

Artículo original

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



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 y Francisco Javier Elola^{b,i}

^a Servicio de Control de Gestión, Hospital Universitario 12 de Octubre, Madrid, España

^b Fundación Instituto para la Mejora de la Asistencia Sanitaria, Madrid, España

^c Servicio de Cardiología, Hospital Universitario Vall d'Hebron, Vall d'Hebron, Institut de Recerca (VHIR), CIBER-CV, Universidad Autónoma de Barcelona, Barcelona, España

^d Servicio de Cardiología, Hospital Álvaro Cunqueiro, Vigo, Pontevedra, España

^e Servicio de Cardiología, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria San Carlos (IdISSC), Universidad Complutense, Madrid, España

^f Fundación Interhospitalaria de Investigación Cardiovascular, Madrid, España

^g Servicio de Medicina Preventiva, Instituto de Investigación Sanitaria San Carlos (IdISSC), Universidad Complutense de Madrid, Madrid, España

^h Servicio de Cardiología, Hospital Universitario de Tarragona Joan XXIII, Institut d'Investigació Sanitaria Pere Virgili (IISPV), Universidad Rovira Virgili, Tarragona, España

ⁱ Elola Consultores, Madrid, España

Historia del artículo:

Recibido el 5 de mayo de 2017

Aceptado el 17 de enero de 2018

On-line el 9 de marzo de 2018

Palabras clave:

Síndrome coronario agudo

Investigación de resultados

Registros clínicos

Bases de datos administrativas

Vinculación de datos

RESUMEN

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 ($\kappa = 0,863$); 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.

© 2018 Sociedad Española de Cardiología. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

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

ABSTRACT

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.

Keywords:

Acute coronary syndrome

Outcomes research

Clinical registries

Administrative database

Data linkage

* Autor para correspondencia: Servicio de Control de Gestión, Hospital Universitario 12 de Octubre, Avda. de Córdoba s/n, 28041 Madrid, España.
Correo electrónico: Jluis.bernal@movistar.es (J.L. Bernal).

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.

Full English text available from: www.revespcardiol.org/en

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

Abreviaturas

CMBD: Conjunto Mínimo Básico de Datos
IAM: infarto agudo de miocardio
RC: registros clínicos
SCA: síndrome coronario agudo

INTRODUCCIÓN

La enfermedad cardiovascular causa anualmente en Europa más de 4 millones de muertes, la mayor parte por enfermedad coronaria¹, y aunque la tasa de mortalidad por cardiopatía isquémica ha descendido en las últimas décadas en los países desarrollados, sigue siendo la causa de aproximadamente un tercio de todas las muertes de mayores de 35 años². En España, se ha estimado un considerable aumento de la incidencia del síndrome coronario agudo (SCA) durante los próximos 35–40 años, paralelo al envejecimiento poblacional³.

La prevalencia del SCA y los costes que implica su atención representan una carga asistencial y económica de primera magnitud que determina la importancia de la evaluación de los resultados de la práctica clínica y la investigación de los factores que la condicionan. En España, el estudio del SCA se ha abordado a partir de registros clínicos (RC)^{4–9} y bases de datos administrativas^{10–14}, fundamentalmente el Conjunto Mínimo Básico de Datos (CMBD) del Sistema Nacional de Salud¹⁵, pero no se ha contrastado la concordancia entre ambas fuentes.

Las ventajas y limitaciones del uso de las bases de datos administrativas en la evaluación de resultados en salud se han analizado profusamente^{16–18}. Son relativamente fáciles de obtener¹⁹, sensiblemente menos costosas que recabar datos primarios mediante la revisión de historias clínicas o mantener registros específicamente diseñados para usos secundarios²⁰, en los que la inclusión irregular de pacientes puede introducir sesgos sistemáticos²¹, y ofrecen información uniforme de grandes poblaciones a lo largo del tiempo²². Sin embargo, su calidad depende de la precisión de la codificación de diagnósticos y procedimientos, realizada por regla general a partir de informes de alta²³, y se han identificado diferentes tipos de sesgos que pueden comprometer su uso: algunas no permiten distinguir entre complicaciones y comorbilidades²⁴ o entre enfermedades crónicas y agudas, y otras no incluyen información clínica relevante, como la medicación administrada o los resultados de las pruebas de laboratorio²⁵; pero, en sentido contrario, numerosos trabajos han avalado su fiabilidad^{26–28} y otros no han encontrado diferencias significativas en el ajuste de riesgo de la mortalidad y reingresos por infarto agudo de miocardio (IAM) o insuficiencia cardiaca al comparar los resultados obtenidos con las bases de datos administrativas y RC²⁹.

Aunque ambas fuentes están sujetas a estrictas reglas de anonimización y no incluyen identificadores directos de los pacientes, se han desarrollado métodos robustos para vincularlas

a partir de identificadores indirectos, que permiten relacionar los registros correspondientes a un mismo episodio en cada origen de datos^{30–33}. Utilizando procedimientos de esta naturaleza, el presente trabajo tuvo por objeto evaluar la concordancia entre los datos del DIOCLES (Descripción de la Cardiopatía Isquémica en el Territorio Español), el último gran registro realizado en España con selección aleatoria de hospitales para estudiar la mortalidad, hospitalaria y a los 6 meses, de los pacientes ingresados por sospecha de SCA y describir su tratamiento⁹, y el CMBD, a fin de analizar las implicaciones de la utilización de RC y las bases de datos administrativas en el estudio del SCA en España.

MÉTODOS

Fuentes de datos

El DIOCLES es un estudio observacional multicéntrico y transversal, con control de calidad, realizado en el primer semestre de 2012, con la participación de 44 centros de 13 comunidades autónomas, que recabó prospectivamente datos de pacientes de edad ≥ 18 años ingresados consecutivamente por sospecha de SCA. Se registraron variables demográficas, factores de riesgo, antecedentes patológicos, presentación clínica, tratamiento prehospitalario, hospitalario y al alta, complicaciones y mortalidad hospitalaria. A los 6 meses se determinó, mediante entrevista telefónica, la ocurrencia de muerte, fecha y causa. La población de estudio fue de 2.557 pacientes de los 3.059 evaluados, tras excluir a 502 por diferentes motivos⁹.

El CMBD de 2012 contiene 400.861 episodios de hospitalización registrados en hospitales del Sistema Nacional de Salud con diagnóstico principal de enfermedades del «área del corazón» o sin ese diagnóstico y alta dada por un servicio de cardiología o cirugía cardiaca. El CMBD incluye información sobre las características demográficas de los pacientes atendidos, variables referidas al proceso de atención y de tipo clínico, relacionadas con las enfermedades y condiciones del paciente y los procedimientos realizados durante su atención, codificados mediante la Clasificación Internacional de Enfermedades, Novena Revisión, Modificación Clínica¹⁵.

Procedimiento de vinculación

El DIOCLES y el CMBD difieren tanto en su alcance (el DIOCLES incluye el seguimiento a los 6 meses después del alta y el CMBD, únicamente el episodio de hospitalización) como en sus modelos de datos. Dado que carecen de atributos comunes que permitan emparejar unívocamente los registros correspondientes a un mismo episodio (identificadores directos), se ensayaron diferentes identificadores indirectos para vincularlas³² y, finalmente, se seleccionó el que más emparejamientos únicos permitió encontrar, compuesto por el código del hospital, las fechas de ingreso y alta, el sexo y la edad del paciente, porque la fecha de nacimiento no quedó registrada en el DIOCLES.

Análisis estadístico

Se realizó un análisis descriptivo de las variables estudiadas; las discretas se expresaron como n (%) y las cuantitativas, cuya normalidad se contrastó mediante el test de Shapiro Wilks, como media \pm desviación estándar. Para comparar variables discretas, se utilizó la prueba de la χ^2 o el test exacto de Fisher, según fuera necesario, y para las variables cuantitativas, el test de la t de Student o la U de Mann-Whitney.

Se utilizó el coeficiente kappa de Cohen para evaluar la concordancia interobservadores, y para medir la validez del CMBD, se calcularon la sensibilidad, la especificidad y los valores predictivos positivo y negativo, para lo cual se consideraron, entre las variables comunes de ambas fuentes, algunas de las más relevantes para el ajuste por riesgo de mortalidad hospitalaria por IAM¹² (tabla 1). Para estimar la discriminación del CMBD considerando conjuntamente los valores de las variables analizadas, se utilizó el área bajo la curva ROC (receiver operating characteristic); la homogeneidad de κ y del área bajo la curva ROC se contrastaron con la prueba de la χ^2 . Las áreas bajo la curva ROC se compararon por hospitales agrupados en quintiles según el número de pacientes incluidos en el DIOCLES. La dependencia entre las observaciones de un mismo paciente se analizó mediante el coeficiente de correlación intragrupal (ρ) y el factor de inflación de la varianza³⁴.

El grado de acuerdo se interpretó según la escala de Landis y Koch³⁵. La significación estadística se definió por un valor de $p < 0,05$ y los análisis se realizaron con Epidat v3.1 y STATA 13.0.

Análisis de sensibilidad

El impacto de los emparejamientos fallidos en los resultados se evaluó mediante un análisis de sensibilidad, admitiendo en la vinculación diferencias de hasta 2 días en las fechas de ingreso y alta entre el DIOCLES y el CMBD y seleccionando, entre los emparejamientos múltiples, los episodios que presentaron menos discrepancias entre las condiciones estudiadas.

Además, como procedimiento de vinculación alternativo, se seleccionaron los episodios del CMBD con SCA como diagnóstico principal y fecha de ingreso en el primer semestre de 2012, y se

utilizaron el código de hospital, la edad, el sexo y las fechas de ingreso y alta como identificadores indirectos para emparejarlos con los episodios registrados en el DIOCLES. Sobre los resultados obtenidos mediante este procedimiento, también se realizó un análisis de sensibilidad siguiendo los mismos criterios que con el procedimiento original.

RESULTADOS

De los 2.557 pacientes registrados en el DIOCLES, se descartó a 28 (1,1%), 17 por falta de alguna de las variables utilizadas para la vinculación y 11 correspondientes a un hospital privado, no incluido en el CMBD del Sistema Nacional de Salud. De los 2.529 restantes, se logró emparejar unívocamente a 1.539 (60,85%) (figura 1), que constituyeron la población del estudio, con una media de edad de $66,95 \pm 13,03$ años y 375 (24,4%) mujeres.

Los registros del DIOCLES emparejados presentaron mayor proporción de IAM y angioplastias y menor proporción de *bypass* que los no emparejados, y no se encontraron diferencias significativas en edad, proporción de mujeres y demás condiciones estudiadas. Entre los episodios emparejados, la prevalencia de todas las condiciones fue significativamente mayor en el DIOCLES que en el CMBD (tabla 1).

En conjunto, el grado de acuerdo entre ambas fuentes de datos fue casi perfecto ($\kappa = 0,863$), con 8.766 (94,93%) acuerdos observados, y osciló entre moderado para peor Killip, sustancial para accidente cerebrovascular, trombolisis y angioplastias, y casi perfecto para IAM y *bypass*. El CMBD, consideradas todas las variables, mostró una sensibilidad del 85,10% y una especificidad del 98,31%, con valor predictivo positivo del 94,55% y valor predictivo negativo del 95,05% (tabla 2), y no se observó dependencia entre las observaciones de un mismo paciente ($\rho \leq 0,04$; factor de inflación de la varianza $\leq 1,06$). En la tabla 1 del material suplementario se presentan los intervalos de confianza del 95% de las estimaciones generales brutas y ajustadas por los métodos utilizados (ρ y factor de inflación de la varianza).

Según su participación en el DIOCLES, los puntos de corte para agrupar hospitales fueron: primer quintil = 41 casos, segundo quintil = 64, tercer quintil = 80 y cuarto quintil = 100. El área bajo la curva ROC fue significativamente diferente por quintiles ($p = 0,007$); las

Tabla 1
Diagnósticos y procedimientos analizados. Prevalencia según el resultado de la vinculación y la fuente de los datos

Diagnósticos y procedimientos	Códigos CIE-9-MC	DIOCLES		p	CMBD (emparejados)	p
		Emparejados	Sin emparejar			
Edad (años)		66,95 \pm 12,99	66,58 \pm 13,03	0,486		
Mujeres		375 (24,37)	271 (27,37)	0,099		
IAM	410.*1	1.094 (71,09)	652 (65,86)	0,006	1.038 (67,45)	0,032
Peor Killip (3 o 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)	0,527	110 (7,15)	< 0,001
ACV	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)	0,896	12 (0,78)	< 0,001
Trombolisis	V45.88, 99.10	133 (8,64)	89 (8,99)	0,818	106 (6,89)	< 0,001
Angioplastia	00.66, 36.01, 36.02, 36.05, 36.06, 36.07	953 (61,92)	542 (54,75)	0,004	836 (54,32)	< 0,001
<i>Bypass</i>	36.1*	27 (1,75)	40 (4,04)	0,008	25 (1,62)	< 0,001
Episodios (n)		1.539	990			

ACV: accidente cerebrovascular; CIE-9-MC: Clasificación Internacional de Enfermedades, Novena Revisión, Modificación Clínica; CMBD: Conjunto Mínimo Básico de Datos; DIOCLES: Descripción de la Cardiopatía Isquémica en el Territorio Español; IAM: infarto agudo de miocardio; Peor Killip (3 o 4): insuficiencia cardíaca grave con edema agudo de pulmón o *shock* cardiogénico durante el ingreso.

Salvo otra indicación, los valores expresan n (%) o media \pm desviación estándar.

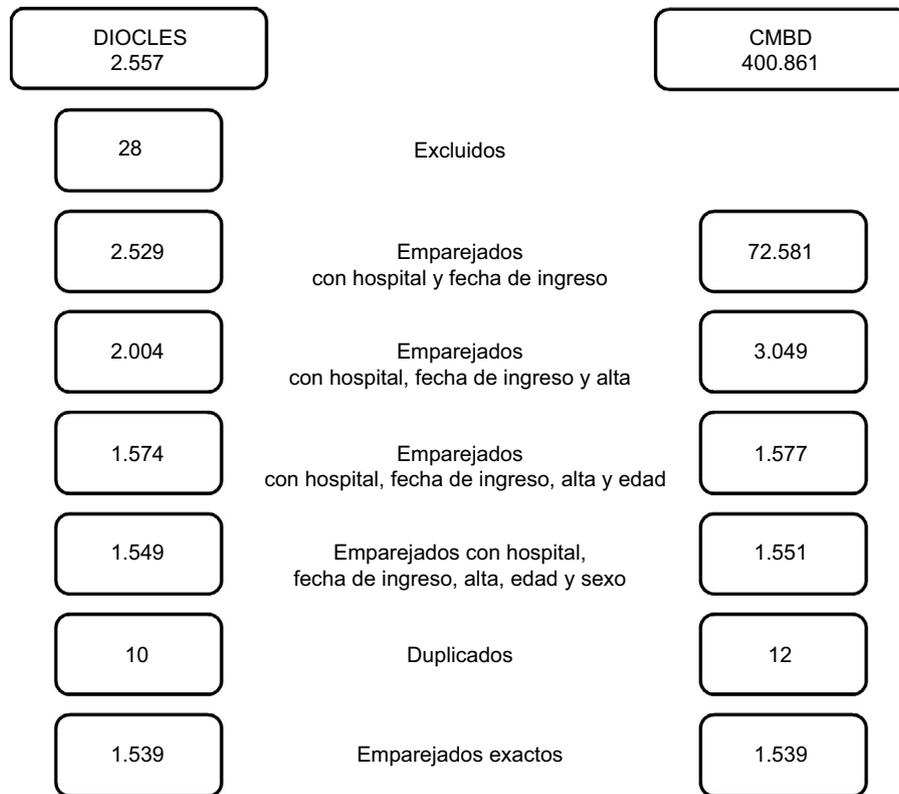


Figura 1. Resultados de la vinculación del DIOCLES y el CMBD. Número de episodios excluidos y emparejados en cada fase del proceso de vinculación. CMBD: Conjunto Mínimo Básico de Datos; DIOCLES: Descripción de la Cardiopatía Isquémica en el Territorio Español.

más altas correspondieron a los quintiles 5 y 3 y las más bajas, a los quintiles 1 y 4 (figura 2).

En el análisis de sensibilidad original (tabla 3), se lograron 2.022 emparejamientos únicos (79,95%) seleccionados entre 2.352 coincidencias, y no se apreciaron diferencias estadísticamente significativas en la concordancia interobservadores para las condiciones analizadas respecto de los resultados obtenidos a partir del procedimiento de vinculación original, excepto IAM y angioplastia, ni en su prevalencia en el DIOCLES ni en el CMBD, excepto también para el IAM y la angioplastia (tabla 4).

Los resultados del procedimiento de vinculación alternativo y su análisis de sensibilidad se presentan en la tabla 2 del material suplementario y la tabla 3 del material suplementario.

DISCUSIÓN

Los resultados obtenidos indican un acuerdo casi perfecto entre el DIOCLES y el CMBD y que la capacidad de este para discriminar episodios con los mismos diagnósticos y procedimientos que el

registro correspondiente del DIOCLES fue alta (área bajo la curva ROC > 0,9), aunque se encontraron diferencias significativas al considerar los hospitales agrupados por quintiles según su contribución al DIOCLES que indican asociación inversa entre la calidad del registro y el volumen de casos.

La concordancia encontrada fue mayor que la publicada por Ribera et al.³⁶ para los diagnósticos y procedimientos incluidos en el estudio ARCA (evaluación del riesgo de la cirugía coronaria en Cataluña) que osciló entre el shock cardiogénico ($\kappa = 0,16$) y el uso de circulación extracorpórea ($\kappa = 0,79$)³⁷, los hallados por Caveró-Carbonell et al.³⁸ para la identificación de anomalías congénitas ($\kappa = 0,70$), y por Hernández Medrano et al.³⁹ para la enfermedad cerebrovascular (tasa de concordancia, 81,87%), en los últimos 2 casos tomando como modelo de referencia la revisión de historias clínicas. La prevalencia fue mayor en el DIOCLES que en el CMBD para todas las condiciones analizadas, y únicamente se encontró $\kappa < 0,7$ en el peor Killip, que incluye el shock cardiogénico, de modo similar a lo encontrado por Lambert et al.⁴⁰ en su evaluación de la precisión de los códigos de enfermedades

Tabla 2

Comparación de diagnósticos y procedimientos registrados en el DIOCLES (Descripción de la Cardiopatía Isquémica en el Territorio Español) frente al Conjunto Mínimo Básico de Datos

Diagnósticos y procedimientos	κ (IC95%)	Sensibilidad, % (IC95%)	Especificidad, % (IC95%)	VPP, % (IC95%)	VPN, % (IC95%)
IAM	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)
Peor Killip (3 o 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)
ACV	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)
Trombolisis	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)
Angioplastia	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)
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 (las 6 condiciones)	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)

ACV: accidente cerebrovascular; IAM: infarto agudo de miocardio; IC95%: intervalo de confianza del 95%; Peor Killip (3 o 4): insuficiencia cardiaca grave con edema agudo de pulmón o shock cardiogénico durante el ingreso; VPN: valor predictivo negativo; VPP: valor predictivo positivo.

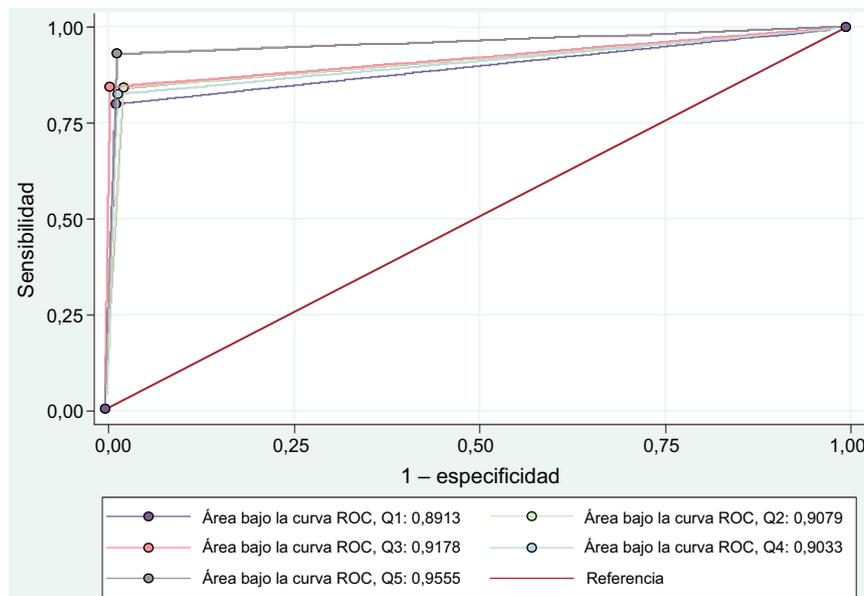


Figura 2. Comparación de las áreas bajo la curva ROC (*receiver operating characteristic*) por quintiles (Q) de hospitales según el número de pacientes registrados en el DIOCLES (Descripción de la Cardiopatía Isquémica en el Territorio Español).

Tabla 3

Análisis de sensibilidad. Comparación de diagnósticos y procedimientos registrados en el DIOCLES (Descripción de la Cardiopatía Isquémica en el Territorio Español) frente al Conjunto Mínimo Básico de Datos

Diagnósticos y procedimientos	κ (IC95%)	Sensibilidad, % (IC95%)	Especificidad, % (IC95%)	VPP, % (IC95%)	VPN, % (IC95%)
IAM	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)
Peor Killip (3 o 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)
ACV	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)
Trombolisis	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)
Angioplastia	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)
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 (las 6 condiciones)	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)

ACV: accidente cerebrovascular; IAM: infarto agudo de miocardio; IC95%: intervalo de confianza del 95%; Peor Killip (3 o 4): insuficiencia cardiaca grave con edema agudo de pulmón o *shock* cardiogénico durante el ingreso; VPN: valor predictivo negativo; VPP: valor predictivo positivo.

Tabla 4

Resultados de la comparación de concordancias y prevalencias en el DIOCLES y el CMBD

Diagnósticos y procedimientos	Igualdad de κ , p	Igualdad de prevalencias en DIOCLES, p	Igualdad de prevalencias en CMBD, p
IAM	0,009	0,649	0,004
Peor Killip (3 o 4)	0,907	0,99	0,467
ACV	0,558	0,965	0,897
Trombolisis	0,218	0,907	0,397
Angioplastia	0,023	0,336	0,02
Bypass	0,379	0,783	0,986

ACV: accidente cerebrovascular; IAM: infarto agudo de miocardio; CMBD: Conjunto Mínimo Básico de Datos; DIOCLES: Descripción de la Cardiopatía Isquémica en el Territorio Español; Peor Killip (3 o 4): insuficiencia cardiaca grave con edema agudo de pulmón o *shock* cardiogénico durante el ingreso.

cardiovasculares en una base de datos administrativa canadiense, en comparación con historias clínicas, aunque su resultado ($\kappa = 0,667$) reveló mayor concordancia que el nuestro.

La explotación de las bases de datos administrativas para la investigación de resultados en salud ofrece importantes ventajas, pero su utilidad depende, en última instancia, de la validez y la consistencia clínica de los datos que, hasta donde alcanza nuestro conocimiento, no se han analizado con anterioridad en relación con el estudio del SCA en España. La variabilidad observada indica que determinadas condiciones (IAM, angioplastia y *bypass*) se codificaron con mayor precisión que otras, de modo que al parecer hay márgenes de mejora de la codificación y, en consecuencia, de la calidad del

CMBD, aunque también cabe la posibilidad de que la falta de concordancia resulte de errores de registro en el DIOCLES²¹.

Limitaciones

Con el hallazgo de que los diagnósticos y los procedimientos analizados tienen una probabilidad de ser verdaderos positivos por encima del 85% si están presentes en el CMBD y de ser verdaderos negativos por encima del 98% si no lo están, el presente trabajo indica un umbral para la validez del CMBD como fuente de datos para el estudio del SCA en España, aunque el porcentaje de emparejamiento que no se pudo resolver es su principal limitación.

Puesto que la utilización de procedimientos alternativos de vinculación no redujo el porcentaje de emparejamientos fallidos, que osciló entre el 55 y el 80%, la dificultad de identificar en el CMBD los episodios incluidos en el DIOCLES tiene que explicarse necesariamente por imprecisiones de registro en cualquiera de las 2 fuentes. A nuestro entender, la consideración en el DIOCLES de la edad, en lugar de la fecha de nacimiento, es un condicionante significativo, pero es posible que existan otros.

La disociación es un requerimiento legal en materia de protección de datos de carácter personal que impide el uso de identificadores directos para vincular los RC y las bases de datos administrativas. Para solventar esta restricción, se desarrollaron procedimientos indirectos que lograron emparejamientos únicos en proporciones que van desde el 58,1% (Setoguchi et al.³³) hasta el 87,5% (Pasquali et al.³¹), el 87,9% (Austin et al.³⁰) y el 90,8% (Hamill et al.³²). Se describieron las razones de que el emparejamiento a veces resultara imposible³²; aunque la mayoría se condicionan por el contexto donde se realizó el estudio y no pueden extrapolarse al nuestro, es probable que en el DIOCLES se registrara, en ocasiones, como fecha de ingreso o alta la de llegada o salida del paciente a urgencias, o al servicio de cardiología, que puede no coincidir con la del CMBD (que refleja cuándo se ingresa al paciente en el hospital y se ocupa una cama de hospitalización⁴¹). Las razones de las discrepancias que Sarkies et al.⁴² encontraron entre la utilización de datos administrativos y registros observacionales para el cálculo de la estancia media por episodio de hospitalización avalarían esta explicación, congruente con que el porcentaje de emparejamientos haya aumentado hasta prácticamente el 80% al admitir diferencias de registro de hasta 2 días en las fechas de ingreso y alta.

En este caso, el análisis de sensibilidad revela que aumentando casi un 32% los casos no cambia significativamente la concordancia entre el DIOCLES y el CMBD, excepto en el IAM y la angioplastia, aunque sigue siendo casi perfecta, y únicamente existen diferencias en la proporción de casos también con IAM y angioplastia en el CMBD, de forma que el riesgo de sesgo de selección como consecuencia de emparejamientos fallidos parece poco probable y, aun así, la sensibilidad quedaría en torno al 80% y la especificidad, por encima del 98%.

Si se considera que la exhaustividad de la muestra es relevante, que no se haya detectado la casi totalidad de los episodios registrados en el DIOCLES condicionaría el uso del CMBD en el estudio de la atención al SCA en la medida que se le pudieran atribuir las causas de los emparejamientos fallidos, cuestión que se debería resolver a partir de la documentación clínica original.

El presente trabajo tiene más limitaciones. Una podría provenir de la participación en el DIOCLES de un hospital privado, cuyos episodios no se pudieron considerar al no estar registrados en el CMBD; otra, de la selección de las variables estudiadas, aunque habiéndose considerado como diagnósticos y procedimientos algunos de los principales factores de riesgo para el ajuste de la mortalidad hospitalaria por IAM, no parece que haya motivo para pensar que una selección diferente habría modificado nuestras observaciones; y otra, de más calado, resulta de no haberse utilizado la información primaria de los pacientes incluidos en el DIOCLES según quedó reflejada en sus historias clínicas, cuya comparación, en una investigación posterior, con las fuentes de datos utilizadas en este trabajo se considera imprescindible para verificar la validez del CMBD.

En todo caso, la vinculación de los RC con el CMBD, según se ha abordado aquí, apunta más a la complementariedad que al antagonismo en la medida que los RC pueden beneficiarse sistemáticamente de la inclusión de los datos administrativos previamente validados (p. ej., la fecha de nacimiento del paciente en lugar de su edad) trazando un camino que, gracias al rápido crecimiento de las tecnologías de la información, generalizará próximamente la disponibilidad de grandes repositorios de información clínica y administrativa para usos secundarios relacionados tanto con la investigación biomédica como con la gestión de la asistencia sanitaria.

CONCLUSIONES

Con un acuerdo sustancial, una discreta probabilidad de falsos positivos y pequeña de falsos negativos, el CMBD puede ser un instrumento útil para la investigación de resultados del SCA en España, aunque parece que hay margen de mejora tanto en la calidad de la codificación y la conciliación del registro observacional de las fechas de ingreso y alta con los datos administrativos como en el propio diseño de los RC. El contraste del DIOCLES y el CMBD con las historias clínicas podría verificar su validez.

AGRADECIMIENTOS

La Sociedad Española de Cardiología ha avalado la realización de este estudio. Los directores del proyecto DIOCLES cedieron los datos del registro, previa anonimización de la identificación del paciente. El Ministerio de Sanidad, Servicios Sociales e Igualdad, a través de la Dirección General de Salud Pública, Calidad e Innovación, cedió la base de datos del CMBD a la Sociedad Española de Cardiología.

FINANCIACIÓN

La investigación para este artículo ha sido financiada mediante una beca no condicionada de la Fundación Interhospitalaria para Investigación Cardiovascular. El estudio DIOCLES se financió mediante una beca sin restricciones de Laboratorios Daiichi-Sankyo-Lilly.

CONFLICTO DE INTERESES

No se declara ninguno.

¿QUÉ SE SABE DEL TEMA?

- Se ha cuestionado el uso de bases de datos administrativas indicando la existencia de sesgos que comprometerían su uso para la investigación de resultados, aunque otros estudios han mostrado su validez.
- El estudio del SCA en España se ha abordado a partir de los RC y las bases de datos administrativas, pero no se ha contrastado hasta ahora la concordancia entre ambas fuentes de información.
- Aunque los datos disponibles están anonimizados, existen métodos robustos para vincular los RC y las bases de datos administrativas mediante identificadores indirectos que permiten su comparación.

¿QUÉ APORTA DE NUEVO?

- El acuerdo observado entre el DIOCLES y el CMBD fue casi perfecto y la capacidad de este para discriminar episodios con los mismos diagnósticos y procedimientos que el DIOCLES fue alta.
- La existencia de diferencias en la discriminación del CMBD entre quintiles de hospitales según su contribución al DIOCLES indica una asociación inversa entre calidad del registro y volumen de casos.
- El CMBD puede ser un instrumento útil para la investigación de resultados del SCA en España. El contraste del DIOCLES y el CMBD con las historias clínicas podría verificar su validez.

MATERIAL SUPLEMENTARIO



Se puede consultar material suplementario a este artículo en su versión electrónica disponible en <https://doi.org/10.1016/j.recesp.2018.01.007>.

BIBLIOGRAFÍA

- 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. Disponible en: https://www.escardio.org/static_file/Escardio/Press-media/press-releases/2013/EU-cardiovascular-disease-statistics-2012.pdf. Consultado 16 Ene 2018.
- 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.
- 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.
- Cabadés A, López-Bescós L, Arós F, et al. Variability in the Management and Prognosis at Short- and Medium-term of Myocardial Infarction in Spain: the PRIAMHO Study. Registration Project of Hospital Acute Myocardial Infarct. *Rev Esp Cardiol*. 1999;52:767–775.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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 Pública*. 2006;80:665–677.
- 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.
- 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.
- 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.
- Rodríguez-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.
- Registro de altas de hospitalización: CMBD del Sistema Nacional de Salud. Glosario de términos y definiciones. Portal estadístico SNS [actualizado Sep 2016]. p. 5–6. Disponible en: <http://pestatistico.inteligenciadegestion.mssi.es/publicoSNS/comun/DescargaDocumento.aspx?IdNodo=6415>. Consultado 16 Ene 2018.
- Demlo LK, Campbell PM. Improving hospital discharge data: lessons from the National Hospital Discharge Survey. *Med Care*. 1981;19:1030–1040.
- Lloyd SS, Rissing JP. Physician and coding errors in patient records. *JAMA*. 1985;254:1330–1336.
- 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.
- 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.
- 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.
- 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.
- Lohr KN. Use of insurance claims data in measuring quality of care. *Int J Technol Assess Health Care*. 1990;6:263.
- 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.
- Hannan EL, Kilburn H, Lindsey ML, Lewis R. Clinical versus administrative data bases for CABG surgery. *Med Care*. 1992;30:892–907.
- 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.
- 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.
- 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.
- 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.
- Krumholz HM, Wang Y, Matterna 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
- 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 Surgery in the Catalan Public Healthcare System. Influence of Hospital Management Type (ARCA study). *Rev Esp Cardiol*. 2006;59:431–440.
- 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.
- Cavero-Carbonell C, Gimeno-Martos S, Zurriaga Oaue, 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.
- Hernández Medrano I, Guillán M, Masjuan J, Alonso Cánovas A, Gogorcena MA. Reliability of the minimum basic dataset for diagnoses of cerebrovascular disease. *Neurología*. 2017;32:74–80.
- 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.
- 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. Disponible en: https://www.mssi.gob.es/estadEstudios/estadisticas/docs/metod_modelo_cmbd_pub.pdf. Consultado 27 Mar 2017.
- 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.