

Validación de los códigos diagnósticos de cáncer de colon y recto del Conjunto Mínimo Básico de Datos

Mirari Márquez Cid / Isabel Valera Niñirola / María Dolores Chirlaque López / Jacinta Tortosa Martínez / Encarnación Párraga Sánchez / Carmen Navarro Sánchez

Servicio de Epidemiología, Consejería de Sanidad de la Región de Murcia, Murcia, España.

(Validation of colorectal cancer diagnostic codes in a hospital administration data set)

Resumen

Objetivo: Validar la capacidad del Conjunto Mínimo Básico de Datos (CMBD) para detectar casos incidentes de cáncer de colon y recto utilizando como estándar de referencia el Registro de Cáncer de Murcia (RCM) y medir la concordancia entre el CMBD y el registro de cáncer en tumores colorrectales.

Material y método: Estudio de validación transversal del CMBD del Hospital Virgen de la Arrixaca de Murcia. La población de estudio son los casos incidentes de colon-recto del año 2000 del RCM y los casos del CMBD en el citado hospital para el mismo año con algún código diagnóstico CIE-9 entre 153.0 y 154.1 eliminando reingresos. Durante el proceso se van a realizar 2 análisis: sólo con el diagnóstico principal y con todos los diagnósticos. Se calcula la sensibilidad, la especificidad, el valor predictivo positivo (VPP) y negativo (VPN) y la concordancia con sus intervalos de confianza (IC) del 95%.

Resultados: Con sólo el primer diagnóstico, el CMBD detecta el 80% de los casos incidentes de cáncer colorrectal con un VPP del 75%; considerando todos los diagnósticos, detecta el 85% con un VPP del 64%. La concordancia en la codificación es elevada tanto con 3 dígitos (índice kappa del 88% [IC del 95%, 0,79-0,97] en primer diagnóstico, y del 89% [IC del 95%, 0,80-0,97] en todos los diagnósticos) como de 4 dígitos (índice kappa del 77% [IC del 95%, 0,68-0,85] en el primer diagnóstico, y del 78% [IC del 95%, 0,70-0,86] en todos los diagnósticos) en ambos análisis.

Conclusiones: El CMBD es una buena fuente para detectar casos incidentes de cáncer al presentar una elevada sensibilidad. La alta concordancia encontrada en las localizaciones tumorales de colon-recto contribuye a consolidar el CMBD como fuente de datos para los registros de cáncer.

Palabras clave: Registro de Cáncer. CMBD. Sensibilidad. Concordancia. Validación. Cáncer colorrectal.

Abstract

Objectives: To validate the ability of a hospital administration data set (minimum data set [MDS]) to detect incident cases of colorectal cancer using the Murcia Cancer Registry (MCR) as the gold standard and to measure agreement between the MDS and registration of colorectal cancer.

Material and method: A cross sectional validation study of the MDS of the main hospital in the region of Murcia (Spain) was conducted. The study population consisted of incident cases of colorectal cancer in 2000 obtained from the MCR and cases in the MDS of the above-mentioned hospital for the same year with an ICD-9 diagnostic code between 153.0 and 154.1, eliminating readmissions. During the process, two analyses were performed: one analysis with the principal diagnosis only and another with all the diagnostic codes. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and agreement was calculated with their 95% confidence intervals (CI).

Results: With the first diagnosis only, the MDS detected 80% of the incident cases of colorectal cancer with a PPV of 75%. With all the diagnoses, the MDS detected 85% of the cases with a PPV of 64%. The agreement in codification was high at three digits (kappa 88% [95% CI, 0.79-0.97] first diagnosis, 89% [95% CI, 0.80-0.97] all diagnoses) as well as at four digits (kappa 77% [IC, 0.68-0.85] first diagnosis, 78% [95% CI, 0.70-0.86] all diagnoses) in both analyses.

Conclusions: Because of its high sensitivity, the MDS is a good source for detecting incident cases of cancer. The high agreement found in the site of colorectal cancer helps to consolidate the MDS as a data source for cancer registration.

Key words: Cancer registry. Hospital discharge database. Sensitivity. Agreement. Validation. Colorectal cancer.

Introducción

Debido al vertiginoso proceso de informatización de las últimas décadas, actualmente los servicios sanitarios disponen de bases de datos de gran utilidad para la medición de problemas de salud. En este sentido, el Conjunto Mínimo Básico de Datos (CMBD)^{1,2} es una fuente de información

Correspondencia: Dra. Mirari Márquez Cid.
Consejería de Sanidad. Servicio de Epidemiología.
Ronda de Levante, 11. Planta Baja. Despacho 12.
30008 Murcia. España.
Correo electrónico: mirari.marquez@carm.es

Recibido: 18 de julio de 2005.

Aceptado: 2 de febrero de 2006.

ampliamente utilizada que nutre diversos registros de información, entre ellos, los registros poblacionales de cáncer. La validación de los códigos diagnósticos contenidos en el CMBD resulta necesaria para su utilización tanto como fuente primaria de datos, como secundaria en estudios de investigación epidemiológica y registros de información.

Los registros poblacionales de cáncer habitualmente utilizan la información del CMBD como fuente de datos, realizando un proceso minucioso de revisión de todos los códigos diagnósticos de cáncer y su comparación con informes anatomopatológicos, historias clínicas y defunciones. Una línea de investigación actual es la automatización de procesos de decisión que permitan obtener el máximo rendimiento de la información electrónica codificada y reducir el tiempo utilizado en la codificación manual^{3,4}. Debido a que el cáncer es un problema de salud pública y la medición y el control de su incidencia un objetivo prioritario, los indicadores de control de calidad (exactitud y exhaustividad)⁵ de los registros no sólo se deben mantener, sino que tienen que mejorar al utilizar fuentes de información codificada, lo que requiere un proceso de validación previo⁶.

Aunque en España la investigación en servicios sanitarios^{7,8} con bases de datos administrativas se ha incrementado en los últimos años, no disponemos de estudios realizados en nuestro entorno que evalúen la validez del CMBD para detectar casos incidentes de cáncer frente a otros países⁹⁻¹³; sin embargo, recientemente el Instituto Catalán de Oncología¹⁴ ha publicado resultados del aumento de fiabilidad de los datos y reducción de los costes de su registro de cáncer hospitalario mediante su automatización. Los tumores más estudiados han sido los de presentación más frecuente (de pulmón en los hombres y de mama en las mujeres), aunque si consideramos ambos sexos conjuntamente, el cáncer colorrectal es el que mayor número de casos aporta al total de tumores malignos, además de ser uno de los que mayor incremento ha presentado en las últimas décadas¹⁵.

El objetivo del trabajo es medir la capacidad del CMBD para detectar casos incidentes de cáncer colorrectal, mediante el cálculo de indicadores de calidad tomando como estándar de referencia el Registro Poblacional de Cáncer de la Región de Murcia (RCM). También se ha calculado el grado de concordancia entre el CMBD y el RCM para el cáncer colorrectal.

Material y método

El estudio se basa en la validación transversal de los datos del CMBD procedentes del hospital Virgen de la Arrixaca, centro de referencia de la Región de Murcia¹⁶. Los casos seleccionados han sido todas las altas

hospitalarias con códigos CIE-9¹⁷ desde 153.0 a 154.1, que corresponden a tumores malignos primarios de colon y recto (año 2000). Los códigos de 3 dígitos representan el órgano y los de 4 dígitos, la parte del órgano afectada. Según los códigos CIE-9 anteriormente descritos, los linfomas primarios de colon y recto quedan excluidos. El conjunto de variables de estudio está formado por los datos de identificación personal (nombre, apellidos, documento nacional de identidad y número de seguridad social) y por los relativos al tumor, como la fecha de incidencia (obtenida siguiendo las recomendaciones de la ENCR¹⁸), la localización y morfología del tumor (CIEO-2)¹⁹ y el comportamiento invasivo de la neoplasia.

Conjunto Mínimo Básico de Datos

En España se aprobó el CMBD en 1987 por el Consejo Interterritorial del Sistema Nacional de Salud, siguiendo como base las recomendaciones de la Unión Europea para desarrollar el Conjunto Mínimo Básico de Datos Europeo²⁰. Este grupo de datos recoge hasta 13 códigos diagnósticos: el primero representa el diagnóstico principal (causa que motiva el ingreso) y el resto lo forman los diagnósticos secundarios. Del CMBD se seleccionan los casos con diagnóstico de tumor maligno de colon y recto con códigos CIE-9 entre 153.0 y 154.1 y fecha de ingreso entre el 1 de enero y el 31 de diciembre de 2000, procedentes del hospital de estudio, lo que supone un total de 218 casos. Se ha realizado un primer análisis que incluye los casos que tienen código de cáncer colorrectal en el diagnóstico principal (166 casos), y un segundo análisis que incluye los casos con diagnóstico principal o secundarios de tumor colorrectal (218 casos).

El CMBD registra episodios, por lo que un paciente puede tener varios ingresos con el mismo diagnóstico principal o secundario en el mismo año. En nuestro estudio sólo se ha incluido el primer episodio producido en el año 2000, lo que ha supuesto la eliminación del 5,2% del total de casos en el primer análisis en el que se analiza exclusivamente el diagnóstico principal, quedando 157 casos pendientes de estudio. Si se aplica el mismo criterio en el segundo análisis, de los 218 casos un 10% presenta más de un episodio, por lo que finalmente se ha incluido 196 casos.

Registro Poblacional de Cáncer

El RCM, de base poblacional^{21,22}, recoge casos de cáncer desde 1983, publicados periódicamente en *Cancer Incidence in Five Continents*²³, y forma parte de la Red Europea de Registros de Cáncer. El RCM ha aportado al estudio los casos incidentes de colon y recto

del año 2000, definiendo como caso incidente los casos de cáncer comprendidos entre 153.0 y 154.1 de la CIE-9 después de convertir los códigos de la CIE-O-2 en CIE-9 mediante el programa IARCTools, con fecha de incidencia del año 2000. Para obtener los casos del RCM se ha tenido en cuenta que entre las fuentes de información conste el hospital de estudio. Con los criterios anteriormente descritos se ha seleccionado 147 casos.

El RCM está inscrito en la agencia de protección de datos, y cumple con la legislación sobre protección de datos personales contemplados en la Ley Orgánica 15/1999 de Protección de Datos de Carácter Personal²⁴. Además, todo el personal adscrito al RCM firma un compromiso de confidencialidad.

Enlace de casos

Se ha realizado la comparación de los casos seleccionados del CMBD con los incidentes del RCM utilizando una estrategia de enlace que mitiga las pérdidas, consistente en una fase previa de normalización de las variables de cruce y la combinación de 2 tipos de enlaces, deterministas y probabilísticos^{25,26}. El procedimiento de identificación de casos coincidentes ha consistido, en primer lugar, en cruces mediante fusiones fijas por diferentes variables de identificación personal y, en segundo lugar, en cruces probabilísticas o por puntuaciones, procedimiento ampliamente utilizado y contrastado²⁷⁻³⁰. Todos los casos coincidentes se han supervisado manualmente, y en alguno de ellos se tuvo que consultar la documentación adicional disponible.

Análisis de los datos

En primer lugar se ha calculado la sensibilidad, la especificidad, el valor predictivo positivo (VPP) y el valor predictivo negativo (VPN), así como sus intervalos de confianza (IC) del 95% para un riesgo $\alpha = 5\%$, teniendo en cuenta sólo el diagnóstico principal y, posteriormente, todos los códigos diagnósticos del CMBD (fig. 1). La sensibilidad se ha obtenido como el porcentaje de casos que el CMBD identifica como cánceres de colon-recto con respecto al total de casos de tales cánceres identificados en el RCM. La especificidad viene dada por el porcentaje de casos que no son cáncer en el CMBD con respecto al sumatorio del total estimado de casos del RCM en que no figura cáncer de colon y recto en los códigos diagnósticos, y los casos que están en el CMBD y no en el RCM. El VPP se calcula incluyendo en el denominador todos los casos identificados por el CMBD como cáncer colon-recto y en el numerador los verdaderos positivos; por último, el VPN indica el porcentaje de casos que realmente no son cán-

Figura 1. Tabla 2 x 2 para el cálculo de los indicadores de validez.

		Registro de cáncer de Murcia Caso incidente de colon y recto	
		Sí	No
CMBD Caso de alta de cólon y recto	Sí	Verdaderos positivos	Falsos positivos
	No	Verdaderos positivos	Falsos positivos

CMBD: Conjunto Mínimo Básico de Datos.

cer entre todos los que no incluye como tal el CMBD, con respecto al sumatorio de casos que no están ni en el CMBD ni en el RCM.

El total de casos ingresados en el año 2000 en el hospital de estudio asciende a 34.122. Para el cálculo de la especificidad se ha realizado una estimación, debido a que el RCM trabaja con casos y el CMBD con episodios. Para el presente estudio se han eliminado del total de casos los reingresos. El porcentaje de registros correspondientes a reingresos eliminados en los casos incluidos en el estudio se ha extrapolado al total de ingresos, eliminando un 5,2% de los casos para el primer diagnóstico y un 10% teniendo en cuenta todos los diagnósticos.

Se ha calculado la concordancia observada y esperada mediante el índice kappa³¹ y sus IC del 95%. Para realizar los estudios estadísticos se utilizan los programas SPSS y Microsoft Excel.

Resultados

En el análisis realizado teniendo en cuenta sólo el diagnóstico principal, se han incluido 147 casos del RCM y 157 del CMBD. La sensibilidad es del 0,80% (IC del 95%, 0,73-0,86), el VPP del 0,75% (IC del 95%, 0,67-0,81), el VPN y la especificidad del 0,99% (IC del 95%, 0,99-1,00), como puede observarse en la tabla 1. Hay 157 casos que están en el CMBD, de los que 39 no aparecen en el RCM (25,4%) debido a varios motivos: tumores prevalentes (21%), el 2,5% pertenece a otras provincias diferentes de la Región de Murcia y el 1,3% restante corresponde a otras localizaciones o no son malignos.

Para el análisis teniendo en cuenta todos los diagnósticos, los casos del RCM continúan siendo 147, aunque los del CMBD ascienden a 196. La sensibilidad es más elevada (0,85%; IC del 95%, 0,78-0,90) y el VPP menor. Del total de 196 casos que se encuentran en

Tabla 1. Tabla de valores de los casos incidentes de cáncer de colon y recto del RCM y el CMBD

	Verdaderos positivos (%)	Falsos positivos (%)	Verdaderos negativos (%)	Falsos negativos (%)	Sensibilidad (%) ^a	Especificidad (%) ^a	VPP (%) ^a	VPN (%) ^a
Sólo primer diagnóstico	118	39	32,190	29	0,80 (0,73-0,86)	0,99 (0,99-1,00)	0,75 (0,67-0,81)	0,99 (0,99-1,00)
Todos los diagnósticos	125	71	33,894	22	0,85 (0,78-0,90)	0,99 (0,99-1,00)	0,64 (0,57-0,70)	0,99 (0,99-1,00)

RCM: Registro de Cáncer de Murcia; CMBD: Conjunto Mínimo Básico de Datos.

^aEntre paréntesis se incluye el intervalo de confianza del 95%.

el CMBD, 71 no están en el RCM (36,2%), de los que el 31,6% son prevalentes, el 3,0% residentes en otras provincias; un caso codificado por el CMBD como colon el RCM lo registra como tumor de páncreas y otro caso codificado por el CMBD como 153.3 (colon sigmoide) y morfología 8140/3 (adenocarcinoma) el RCM lo confirma como linfoma de colon (200._), cuyo código no se analiza en este estudio.

De los 147 casos incluidos en el RCM, el 15,6% no se encuentran en el CMBD debido a varias causas. La mayoría de los casos difieren en la fecha de incidencia (12 con alta en el año 2001), 7 no han sido codificados en el CMBD como cáncer de colon o recto, y 3 fueron notificados al RCM por anatomía patológica sin ingreso.

La concordancia en la codificación de los casos de cáncer de colon y recto, incluyendo únicamente el diagnóstico principal, es elevada, aunque hay que tener en cuenta que sólo se realiza con los casos coincidentes en ambas fuentes como casos de cáncer colorrectal. Los casos concordantes en CIE-9 de 3 dígitos se muestran en la tabla 2. El porcentaje de concordancia simple es del 94%, con un índice kappa de 0,88 (IC del 95%, 0,79-0,97). Los datos obtenidos al analizar los

casos concordantes en la codificación entre el CMBD y el RCM de 4 dígitos se muestran en la tabla 3, donde se observa una menor concordancia prácticamente en todos los códigos superiores a 153.3. En el código 153.8 la concordancia es cero. El porcentaje de concordancia simple es del 81%, con un índice kappa de 0,77 (IC del 95%, 0,68-0,85).

En la tabla 4 se muestran los casos concordantes de los códigos CIE-9 de 3 dígitos entre el CMBD y el

Tabla 2. Número de casos concordantes de 3 dígitos entre los casos de cáncer de colon y recto del RCM y del CMBD teniendo en cuenta sólo el primer diagnóstico

	RCM		Total
	153	154	
CMBD	153	75	78
	154	3	40
Total	78	40	118

CMBD: Conjunto Mínimo Básico de Datos; RCM: Registro de Cáncer de Murcia.

Tabla 3. Número de casos concordantes de 4 dígitos entre los casos de cáncer de colon y recto del RCM y del CMBD teniendo en cuenta sólo el primer diagnóstico

CMBD	RCM											Total
	153.0	153.1	153.2	153.3	153.4	153.5	153.6	153.8	153.9	154.0	154.1	
153.0	2	0	0	0	0	0	0	0	0	0	0	2
153.1	0	2	0	0	0	0	0	0	0	0	0	2
153.2	0	0	10	0	0	0	0	0	0	0	0	10
153.3	0	0	0	35	0	0	0	1	0	2	1	39
153.4	0	0	0	0	4	0	1	1	0	0	0	6
153.5	0	0	0	0	0	1	0	0	0	0	0	1
153.6	0	0	0	0	1	0	9	1	0	0	0	11
153.8	0	0	0	1	1	0	0	0	0	0	0	2
153.9	0	0	0	1	1	0	1	0	2	0	0	5
154.0	0	0	0	3	0	0	0	0	0	7	4	14
154.1	0	0	0	0	0	0	0	0	0	2	24	26
Total	2	2	10	40	7	1	11	3	2	11	29	118

CMBD: Conjunto Mínimo Básico de Datos; RCM: Registro de Cáncer de Murcia.

RCM, que incluyen todos los diagnósticos. La diagonal de la tabla indica los valores concordantes. El porcentaje de concordancia es del 95% en el caso de los códigos de 3 dígitos con un índice kappa de 0,89 (IC del 95%, 0,80-0,97). En la tabla 5 se indican los casos concordantes en la codificación entre el CMBD y el RCM de 4 dígitos. También en este caso el código 153.8 presenta concordancia cero. El porcentaje de concordancia simple es del 83% en los códigos de 4 dígitos con un índice kappa de 0,78 (IC del 95%, 0,70-0,86). La figura 2 ilustra el número de casos concordantes de 4 dígitos entre CMBD y RCM teniendo en cuenta el primer diagnóstico, mientras que la figura 3 indica el número de casos concordantes de 4 dígitos teniendo en cuenta todos los diagnósticos.

Discusión

La utilización de CMBD como fuente de datos del RCM para la detección de casos incidentes de cáncer es una buena práctica, ya que es capaz de detectar el 80% de los casos si tenemos en cuenta sólo el diagnóstico principal y el 85% si tenemos en cuenta todos los diagnósticos; sin embargo, esta fuente debe ser complementada con otras para conseguir los niveles de exhaustividad y exactitud deseados en los registros de cáncer poblacionales (CIFC, Vol. VIII). Estos resultados no se alejan de los mostrados por otro estudio sobre la calidad del CMBD³², donde se calculó el grado de desacuerdo entre la codificación del CMBD para todas las causas en 8 hospitales de la Región de Murcia y la que realizó un codificador experto y externo al CMBD; se determinó un grado de desacuerdo del 41,6% para el

diagnóstico principal y del 33,5% para el procedimiento quirúrgico principal.

Las estrategias que se han mostrado más fiables para mejorar la capacidad predictiva del CMBD con el propósito de detectar casos incidentes de cáncer han sido las que combinan la información diagnóstica y de procedimientos, cualquiera que sea el patrón de comparación, tanto la historia clínica como el registro de cáncer poblacional. Esta capacidad se incrementa adoptando ventanas temporales que permitan desechar los casos prevalentes. La eficacia de la propuesta se afianza al observar la variabilidad de los resultados encontrada en los diferentes estudios. Mientras que en nuestro estudio el VPP es del 74% si tenemos en cuenta sólo el diagnóstico principal y del 84% considerando todos los diagnósticos, en el cáncer de mama podemos encontrar un rango de VPP para la detección de casos incidentes entre el 43 y el 88%.

Las razones citadas por los autores para la incorrecta identificación de los casos incidentes son numerosas y no difieren en exceso de las del presente trabajo: casos

Tabla 4. Número de casos concordantes de 3 dígitos entre los casos de cáncer de colon y recto del RCM y del CMBD teniendo en cuenta todos los diagnósticos

	RCM		Total	
	153	154		
CMBD	153	80	3	83
	154	3	39	42
Total	83	42	125	

CMBD: Conjunto Mínimo Básico de Datos; RCM: Registro de Cáncer de Murcia.

Tabla 5. Número de casos concordantes de 4 dígitos entre los casos de cáncer de colon y recto del RCM y del CMBD teniendo en cuenta todos los diagnósticos

CMBD	RCM											Total
	153.0	153.1	153.2	153.3	153.4	153.5	153.6	153.8	153.9	154.0	154.1	
153.0	2	0	0	0	0	0	0	0	0	0	0	2
153.1	0	3	0	0	0	0	0	0	0	0	0	3
153.2	0	0	10	0	0	0	0	0	0	0	0	10
153.3	0	0	0	35	0	0	0	1	1	3	0	40
153.4	0	0	0	0	7	0	1	1	1	0	0	10
153.5	0	0	0	0	0	1	0	0	0	0	0	1
153.6	0	0	0	0	1	0	9	0	1	0	0	11
153.8	0	0	0	1	1	0	1	0	0	0	0	3
153.9	0	0	0	0	0	0	0	0	3	0	0	3
154.0	0	0	0	2	0	0	0	0	0	8	2	12
154.1	0	0	0	1	0	0	0	0	0	4	25	30
Total	2	3	10	39	9	1	11	2	6	15	27	125

CMBD: Conjunto Mínimo Básico de Datos; RCM: Registro de Cáncer de Murcia.

Figura 2. Número de casos concordantes de 4 dígitos entre el Conjunto Mínimo Básico de Datos (CMBD) y el Registro de Cáncer de Murcia (RCM), teniendo en cuenta el primer diagnóstico.

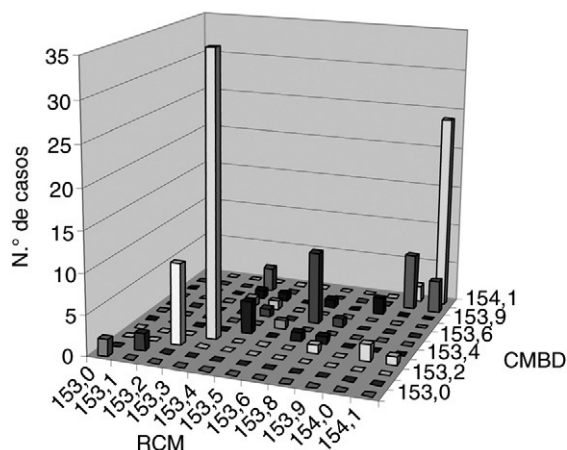
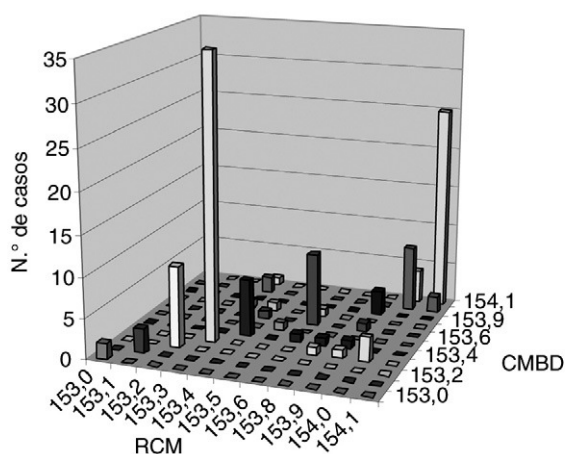


Figura 3. Número de casos concordantes de 4 dígitos entre el Conjunto Mínimo Básico de Datos (CMBD) y el Registro de Cáncer de Murcia (RCM), teniendo en cuenta todos los diagnósticos.



prevalentes, otras provincias, diferente codificación o formar parte de los criterios de exclusión. En el trabajo de Leung et al¹⁰, se indica que el 63% de los casos se deben a una incorrecta codificación de biopsia negativa como cáncer de mama, el 10% se trataba de recidiva tras cirugía conservadora y un 7% se debía a la utilización inadecuada de algún código. Otro estudio sobre cáncer de mama ha mostrado que la concordancia entre los registros de Medicare y el estándar fue superior al 85% cuando a los pacientes se les había realizado cirugía, y menor del 50% cuando a los pacien-

tes se les había realizado sólo una biopsia³³. La relevancia de los falsos positivos se realiza al considerar que el 75% de éstos tiene patrones de cuidados (diagnósticos y procedimientos) similares a los de casos que han resultado verdaderos positivos.

La sensibilidad varía entre el 80% teniendo en cuenta sólo el diagnóstico principal y el 85% si consideramos los 13 códigos diagnósticos, que son valores altos si comparamos estos resultados con los estudios sobre otras localizaciones tumorales. En un ensayo reciente sobre cáncer de próstata, pulmón, colorrectal, mama, páncreas y endometrio, los autores estudiaron la sensibilidad de los registros de Medicare con los datos del Surveillance Epidemiology and End Results. La sensibilidad global para diagnóstico de cáncer en 1993 fue del 66%; la más baja fue para el cáncer de próstata (45,6%) y la más elevada para el de colon y recto (82,1%).

Cabe destacar que la concordancia es elevada para 3 y 4 dígitos de códigos CIE-9 al analizar tanto el diagnóstico principal como todos los diagnósticos; el índice kappa presenta valores en nuestro estudio entre 0,77 y 0,89, algo menor a los hallados por Middleton, que además muestra datos sobre los motivos de las discordancias que coinciden con el presente trabajo, pues se dan principalmente entre sublocalizaciones contiguas, como colon sigmoide, unión rectosigmoidea y recto. También observamos que en los 2 análisis realizados, el código 153.8 (neoplasia maligna que afecta a localizaciones en contigüidad del colon cuyo lugar de origen no puede determinarse) presenta una total discordancia.

Los registros de cáncer poblacionales son los instrumentos idóneos para utilizar como estándar de calidad en los estudios de validación de diagnóstico de cáncer; de hecho, la mayor parte de los estudios utilizan registros de esta naturaleza en lugar de revisión de historias clínicas³⁴. Una de las ventajas de estos registros es que la mayoría presenta coberturas muy elevadas muestra óptimos indicadores de calidad, sigue las normas establecidas por la ENCR y, además, el RCM dispone de un manual de procedimientos para evitar sesgos de clasificación inter/intra-codificación. Una de las limitaciones del CMBD es la falta de exhaustividad al codificarse en 1996 el 86,6% de los registros, aunque en la actualidad están codificados el 99% de los diagnósticos principales de los pacientes ingresados.

El motivo para realizar 2 análisis, teniendo en cuenta sólo el primero o todos los diagnósticos, es porque a pesar de que para medir la validez del CMBD para detectar casos incidentes se debería tomar sólo el primer diagnóstico, manteniendo esta restricción corremos el riesgo de perder casos incidentes que ingresan por otra causa ajena al cáncer. El análisis que se realiza teniendo en cuenta todos los códigos diagnósticos; al incluir más casos, los indicadores de calidad mejoran, a pesar de aumentar los casos prevalentes. No se han

eliminado casos prevalentes en el estudio dado que no se pretende medir la incidencia a partir del CMBD, sino detectar los casos incidentes que éste proporciona.

Se puede concluir que el CMBD es una buena fuente para detectar los casos incidentes de cáncer, al presentar una elevada sensibilidad que, además, aumenta al considerar todos los diagnósticos codificados en un mismo ingreso, aunque de este modo disminuya el VPP. A su vez, la alta concordancia encontrada en las localizaciones tumorales de colon y recto consolida al CMBD como fuente de datos para los registros de cáncer y ofrece una buena calidad en la codificación, aunque es susceptible de mejora.

Agradecimientos

Este proyecto fue financiado parcialmente por el Instituto de Salud Carlos III (Red de Centros RCESP, C03/09 y FIS PI042384).

Bibliografía

- Decreto número 34/1999, de 26 de mayo, por el que se crea el Registro Regional del Conjunto Mínimo Básico de Datos del Alta Hospitalaria. BORM n.º 124 de 1/06/1999.
- López-Abente G, Pollán M, Aragonés N, Pérez B, Hernández V, et al. La situación del cáncer en España. Madrid: Ministerio de Sanidad y Consumo; 2005. p. 131-57.
- Black RJ, Simonato L, Storm HH, Demaret E. Automated data collection in cancer registration. Lyon: IARC Technical reports n.º 32; 1998.
- Tyczynski JE, Demaret E, Parkin DM. Standards and Guidelines for Cancer registration in Europe. Lyon: IARC Technical Publication n.º 40; 2003.
- Middleton RJ, Gavin AT, Reid JS, O'Reilly D. Accuracy of hospital discharge data for cancer registration and epidemiological research in Northern Ireland. *Cancer Causes Control*. 2000;11:899-905.
- Simonato L, Zambon P, Rodella S, et al. A computerized cancer registration network in the Veneto region, north-east of Italy: a pilot study. *Br J Cancer*. 1996;73:1436-9.
- Librero J, Peiro S, Ordiñana R. Chronic comorbidity and outcomes of hospital care: length of stay, mortality and readmission at 30 and 365 days. *J Clin Epidemiol*. 1999;52:171-9.
- Warren JL, Feuer E, Potosky AL, Riley GF, Lynch CF. Use of Medicare hospital and physician data to assess breast cancer incidence. *Med Care*. 1999;37:445-56.
- Freeman JL, Zhang D, Freeman DH, Goodwin JS. An approach to identifying incident breast cancer cases using Medicare Claims data. *J Clin Epidemiol*. 2000;53:605-14.
- Leung KM, Hasan AG, Rees KS, Parker RG, Legorreta AP. Patients with newly diagnosed carcinoma of the breast: validation of a claim-based identification algorithm. *J Clin Epidemiol*. 1999;52:57-64.
- Cooper GC, Yuan Z, Stange KC, Dennos LK, Amino SB, Rimm AA. The sensitivity of Medicare Claims Data for case ascertainment of six common cancers. *Med Care*. 1999;37:436-44.
- Koroukian SM, Cooper GS, Rimm AA. Ability of Medicaid Claims Data to identify incident cases of breast cancer in the Ohio Medical Population. *HSR*. 2003;38:947-60.
- Warren JL, Feuer E, Potosky AL, Riley GF, Lynch CF. Use of Medicare hospital and physician data to assess breast cancer incidence. *Med Care*. 1999;37:445-56.
- Ribes J, Gálvez J, Melià A, Cléries R, Messeguer X, Bosch FX. Automatización de un registro hospitalario de tumores. *Gac Sanit*. 2005;19:221-8.
- González P, Chirlaque M, Jiménez M, Navarro C. Tendencia temporal de la incidencia de cáncer en la Región de Murcia, 1983-1996. *Bol Epidemiol Murcia*. 2001;22:49-55.
- Navarro C, Valera I, Tortosa J, Párraga ME, Garrido S, Chirlaque D, et al. Registro de cáncer de Murcia. Murcia: Manual de Procedimiento; 2003.
- Ministerio de Sanidad y Consumo. Clasificación Internacional de Enfermedades. 9.ª revisión. Modificación clínica. 3.ª ed. Madrid: Ministerio de Sanidad y Consumo; 1996.
- Comisión Europea. Red Europea de Registros de Cáncer. Luxemburgo. Oficina de Publicaciones Oficiales de las Comunidades Europeas; 1995.
- Percy C, Van Loten V, Muir C. International Classification of Diseases for Oncology. 2nd ed. Genève: World Health Organisation (WHO); 1990.
- Roger FH. The minimum basic data set for hospitals statistics in Europe. En: Lambert PM, Roger FH, editors. *Hospital Statistics in Europe*. Brussels, Luxembourg: North Holland; 1982. p. 83-111.
- Miñarro R, Black RJ, Martínez C, Navarro C, Garau I, Izarzugaza I, et al. Incidencia y mortalidad por cáncer en España. Patrones y tendencias. Lyon: IARC Technical Report n.º 36; 2000.
- Navarro C, Chirlaque MD, Rodríguez M, Garrido S, Párraga E, Tortosa J, et al. Estadísticas básicas del Registro de Cáncer de Murcia 1993-1996. Serie informes 33. Murcia: Consejería de Sanidad. Dirección General de Salud Pública; 2003.
- Parkin DM, Whellan SL, Ferlay J, Teppo L, Thomas DB. Cancer Incidence in five continents. Lyon: IARC Sci Publ; 2002.
- Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal. BOE n.º 298, de 14 de diciembre de 1999.
- Couris CM, Rabilloud CCM, Schott AM, Ecochard R. Method of correction to assess the number of hospitalized incident breast cancer cases based on claims databases. *J Clin Epidemiol*. 2002;55:386-91.
- Van den Brandt PA, Schouten LJ, Goldbohm RA, et al. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. *Int J Epidemiol*. 1990;19:553-8.
- Cook LJ, Olson LM, Dean JM. Probabilistic record linkage: relationships between file sizes, identifiers and match weights. *Methods Inf Med*. 2001;40:196-203.
- Fellegi I, Sunter A. A theory for record linkage. *JASA*. 1969;64:1183-210.
- Gill L. Methods for automatic record matching and linking and their use in national statistics. London: National Statistics Methodology Series n.º 25; 2001.
- Winkler W. The State of record linkage and current research problems. Technical Report RR/1999/04. Statistical Research Report Series. Washington DC: US Bureau of the Census; 1999.
- Clackman JM, Koval J. Interval estimation for Cohen's kappa as a measure of agreement. *Statistics in Medicine*. *Stat Med*. 2000;19:723-41.
- Calle JE, Saturno JP, Parra P, Rodenas J, Pérez MJ, San Eustaquio F, et al. Quality of the information container in the minimum basic data set: results from an evaluation in eight hospitals. *Eur J Epidemiol*. 2000;16:1073-80.
- Cooper G, Yuang Z, Stange KC, Denise LK, Amini SB, Rimm AA. Agreement of Medicare Claims and Cancer Registry Data form Assessment of Cancer-related treatment. *Med Care*. 2000;38:411-21.
- Couris CM, Schott AM, Ecochard R, Morgon E, Colin C. A literature review to assess the use of claims databases in identifying incident cancer cases. *Health Serv Outcomes Res Method*. 2003;4:49-63.