2016b;223:953–7.
- Kang SH, Choi EK, Han KD, Lee SR, Lim WH, Cha MJ, Cho Y, Oh IY, Oh S. Underweight is a risk factor for atrial fibrillation: a nationwide population-based study. *Int J Cardiol*. 2016;215:449–56.
- Lee HJ, Choi EK, Lee SH, Kim YJ, Han KD, Oh S. Risk of ischemic stroke in metabolically healthy obesity: a nationwide population-based study. *PLoS ONE*. 2018b;13:e0195210.
- Lee S-R, Choi E-K. Prevalence of atrial fibrillation in Korean population. *Int J Arrhythm*. 2017;18:195–204.
- Kim J, Lee J, Shin CM, Lee DH, Park BJ. Risk of gastrointestinal bleeding and cardiovascular events due to NSAIDs in the diabetic elderly population. *BMJ Open Diabetes Res Care*. 2015a;3:e000133.
- Cheng C-L, Lee C-H, Chen P-S, Li Y-H, Lin S-J, Yang Y-HK. Validation of acute myocardial infarction cases in the national health insurance research database in Taiwan. *J Epidemiol*. 2014;24:500–7.
- Petersen LA, Wright S, Normand SL, Daley J. Positive predictive value of the diagnosis of acute myocardial infarction in an administrative database. *J Gen Intern Med*. 1999;14:555–8.
- Thygesen K, Alpert JS, Jaffe AS, Simoons ML, Chaitman BR, White HD, Joint ESCAAHAWHFTftUDoMI, Katus HA, Lindahl B, Morrow DA, Clemmensen PM, Johanson P, Hod H, Underwood R, Bax JJ, Bonow RO, Pinto F, Gibbons RJ, Fox KA, Atar D, Newby LK, Galvani M, Hamm CW, Uretsky BF, Steg PG, Wijns W, Bassand JP, Menasché P, Ravkilde J, Ohman EM, Antman EM, Wallentin LC, Armstrong PW, Simoons ML, Januzzi JL, Nieminen MS, Gheorghiade M, Filippatos G, Luepker RV, Fortmann SP, Rosamond WD, Levy D, Wood D, Smith SC, Hu D, Lopez-Leydon JL, Robertson RM, Weaver D, Tendera M, Bove AA, Parkhomenko AN, Vasilieva EJ, Mendis S. Third universal definition of myocardial infarction. *Circulation*. 2012;126:2020–35.
- Arnason T, Wells PS, van Walraven C, Forster AJ. Accuracy of coding for possible warfarin complications in hospital discharge abstracts. *Thromb Res*. 2006;118:253–62.
- Kim BJ, Park JM, Kang K, Lee SJ, Ko Y, Kim JG, Cha JK, Kim DH, Nah HW, Han MK, Park TH, Park SS, Lee KB, Lee J, Hong KS, Cho YJ, Lee BC, Yu KH, Oh MS, Kim DE, Ryu WS, Cho KH, Kim JT, Choi JC, Kim WJ, Shin DI, Yeo MJ, Sohn SI, Hong JH, Lee J, Lee JS, Yoon BW, Bae HJ. Case characteristics, hyperacute treatment, and outcome information from the clinical research center for stroke-fifth division registry in South Korea. *J Stroke*. 2015b;17:38–53.
- McManus DD, Gore J, Yarzebski J, Spencer F, Lessard D, Goldberg RJ. Recent trends in the incidence, treatment, and outcomes of patients with STEMI and NSTEMI. *Am J Med*. 2011;124:40–7.
- Kimm H, Yun JE, Lee SH, Jang Y, Jee SH. Validity of the diagnosis of acute myocardial infarction in Korean national medical health insurance claims data: the Korean heart study (1). *Korean Circ J*. 2012;42:10–5.
- Kim HY. Validity of the diagnosis of acute myocardial infarction in Korean national medical health insurance claims data: the Korean Heart Study [Master thesis]. Yonsei University, Seoul, Korea, June 2013.

19. Cunningham A, Stein CM, Chung CP, Daugherty JR, Smalley WE, Ray WA. An automated database case definition for serious bleeding related to oral anticoagulant use. *Pharmacoepidemiol Drug Saf.* 2011;20:560–6.
20. Steg PG, Huber K, Andreotti F, Arnesen H, Atar D, Badimon L, Bassand JP, De Caterina R, Eikelboom JA, Gulba D, Hamon M, Helft G, Fox KA, Kristensen SD, Rao SV, Verheugt FW, Widimsky P, Zeymer U, Collet JP. Bleeding in acute coronary syndromes and percutaneous coronary interventions: position paper by the Working Group on Thrombosis of the European Society of Cardiology. *Eur Heart J.* 2011;32:1854–64.
21. Andrade SE, Harrold LR, Tjia J, Cutrona SL, Saczynski JS, Dodd KS, Goldberg RJ, Gurwitz JH. A systematic review of validated methods for identifying cerebrovascular accident or transient ischemic attack using administrative data. *Pharmacoepidemiol Drug Saf.* 2012;21(Suppl 1):100–28.
22. Krarup LH, Boysen G, Janjua H, Prescott E, Truelsen T. Validity of stroke diagnoses in a national register of patients. *Neuroepidemiology.* 2007;28:150–4.
23. Park JK, Kim KS, Kim CB, Lee TY, Lee KS, Lee DH, Lee S, Jee SH, Suh I, Koh KW, Ryu SY, Park KH, Park W, Wang S, Lee H, Chae Y, Hong H, Suh JS. The accuracy of ICD codes for cerebrovascular diseases in medical insurance claims. *Korean J Prev Med.* 2000;33:76–82.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

