Original article

Clinical and Administrative Data on the Research of Acute Coronary Syndrome in Spain. Minimum Basic Data Set Validity

José Luis Bernal,[a,b,*] José A. Barrabés,[c] Andrés Íñiguez,[d] Antonio Fernández-Ortiz,[e,f] Cristina Fernández-Pérez,[b,g] Alfredo Bardají,[h] and Francisco Javier Elola[b,i]

[a] *Servicio de Control de Gestión, Hospital Universitario 12 de Octubre, Madrid, Spain*
[b] *Fundación Instituto para la Mejora de la Asistencia Sanitaria, Madrid, Spain*
[c] *Servicio de Cardiología, Hospital Universitari Vall d'Hebron, Vall d'Hebron, Institut de Recerca (VHIR), CIBER-CV, Universidad Autónoma de Barcelona, Barcelona, Spain*
[d] *Servicio de Cardiología, Hospital Álvaro Cunqueiro, Vigo, Pontevedra, Spain*
[e] *Servicio de Cardiología, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria San Carlos (IdISSC), Universidad Complutense, Madrid, Spain*
[f] *Fundación Interhospitalaria de Investigación Cardiovascular, Madrid, Spain*
[g] *Servicio de Medicina Preventiva, Instituto de Investigación Sanitaria San Carlos (IdISSC), Universidad Complutense de Madrid, Madrid, Spain*
[h] *Servicio de Cardiología, Hospital Universitario de Tarragona Joan XXIII, Institut d'Investigació Sanitària Pere Virgili (IISPV), Universidad Rovira Virgili, Tarragona, Spain*
[i] *Elola Consultores, Madrid, Spain*

A B S T R A C T

Introduction and objectives: Health outcomes research is done from clinical registries or administrative databases. The aim of this work was to evaluate the concordance of the Minimum Basic Data Set (MBDS) with the DIOCLES (*Descripción de la Cardiopatía Isquémica en el Territorio Español*) registry and to analyze the implications of use of the MBDS in the study of acute coronary syndrome in Spain.
Methods: Through indirect identifiers, DIOCLES was linked with MBDS and unique matches were selected. Some of most relevant variables for risk adjustment of in-hospital mortality due to acute myocardial infarction were considered. Kappa coefficient was used to evaluate the concordance; sensitivity, specificity and positive and negative predictive values to measure the validity of the MBDS, and the area under ROC (receiver operating characteristic) curve to calculate its discrimination. The results were compared among hospitals quintiles according to their contribution to DIOCLES. The influence of unmatched episodes on results was assessed by a sensitivity analysis, using looser linking criteria.
Results: Overall, 1539 (60.85%) unique matches were achieved. The prevalence was higher in DIOCLES (acute myocardial infarction: 71.09%; Killip 3-4: 9.17%; cerebrovascular accident: 0.97%; thrombolysis: 8.64%; angioplasty: 61.92% and coronary bypass: 1.75%) than in the MBDS ($P < .001$). The agreement level observed was almost perfect ($\kappa = 0.863$). The MBDS showed a sensitivity of 85.10% and a specificity of 98.31%. Most results were confirmed by using sensitivity analysis (79.95% episodes matched).
Conclusions: The MBDS can be a useful tool for outcomes research of acute coronary syndrome in Spain. The contrast of DIOCLES and MBDS with medical records could verify their validity.

© 2018 Sociedad Española de Cardiología. Published by Elsevier España, S.L.U. All rights reserved.

Datos clínicos y administrativos en la investigación de resultados del síndrome coronario agudo en España. Validez del Conjunto Mínimo Básico de Datos

R E S U M E N

Introducción y objetivos: La investigación de resultados en salud utiliza tanto registros clínicos como bases de datos administrativas. El objetivo de este trabajo es evaluar la concordancia del Conjunto Mínimo Básico de Datos (CMBD) con el registro DIOCLES (Descripción de la Cardiopatía Isquémica en el Territorio Español) y su utilidad en la investigación de resultados del síndrome coronario agudo en España.
Métodos: Mediante identificadores indirectos, se vinculó el DIOCLES con el CMBD y se seleccionaron los emparejamientos únicos. Considerando algunas de las variables más relevantes para ajustar por riesgo la mortalidad intrahospitalaria por infarto agudo de miocardio, se calculó la concordancia interobservadores, la sensibilidad, la especificidad y los valores predictivos positivo y negativo para medir la validez del CMBD, y el área bajo la curva ROC (*receiver operating characteristic*) para determinar su discriminación. Los resultados se compararon entre quintiles de hospitales según su contribución

---

a DIOCLES. El impacto de los emparejamientos fallidos se evaluó mediante un análisis de sensibilidad con criterios de vinculación más laxos.

Resultados: Se lograron 1.539 (60,85%) emparejamientos únicos. Entre los episodios emparejados, la prevalencia fue mayor en el DIOCLES (infarto agudo de miocardio, el 71,09%; Killip 3-4, el 9,17%; accidente cerebrovascular, el 0,97%; trombolisis, el 8,64%; angioplastia, el 61,92%, y *bypass*, el 1,75%) que en el CMBD (p < 0,001). El acuerdo observado fue casi perfecto (κ = 0,863); la sensibilidad del CMBD resultó del 85,10% y su especificidad, del 98,31%. El análisis de sensibilidad (el 79,95% de emparejamientos) confirmó, en general, estos resultados.

Conclusiones: El CMBD puede ser un instrumento útil para la investigación de resultados del síndrome coronario agudo en España. El contraste de DIOCLES y CMBD con las historias clínicas podría verificar su validez.

Abbreviations

ACS: acute coronary syndrome
AMI: acute myocardial infarction
CRs: clinical registries
MBDS: Minimum Basic Data Set

INTRODUCTION

Cardiovascular disease causes more than 4 million deaths annually in Europe, mostly due to coronary heart disease,[1] and although the mortality rate from ischemic heart disease has declined in recent decades in developed countries, it is still responsible for about one-third of all deaths in the population older than 35 years.[2] In Spain, a considerable increase has been predicted in the incidence of acute coronary syndrome (ACS) over the next 35 to 40 years in parallel with population aging.[3]

Because the prevalence of ACS and the costs involved in its care represent a substantial economic and health care burden, it is important to evaluate clinical practice outcomes and investigate the factors involved. In Spain, the study of ACS has been addressed using clinical registries (CRs)[4–9] and administrative databases,[10–14] mainly the Minimum Basic Data Set (MBDS) of the Spanish National Health System.[15] However, the concordance between the 2 sources has not been examined.

The advantages and limitations of the use of administrative databases to evaluate health outcomes have been extensively analyzed.[16–18] These databases are relatively easy to obtain,[19] significantly less expensive than collection of primary data via medical record review or registries specifically designed and maintained for secondary uses,[20] which can suffer from systematic biases due to irregular patient inclusion,[21] and offer long-term uniform information on large populations.[22] However, their quality depends on the accuracy of diagnostic and procedural coding, usually performed via discharge reports,[23] and different types of biases can compromise their use: some do not allow a distinction between complications and comorbidities[24] or between chronic and acute diseases, and others do not include relevant clinical information, such as medications administered or the results of laboratory tests[25]; nonetheless, in contrast, numerous studies have confirmed their reliability[26–28] and others have not found significant differences in risk-adjusted rates of mortality and readmissions due to acute myocardial infarction (AMI) or heart failure upon comparison of the results obtained with administrative databases and CRs.[29]

Although both sources are subject to strict anonymization rules and do not include direct patient identifiers, robust linkage methods have been developed using indirect identifiers, which allow the records corresponding to the same episode to be correlated in each data source.[30–33] Using procedures of this nature, the aim of this study was to evaluate the agreement between the DIOCLES (*Descripción de la Cardiopatía Isquémica en el Territorio Español* [Description of Ischemic Heart Disease in the Spanish Territory]) registry, the most recent large registry performed in Spain that used random hospital selection to study the in-hospital and 6-month mortality of patients admitted for suspected ACS and to characterize their management,[9] and the MBDS, in order to analyze the implications of the use of CRs and administrative databases in the study of ACS in Spain.

METHODS

Data Sources

The DIOCLES is a multicenter, cross-sectional, quality controlled, observational study conducted in the first half of 2012, with the participation of 44 centers from 13 autonomous communities in Spain. The study prospectively collected data from patients aged ≥ 18 years consecutively admitted with suspected ACS. Demographic variables, risk factors, medical history, clinical presentation, complications, in-hospital mortality, and prehospital, in-hospital, and discharge management were recorded. At 6 months, the occurrence of death and its date and cause were determined by telephone interview. The study population comprised 2557 patients out of the 3059 evaluated, after 502 were excluded for various reasons.[9]

The 2012 MBDS contains 400 861 hospitalization episodes recorded in hospitals of the Spanish National Health System with a primary diagnosis of cardiovascular diseases or without that diagnosis and discharged from a cardiology or cardiac surgery department. The MBDS includes information on patients' demographic characteristics and on variables related to the health care and clinical process, to patients' diseases and conditions, and to the procedures performed during their care, coded by the International Classification of Diseases, Ninth Revision, Clinical Modification.[15]

Linkage Procedure

The DIOCLES registry and the MBDS differ in their scope–DIOCLES includes 6-month follow-up after discharge, whereas the MBDS includes only the hospitalization episode–and in their data models. Given that these models lack common attributes that allow the records corresponding to the same episode to be unequivocally matched (direct identifiers), different indirect identifiers were tested to link them.[32] The identifier found to obtain the most unique matches was selected. This model comprised the hospital code, dates of admission and discharge, and patient age and sex, because date of birth was not recorded in DIOCLES.

Statistical Analysis

A descriptive analysis was performed of the variables studied; discrete variables are expressed as No. (%), and quantitative variables, whose normality was assessed by the Shapiro-Wilk test, as the mean ± standard deviation. To compare discrete variables, the chi-square test or Fisher exact test was applied, as necessary, with the Student *t* test or Mann-Whitney *U* test used for quantitative variables.

The Cohen kappa coefficient was used to evaluate interobserver agreement. Sensitivity, specificity, and positive and negative predictive values were calculated to measure the validity of the MBDS. Among the variables common to the 2 sources, this analysis included some of those most relevant for the adjustment of the risk of hospital mortality due to AMI[12] (Table 1). To estimate the discrimination power of the MBDS, considering the values of the analyzed variables together, the area under the ROC (receiver operating characteristic) curve was used; the homogeneity of κ and of the area under the ROC curve were examined with the chi-square test. The areas under the ROC curve were compared by hospitals grouped into quintiles according to the number of patients included in DIOCLES. The dependence between observations in the same patient was analyzed by the intraclass correlation coefficient (ρ) and the variance inflation factor.[34]

The degree of agreement was interpreted according to the Landis and Koch scale.[35] Statistical significance was defined by *P* < .05 and analyses were performed with Epidat v3.1 and STATA 13.0.

Sensitivity Analysis

The impact of match failures on the results was evaluated through a sensitivity analysis by allowing in the linkage differences of up to 2 days in the admission and discharge dates between DIOCLES and the MBDS and by selecting, among multiple matches, the episodes showing fewer discrepancies among the conditions studied.

In addition, as an alternative linkage procedure, MBDS episodes with ACS as the primary diagnosis and an admission date in the first half of 2012 were selected, and the hospital code, age, sex, and dates of admission and discharge were used as indirect identifiers to match these episodes with those recorded in DIOCLES. On the results obtained through this procedure, a sensitivity analysis was also carried out following the same criteria as in the original procedure.

RESULTS

Of the 2557 patients registered in DIOCLES, 28 were excluded (1.1%), 17 due to a lack of at least 1 of the variables used for the linkage and 11 corresponding to a private hospital not included in the MBDS of the Spanish National Health System. Of the remaining 2529 patients, 1539 (60.85%) were unequivocally matched (Figure 1). These patients constituted the study population, had an average age of 66.95 ± 13.03 years, and comprised 375 women (24.4%).

Matched DIOCLES records showed a higher proportion of AMI and angioplasties and a lower proportion of coronary bypass than nonmatched records, and no significant differences were found in age, the proportion of women, or the other conditions studied. Among the matched episodes, the prevalence of all conditions was significantly higher in DIOCLES than in the MBDS (Table 1).

Overall, the degree of agreement between the 2 data sources was almost perfect (κ = 0.863), with 8766 matches observed (94.93%), and ranged from moderate for worst Killip (classes 3 and 4), substantial for stroke, thrombolysis, and angioplasty, and almost perfect for AMI and coronary bypass. The MBDS, considering all of the variables, showed a sensitivity of 85.10% and a specificity of 98.31%, with a positive predictive value of 94.55% and a negative predictive value of 95.05% (Table 2); there was no dependence between observations in the same patient (ρ ≤ 0.04, variance inflation factor ≤ 1.06). The 95% confidence intervals of the crude general estimates and those adjusted by the methods used (ρ and variance inflation factor) are shown in Table 1 of the supplementary material.

According to their participation in DIOCLES, the cutoff points for hospital grouping were as follows: first quintile = 41 patients; second quintile = 64; third quintile = 80; and fourth quintile = 100. The area under the ROC curve was significantly different by quintile (*P* = .007); the highest corresponded to quintiles 5 and 3 and the lowest to quintiles 1 and 4 (Figure 2).

In the original sensitivity analysis (Table 3), 2022 unique matches (79.95%) were obtained, selected from 2352 coincidences, and there were no significant differences in the interobserver agreement for the conditions analyzed vs the results obtained with

Table 1
Diagnoses and Procedures Analyzed. Prevalence According to the Linkage Result and Data Source

Diagnoses and procedures	ICD-9-CM codes	DIOCLES		*P*	MBDS (matched)	*P*
		Matched	Unmatched			
Age, y		66.95 ± 12.99	66.58 ± 13.03	.486		
Women		375 (24.37)	271 (27.37)	.099		
AMI	410.*1	1094 (71.09)	652 (65.86)	.006	1038 (67.45)	.032
Worse Killip (3 or 4)	427.41, 427.42, 427.5, 518.4, 518.5, 518.51, 518.52, 518.53, 518.81, 518.82, 518.83, 518.84,785.50, 785.51, 798.0, 798.1, 798.2, 798.9, 799.01, 799.02, 998.01	141 (9.17)	99 (10.00)	.527	110 (7.15)	< .001
Stroke	094.87, 430, 431, 432.0, 432.1, 432.9,433.01, 433.11, 433.21, 433.31, 433.81, 433.91, 434.01, 434.11, 434.91, 436	15 (0.97)	11 (1.11)	.896	12 (0.78)	< .001
Thrombolysis	V45.88, 99.10	133 (8.64)	89 (8.99)	.818	106 (6.89)	< .001
Angioplasty	00.66, 36.01, 36.02, 36.05, 36.06, 36.07	953 (61.92)	542 (54.75)	.004	836 (54.32)	< .001
Coronary bypass	36.1*	27 (1.75)	40 (4.04)	.008	25 (1.62)	< .001
Episodes, No.		1539	990			

AMI, acute myocardial infarction; DIOCLES, *Descripción de la Cardiopatía Isquémica en el Territorio Español* (Description of Ischemic Heart Disease in the Spanish Territory); ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification; MBDS, Minimum Basic Data Set; Worse Killip (3 or 4), severe heart failure with acute pulmonary edema or cardiogenic shock during admission.
Unless otherwise indicated, the data represent No. (%) or mean ± standard deviation.
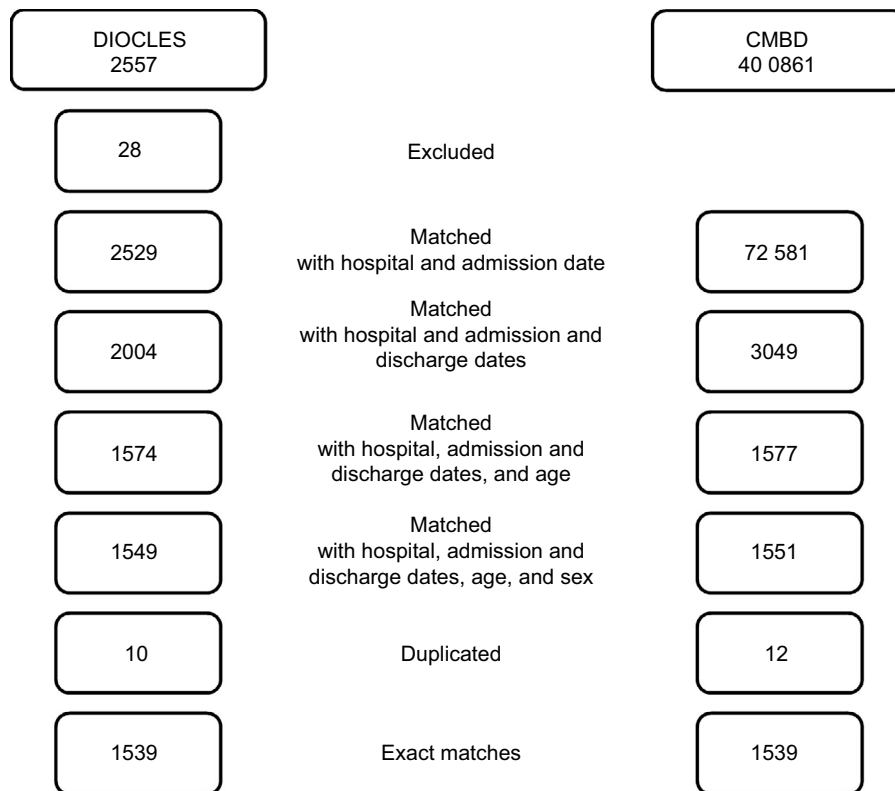
Figure 1. Linkage results for DIOCLES and the MBDS. Number of episodes excluded and matched in each phase of the linkage process. DIOCLES, *Descripción de la Cardiopatía Isquémica en el Territorio Español* (Description of Ischemic Heart Disease in the Spanish Territory); MBDS, Minimum Basic Data Set.

the original linkage procedure, except for AMI and angioplasty, or in either their prevalence in DIOCLES or in the MBDS, except also for AMI and angioplasty (Table 4).

The results of the alternative linkage procedure and its sensitivity analysis are shown in Table 2 of the supplementary material and in Table 3 of the supplementary material.

DISCUSSION

Our results indicate an almost perfect agreement between DIOCLES and the MBDS and a high ability of the latter to discriminate episodes with the same diagnoses and procedures as the corresponding DIOCLES record (area under the ROC curve > 0.9). However, significant differences were found when hospitals were grouped by quintiles according to their contribution to DIOCLES, indicating an inverse association between registry quality and patient volume.

The concordance was greater than that published by Ribera et al.[36] for the diagnoses and procedures included in the ARCA study (evaluation of the risk of coronary surgery in Catalonia), which varied widely from cardiogenic shock (κ = 0.16) to the use of extracorporeal circulation (κ = 0.79),[37] the findings of Cavero-Carbonell et al.[38] for the identification of congenital anomalies (κ = 0.70), and those of Hernández Medrano et al.[39] for cerebrovascular disease (concordance rate, 81.87%); the last 2 examples used medical record review as the reference model. The prevalence was higher in DIOCLES than in the MBDS for all the conditions analyzed, and a κ < 0.7 was only found for the worst Killip classes, which include cardiogenic shock. This finding is similar to that of Lambert et al.[40] in their evaluation of the accuracy of cardiovascular disease coding in a Canadian administrative database vs medical records, although their result (κ = 0.667) showed higher agreement than ours.

The use of administrative databases to investigate health outcomes offers important advantages, but their usefulness

Table 2

the DIOCLES (*Descripción de la Cardiopatía Isquémica en el Territorio Español* [Description of Ischemic Heart Disease in the Spanish Territory]) vs the Minimum Basic Data Set

Diagnoses and procedures	κ (95%CI)	Sensitivity, % (95%CI)	Specificity, % (95%CI)	PPV, % (95%CI)	NPV, % (95%CI)
AMI	0.814 (0.783-0.846)	91.86 (90.20-93.53)	92.58 (90.04-95.13)	96.82 (95.71-97.94)	82.24 (78.79-85.68)
Worse Killip (3 or 4)	0.45 (0.37-0.53)	43.97 (35.42-52.52)	96.57 (95.58-97.56)	56.36 (46.64-66.09)	94.47 (93.25-95.69)
Stroke	0.739 (0.552-0.925)	66.67 (39.48-93.86)	99.87 (99.65-100.00)	83.33 (58.08-100.00)	99.67 (99.35-99.99)
Thrombolysis	0.742 (0.678-0.806)	68.42 (60.15-76.70)	98.93 (98.36-99.51)	85.85 (78.74-92.96)	97.07 (96.16-97.98)
Angioplasty	0.8 (0.77-0.83)	85.94 (83.68-88.20)	97.10 (95.65-98.54)	97.97 (96.95-98.98)	80.94 (77.96-83.91)
Coronary bypass	0.922 (0.845-0.998)	88.89 (75.18-100.00)	99.93 (99.77-100.00)	96.00 (86.32-100.00)	99.80 (99.54-100.00)
Total (all 6 conditions)	0.863 (0.850-0.875)	85.10 (83.64-86.56)	98.31 (98.00-98.62)	94.55 (93.56-95.53)	95.05 (94.54-95.56)

95%CI, 95% confidence interval; AMI, acute myocardial infarction; NPV, negative predictive value; PPV, positive predictive value; Worse Killip (3 or 4), severe heart failure with acute pulmonary edema or cardiogenic shock during admission.
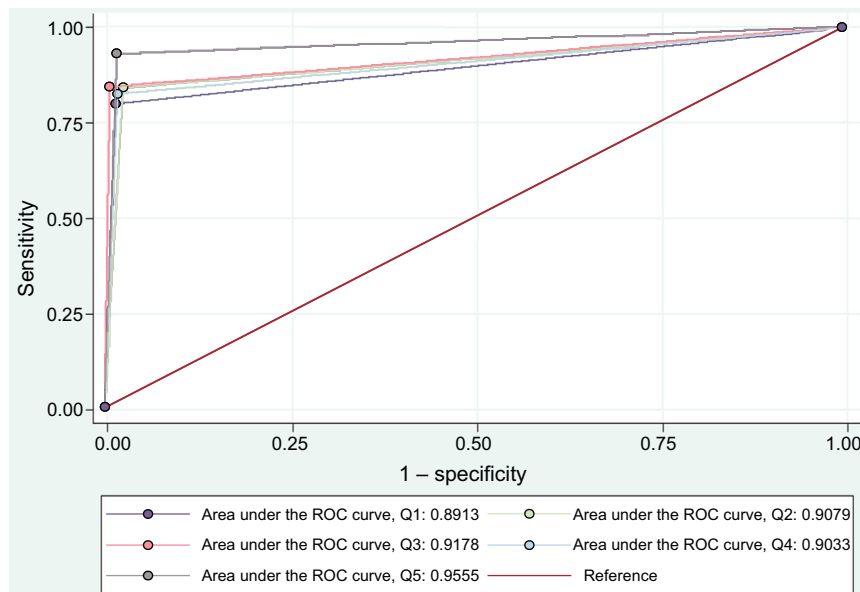
Figure 2. Comparison of the areas under the ROC (receiver operating characteristic) curve by hospital quintile (Q) according to the number of patients registered in DIOCLES (*Descripción de la Cardiopatía Isquémica en el Territorio Español* [Description of Ischemic Heart Disease in the Spanish Territory]).

Table 3

Sensitivity Analysis. Comparison of the Diagnoses and Procedures Recorded in DIOCLES (*Descripción de la Cardiopatía Isquémica en el Territorio Español* [Description of Ischemic Heart Disease in the Spanish Territory]) vs the Minimum Basic Data Set

Diagnoses and procedures	κ (95%CI)	Sensitivity, % (95%CI)	Specificity, % (95%CI)	PPV, % (95%CI)	NPV, % (95%CI)
AMI	0.740 (0.709-0.770)	86.29 (84.46-88.11)	93.17 (91.06-95.27)	96.77 (95.75-97.78)	74.14 (70.95-77.33)
Worse Killip (3 or 4)	0.45 (0.37-0.53)	41.62 (34.25-48.99)	97.01 (96.20-97.81)	58.33 (49.94-67.12)	94.29 (93.21-95.36)
Stroke	0.659 (0.473-0.846)	60.00 (36.03-83.97)	99.85 (99.66-100.00)	80.00 (56.42-100.00)	99.60 (99.30-99.90)
Thrombolysis	0.686 (0.624-0.747)	60.45 (52.97-67.94)	99.02 (98.55-99.50)	85.60 (79.05-92.15)	96.31 (95.44-97.18)
Angioplasty	0.753 (0.725-0.781)	81.56 (79.34-83.77)	97.01 (95.77-98.25)	97.64 (96.66-98.62)	77.57 (74.94-80.20)
Coronary bypass	0.871 (0.788-0.954)	81.58 (67.94-95.22)	99.90 (99.73-100.00)	93.94 (84.28-100.00)	99.65 (99.36-99.93)
Total (all 6 conditions)	0.826 (0.814-83.79)	79.98 (78.55-81.41)	98.42 (98.16-98.69)	94.48 (93.58-95.38)	93.57 (93.08-94.07)

95%CI, 95% confidence interval; AMI, acute myocardial infarction; NPV, negative predictive value; PPV, positive predictive value; Worse Killip (3 or 4), severe heart failure with acute pulmonary edema or cardiogenic shock during admission.

Table 4

Results of the Comparison of Concordances and Prevalences Between DIOCLES and the MBDS

Diagnoses and procedures	κ coefficient, *P*	Coefficient of prevalence in the DIOCLES, *P*	Coefficient of prevalence in the MBDS, *P*
AMI	.009	.649	.004
Worse Killip (3 or 4)	.907	.99	.467
Stroke	.558	.965	.897
Thrombolysis	.218	.907	.397
Angioplasty	.023	.336	.02
Coronary bypass	.379	.783	.986

AMI, acute myocardial infarction; DIOCLES, *Descripción de la Cardiopatía Isquémica en el Territorio Español* (Description of Ischemic Heart Disease in the Spanish Territory); MBDS, Minimum Basic Data Set; Worse Killip (3 or 4), severe heart failure with acute pulmonary edema or cardiogenic shock during admission.

ultimately depends on the clinical validity and consistency of the data, which, as far as we know, have not previously been analyzed in relation to the study of ACS in Spain. The observed variability indicates that certain conditions (AMI, angioplasty, and coronary bypass) were coded with greater accuracy than others. Thus, there seems to be room for improvement in hospital coding and, consequently, in the quality of the MBDS, although the lack of concordance may also result from registration errors in DIOCLES.[21]

Limitations

More than 85% of the diagnoses and procedures analyzed had a probability of being true positives if present in the MBDS and more than 98% had a probability of being true negatives if not. Thus, the present work indicates a threshold for the validity of the MBDS as a data source for the study of ACS in Spain, although its main limitation is the percentage of matches that could not be resolved.

Because the use of alternative linkage procedures did not reduce the percentage of failed matches, which ranged between 55% and 80%, the difficulty of identifying the episodes included in DIOCLES in the MBDS must be due to registration inaccuracies in either of the 2 sources. To our knowledge, the inclusion of age instead of date of birth in DIOCLES is a significant factor, but there may be others.

Dissociation is a legal requirement for personal data protection that prevents the use of direct identifiers to link CRs and administrative databases. To resolve this restriction, indirect procedures have been developed that achieved unique matches in proportions ranging from 58.1% (Setoguchi et al.[33]) to 87.5% (Pasquali et al.[31]), 87.9% (Austin et al.[30]), and 90.8% (Hammill et al.[32]). The reasons for the match failures[32] have been described; although most are determined by the study context and cannot be applied to our situation, it is likely that, in DIOCLES, the date of patient arrival at or release from the emergency department or cardiology department was sometimes recorded as the date of patient admission or discharge, which might not coincide with that of the MBDS (which reflects when the patient is admitted to the hospital and occupies a hospital bed[41]). The reasons for the discrepancies found by Sarkies et al.[42] between the use of administrative data and observational registries for the calculation of the average stay per hospitalization episode would support this explanation, consistent with the increase in the percentage of matches to practically 80% when registry differences of up to 2 days are allowed for the dates of admission and discharge.

In this case, the sensitivity analysis revealed that an increase in the number of cases by almost 32% did not significantly modify the concordance between DIOCLES and the MBDS, except for AMI and angioplasty. However, the concordance is still almost perfect, and there are only differences in the proportion of patients also with AMI and angioplasty in the MBDS. Thus, the risk of a selection bias as a result of mismatches seems unlikely and, even so, the sensitivity would still be around 80% and the specificity above 98%.

If the completeness of the sample is deemed important, the failure to detect all of the episodes recorded in the DIOCLES would affect the use of the MBDS in the study of ACS-related health care and could be due to match failures, an issue that should be resolved using the original clinical documentation.

The present work has other limitations. One limitation might be due to the participation in DIOCLES of a private hospital, whose episodes could not be included because they were not registered in the MBDS. Another might be due to the selection of the variables studied; however, because some of the main risk factors for the adjustment of in-hospital mortality due to AMI were considered as diagnoses and procedures, there seems to be no reason to believe that a different selection would have altered our observations. Another, more important, limitation is the result of not having used the primary information on the patients included in DIOCLES as reflected in the medical records, whose future comparison with the data sources used in this work is considered essential to verify the validity of the MBDS.

Irrespective of these considerations, the linkage of the CRs with the MBDS, as discussed here, points more to complementarity rather than to antagonism, to the extent that CRs can systematically benefit from the inclusion of previously validated administrative data (eg, patients' date of birth instead of their age), setting a course that, due to the rapid growth in information technologies, will soon increase the availability of large repositories of clinical and administrative information for secondary uses related both to biomedical research and health care management.

CONCLUSIONS

With substantial agreement, a modest probability of false positives, and a low rate of false negatives, the MBDS can be a useful tool for investigating ACS results in Spain. Nonetheless, there seems to be room for improvement in both hospital coding quality and the agreement between the observational registry and administrative data for admission and discharge dates and in the design of the CRs. A comparison of the DIOCLES and MBDS with medical records could verify its validity.

CONFLICTS OF INTEREST

None declared.

WHAT IS KNOWN ABOUT THE TOPIC?

- The use of administrative databases has been questioned, due to the existence of biases that could compromise their use for outcomes research, although other studies have shown their validity.
- The study of ACS in Spain has been addressed using CRs and administrative databases, but the concordance between the 2 information sources has not yet been assessed.
- Although the available data are anonymized, there are robust methods to link CRs and administrative databases through indirect identifiers allowing their comparison.

WHAT DOES THIS STUDY ADD?

- The agreement between DIOCLES and the MBDS was almost perfect and the latter showed a strong ability to discriminate episodes with the same diagnoses and procedures as DIOCLES.
- The presence of differences in the discrimination power of the MBDS between hospital quintiles according to their contribution to DIOCLES indicates an inverse association between registry quality and patient volume.
- The MBDS can be a useful tool for the investigation of ACS outcomes in Spain. A comparison of DIOCLES and the MBDS with medical records could verify its validity.

SUPPLEMENTARY MATERIAL

Supplementary material associated with this article can be found in the online version available at https://doi.org/10.1016/j.rec.2018.01.026.

REFERENCES

1. Nichols M, Townsend N, Luengo-Fernandez R, et al. European Cardiovascular Disease Statistics 2012. European Heart Network, Brussels, European Society of Cardiology. *Sophia Antipolis.* 2012. Available at: https://www.escardio.org/static_file/Escardio/Press-media/press-releases/2013/EU-cardiovascular-disease-statistics-2012.pdf. Accessed 16 Jan 2018.
2. Benjamin EJ, Blaha MJ, Chiuve SE, et al. Heart disease and stroke statistics—2017 update: a report from the American Heart Association. *Circulation.* 2017;135:e146–e603.
3. Dégano IR, Elosua R, Marrugat J. Epidemiology of Acute Coronary Syndromes in Spain: Estimation of the Number of Cases and Trends From 2005 to 2049. *Rev Esp Cardiol.* 2013;66:472–481.
4. Cabadés A, López-Bescós L, Arós F, et al. Variability in the Management and Prognosis at Short- and Medium-term of Myocardial Infarct in Spain: the PRIAMHO Study. Registration Project of Hospital Acute Myocardial Infarct. *Rev Esp Cardiol.* 1999;52:767–775.
5. Arós F, Cuñat J, Loma-Osorio A, et al. Management of Myocardial Infarction in Spain in the Year 2000. The PRIAMHO II Study. *Rev Esp Cardiol.* 2003;56:1165–1173.
6. Bueno H, Bardají A, Fernández-Ortiz A, Marrugat J, Martí H, Heras M. Management of Non-ST-Segment-Elevation Acute Coronary Syndromes in Spain. The DES-CARTES (Descripción del Estado de los Síndromes Coronarios Agudos en un Registro Temporal ESpañol) Study. *Rev Esp Cardiol.* 2005;58:244–252.
7. Ruiz-Bailén M, Macías-Guaras I, Rucabado-Aguilar L, et al. Estancia media y pronóstico en la angina inestable: resultados del registro ARIAM. *Med Clin (Barc).* 2007;128:281–290.
8. Ferreira-González I, Permanyer-Miralda G, Marrugat J, et al. MASCARA (*Manejo del Síndrome Coronario Agudo. Registro Actualizado*) Study. General findings. *Rev Esp Cardiol.* 2008;61:803–816.
9. Barrabés JA, Bardají A, Jiménez-Candil J, et al. Prognosis and Management of Acute Coronary Syndrome in Spain in 2012: The DIOCLES Study. *Rev Esp Cardiol.* 2015;68:98–106.
10. Sendra JM, Sarría A, Iñigo MJ. Desarrollo de un modelo de ajuste por el riesgo para el infarto agudo de miocardio en España. Comparación con el modelo de Charlson y el modelo ICES. Aplicaciones para medir resultados asistenciales. *Rev Esp Salud Publica.* 2006;80:665–677.
11. Andrés E, Cordero A, Magán P, et al. Long-Term Mortality and Hospital Readmission After Acute Myocardial Infarction: an Eight-Year Follow-Up Study. *Rev Esp Cardiol.* 2012;65:414–420.
12. Bertomeu V, Cequier A, Bernal JL, et al. In-hospital Mortality Due to Acute Myocardial Infarction. Relevance of Type of Hospital and Care Provided. RECALCAR Study. *Rev Esp Cardiol.* 2013;66:935–942.
13. Cequier Á, Ariza-Solé A, Elola FJ, et al. Impact on Mortality of Different Network Systems in the Treatment of ST-Segment Elevation Acute Myocardial Infarction. The Spanish Experience. *Rev Esp Cardiol.* 2017;70:155–161.
14. Rodriguez-Padial L, Elola FJ, Fernández-Pérez C, et al. Patterns of inpatient care for acute myocardial infarction and 30-day, 3-month and 1-year cardiac diseases readmission rates in Spain. *Int J Cardiol.* 2017;230:14–20.
15. Registro de altas de hospitalización: CMBD del Sistema Nacional de Salud. Glosario de términos y definiciones. Portal estadístico SNS [updated Sep 2016]. p. 5-6. Available at: http://pestadistico.inteligenciadegestion.msssi.es/publicoSNS/comun/DescargaDocumento.aspx?IdNodo=6415. Accessed 16 Jan 2018.
16. Demlo LK, Campbell PM. Improving hospital discharge data: lessons from the National Hospital Discharge Survey. *Med Care.* 1981;19:1030–1040.
17. Lloyd SS, Rissing JP. Physician and coding errors in patient records. *JAMA.* 1985;254:1330–1336.
18. Fisher ES, Whaley FS, Krushat WM, et al. The accuracy of Medicare's hospital claims data: progress has been made, but problems remain. *Am J Public Health.* 1992;82:243–248.
19. Iezzoni LI, Ash AS, Shwartz M, Mackiernan YD. Differences in procedure use, in-hospital mortality, and illness severity by gender for acute myocardial infarction patients: are answers affected by data source and severity measure? *Med Care.* 1997;35:158–171.
20. Iezzoni LI, Ash AS, Shwartz M, et al. Predicting who dies depends on how severity is measured: implications for evaluating patient outcomes. *Ann Intern Med.* 1995;123:763–770.
21. Ferreira-González I, Marsal JR, Mitjavila F, et al. Patient registries of acute coronary syndrome: assessing or biasing the clinical real world data? *Circ Cardiovasc Qual Outcomes.* 2009;2:540–547.
22. Lohr KN. Use of insurance claims data in measuring quality of care. *Int J Technol Assess Health Care.* 1990;6:263.
23. Hsia DC, Krushat WM, Fagan AB, Tebbutt JA, Kusserow RP. Accuracy of diagnostic coding for Medicare patients under the prospective-payment system. *N Engl J Med.* 1988;318:352–355.
24. Hannan EL, Kilburn H, Lindsey ML, Lewis R. Clinical versus administrative data bases for CABG surgery. *Med Care.* 1992;30:892–907.
25. Jollis JG, Ancukiewicz M, DeLong ER, Pryor DB, Muhlbaier LH, Mark DB. Discordance of databases designed for claims payment versus clinical information systems. *Ann Intern Med.* 1993;119:844–850.
26. Weingart SN, Iezzoni LI, Davis RB, et al. Use of administrative data to find substandard care: validation of the complications screening program. *Med Care.* 2000;38:796–806.
27. Humphries KH, Rankin JM, Carere RG, Buller CE, Kiely FM, Spinelli JJ. Co-morbidity data in outcomes research: are clinical data derived from administrative databases a reliable alternative to chart review? *J Clin Epidemiol.* 2000;53:343–349.
28. Parker JP, Li Z, Damberg CL, Danielsen B, Carlisle DM. Administrative versus clinical data for coronary artery bypass graft surgery report cards: the view from California. *Med Care.* 2006;44:687–695.
29. Krumholz HM, Wang Y, Mattera JA, et al. Administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with an acute myocardial infarction. *Circulation.* 2006;113:1683–1692.
30. Austin PC, Daly PA, Tu JV. A multicenter study of the coding accuracy of hospital discharge administrative data for patients admitted to cardiac care units in Ontario. *Am Heart J.* 2002;144:290–296.
31. Pasquali SK, Jacobs JP, Shook GJ, et al. Linking clinical registry data with administrative data using indirect identifiers: implementation and validation in the congenital heart surgery population. *Am Heart J.* 2010;160:1099–1104.
32. Hammill BG, Hernandez AF, Peterson ED, Fonarow GC, Schulman KA, Curtis LH. Linking inpatient clinical registry data to Medicare claims data using indirect identifiers. *Am Heart J.* 2009;157:995–1000.
33. Setoguchi S, Zhu Y, Jalbert JJ, Williams LA, Chen CY. Validity of deterministic record linkage using multiple indirect personal identifiers: linking a large registry to claims data. *Circ Cardiovasc Qual Outcomes.* 2014;7:475–480.
34. Genders TS, Spronk S, Stijnen T, Steyerberg EW, Lesaffre E, Hunink M. Methods for calculating sensitivity and specificity of clustered data: a tutorial. *Radiology.* 2012;265:910–916.
35. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.
36. Ribera A, Ferreira-González I, Cascant P, Pons JM, Permanyer-Miralda G. ARCA study investigators. Evaluation of Risk-Adjusted Hospital Mortality After Coronary Artery Bypass Graft Surgery in the catalan Public Healthcare System. Influence of Hospital Management Type (ARCA study). *Rev Esp Cardiol.* 2006;59:431–440.
37. Ribera A, Marsal JR, Ferreira-González I, et al. Predicting In-Hospital Mortality With Coronary Bypass Surgery Using Hospital Discharge Data: Comparison With a Prospective Observational Study. *Rev Esp Cardiol.* 2008;61:843–852.
38. Cavero-Carbonell C, Gimeno-Martos S, Zurriaga Óaue, Rabanaque-Hernández MJ, Martos-Jiménez C. La validez del Conjunto Mínimo Básico de Datos como fuente de identificación de las anomalías congénitas en la Comunitat Valenciana. *Gac Sanit.* 2017;31:220–226.
39. Hernández Medrano I, Guillán M, Masjuan J, Alonso Cánovas A, Gogorcena MA. Reliability of the minimum basic dataset for diagnoses of cerebrovascular disease. *Neurologia.* 2017;32:74–80.
40. Lambert L, Blais C, Hamel D, et al. P. Evaluation of care and surveillance of cardiovascular disease: can we trust medico-administrative hospital data? *Can J Cardiol.* 2012;28:162–168.
41. Alfaro M, Gogorcena M, Cózar R, López O, López P, Salmador E. *Metodología de análisis de la hospitalización en el Sistema Nacional de Salud. Modelo de indicadores basado en el Registro de Altas (CMBD) documento base.* Madrid: Ministerio de Sanidad y Consumo; 2008. Available at: https://www.msssi.gob.es/estadEstudios/estadisticas/docs/metod_modelo_cmbd_pub.pdf. Accessed 27 Mar 2017.
42. Sarkies MN, Bowles KA, Skinner EH, et al. Data collection methods in health services research: hospital length of stay and discharge destination. *Appl Clin Inform.* 2015;6:96–109.