

Original breve

# Validez y concordancia de la historia clínica electrónica de atención primaria (AP-Madrid) en la vigilancia epidemiológica de la diabetes mellitus. Estudio PREDIMERC



Elisa Gil Montalbán<sup>a,\*</sup>, Honorato Ortiz Marrón<sup>a</sup>, Dulce López-Gay Lucio-Villegas<sup>a</sup>, Belén Zorrilla Torrás<sup>a</sup>, Francisco Arrieta Blanco<sup>b</sup> y Pedro Nogales Aguado<sup>c</sup>

<sup>a</sup> Subdirección de Promoción de la Salud y Prevención, Consejería de Sanidad, Madrid, España

<sup>b</sup> Servicio de Endocrinología y Nutrición, Hospital Universitario Ramón y Cajal, Madrid, España

<sup>c</sup> Centro de Salud Las Águilas, Servicio Madrileño de Salud, Madrid, España

## INFORMACIÓN DEL ARTÍCULO

### Historia del artículo:

Recibido el 3 de diciembre de 2013

Aceptado el 6 de marzo de 2014

On-line el 13 de abril de 2014

### Palabras clave:

Diabetes mellitus  
Epidemiología  
Vigilancia epidemiológica  
Sensibilidad y especificidad  
Historia clínica electrónica

## R E S U M E N

**Objetivo:** Estudiar la validez y la concordancia de los datos de diabetes en la historia clínica electrónica de atención primaria (AP-Madrid) comparándola con el estudio PREDIMERC.

**Métodos:** En una muestra poblacional de 2268 individuos se calcularon la sensibilidad, la especificidad, el valor predictivo positivo, el valor predictivo negativo y el índice kappa de los casos de diabetes registrados en la historia clínica AP-Madrid, comparándolos con el estudio PREDIMERC (como referencia). Se estimó la prevalencia de diabetes según cada fuente de información.

**Resultados:** La sensibilidad para la diabetes registrada en AP-Madrid fue del 74%, la especificidad del 98,8%, el valor predictivo positivo del 87,9%, el valor predictivo negativo del 97,3% y el índice kappa de 0,78. La prevalencia de diabetes registrada fue del 6,7%, frente al 8,1% de PREDIMERC, siendo la diabetes conocida del 6,3%.

**Conclusiones:** La AP-Madrid es un sistema de información válido para realizar estudios de vigilancia epidemiológica de diabetes en la Comunidad de Madrid.

© 2013 SESPAS. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

## Validity and concordance of electronic health records in primary care (AP-Madrid) for surveillance of diabetes mellitus. PREDIMERC study

### A B S T R A C T

**Objective:** To assess the validity and concordance of diabetes data in the electronic health records of primary care (Madrid-PC) by comparing with those from the PREDIMERC study.

**Methods:** The sensitivity, specificity, positive predictive value, negative predictive value and kappa index of diabetes cases recorded in the health records of Madrid-PC were calculated by using data from PREDIMERC as the gold standard. The prevalence of diabetes was also determined according to each data source.

**Results:** The sensitivity of diabetes recorded in Madrid-PC was 74%, the specificity was 98.8%, the positive predictive value was 87.9%, the negative predictive value was 97.3%, and the kappa index was 0.78. The prevalence of diabetes recorded in Madrid-PC was 6.7% versus 8.1% by PREDIMERC, where known diabetes was 6.3%.

**Conclusions:** The electronic health records of primary care are a valid source for epidemiological surveillance of diabetes in Madrid.

© 2013 SESPAS. Published by Elsevier España, S.L.U. All rights reserved.

### Keywords:

Diabetes mellitus  
Epidemiology  
Public health surveillance  
Sensitivity and specificity  
Electronic health records

## Introducción

La diabetes representa un problema de salud pública a escala mundial que ha aumentado de manera alarmante en los últimos años<sup>1</sup>. En 2007 se realizó en la Comunidad de Madrid el estudio PREDIMERC, para estimar la prevalencia de la diabetes, seleccionando

una muestra aleatoria y representativa de la población de 30-74 años de edad, con determinación de la glucemia basal. En este estudio se detectó que el 22% de los diabéticos estaban sin diagnosticar<sup>2</sup>.

La atención primaria ofrece un ámbito privilegiado para realizar estudios epidemiológicos, por su alta cobertura poblacional y atención continuada<sup>3</sup>. Los avances tecnológicos están favoreciendo el acceso a la información clínica. Cada día es más frecuente la utilización de datos de la historia clínica electrónica para realizar estudios epidemiológicos: GPRD<sup>4</sup> (*General Practice Research Database*), BIFAP<sup>5</sup> (*Base de datos para la Investigación*

\* Autora para correspondencia.

Correo electrónico: elisa.gil@salud.madrid.org (E. Gil Montalbán).

**Tabla 1**  
Características demográficas de la población del estudio, diabéticos registrados en AP-Madrid y diabéticos del estudio PREDIMERC

	Población de estudio n (%) <sup>a</sup>	AP-Madrid		PREDIMERC	
		Diabetes registrada n (%) <sup>a</sup>	Diabetes n (%) <sup>a</sup>	Diabetes conocida n (%) <sup>a</sup>	
<i>Total</i>	2268	181	215	170	
<i>Sexo</i>					
Hombres	1085 (48,4)	110 (58,8)	136 (61,4)	101 (57,1)	
Mujeres	1183 (51,6)	71 (41,5)	79 (38,6)	69 (42,6)	
<i>Edad</i>					
30-59 años	1608 (77,7)	56 (39,2)	73 (42,4)	54 (39,4)	
60-74 años	660 (22,3)	125 (60,8)	142 (57,6)	116 (60,6)	
<i>Nacionalidad</i>					
España	1961 (85,5)	172 (93,6)	201 (91,9)	160 (92,7)	
Extranjero	307 (14,5)	9 (6,4)	14 (8,1)	10 (7,3)	
<i>Nivel de estudios</i>					
Primarios o inferiores	562 (21,4)	93 (49,4)	111 (49,3)	92 (51,8)	
Secundarios	1173 (53,6)	70 (41,4)	80 (39,3)	60 (37,1)	
Universitarios	533 (25,0)	18 (9,4)	24 (11,4)	18 (11,1)	

<sup>a</sup>Porcentaje ponderado.

Farmacoepidemiológica en Atención Primaria) y SIDIAP<sup>6</sup> (Sistema de Información para el Desarrollo de la Investigación en Atención Primaria).

El programa AP-Madrid, que inició su implantación a principios de los años noventa, es el programa utilizado en la Comunidad de Madrid para la gestión de la historia clínica electrónica en atención primaria.

El objetivo de este estudio es analizar la validez, la concordancia y la prevalencia de los casos de diabetes registrados en AP-Madrid en comparación con el estudio PREDIMERC.

## Métodos

Estudio transversal de validación de la diabetes registrada en AP-Madrid. La población de estudio fue la muestra de PREDIMERC (n=2268), utilizado como referencia. Los detalles del diseño de PREDIMERC han sido publicados previamente<sup>2</sup>.

A partir de AP-Madrid se definió «diabetes registrada» como tener registrado un episodio de diabetes (código T89 o T90 de la Clasificación CIAP2) en 2007 o en años anteriores y/o en tratamiento con antidiabéticos orales/insulina activo en 2007. En PREDIMERC se consideró «diabetes» una glucemia basal  $\geq 126$  mg/dl y/o estar en tratamiento farmacológico hipoglucemiante (antidiabéticos orales o insulina) y/o responder afirmativamente a una de estas preguntas: «¿le ha dicho alguna vez un médico o enfermero que tiene diabetes?» o «¿ha tenido alto el azúcar o la glucosa en sangre en más de una ocasión?». Se consideraron «diabetes conocida» los casos que contestaron afirmativamente alguna de las preguntas anteriores o estaban en tratamiento con antidiabéticos orales/insulina.

Se calcularon la sensibilidad, la especificidad, el valor predictivo positivo (VPP) y el valor predictivo negativo (VPN) de la diabetes registrada en AP-Madrid con respecto a la diabetes y la diabetes conocida del estudio PREDIMERC, con sus respectivos intervalos de confianza del 95%. Para el grado de concordancia se calculó el índice kappa. Se estimaron las prevalencias de diabetes, diabetes conocida de PREDIMERC y diabetes registrada en AP Madrid. Dado que la asignación muestral por grupos de edad fue constante, las prevalencias se ponderaron según la estructura de edad de la población de la Comunidad de Madrid. Los intervalos de confianza se calcularon teniendo en cuenta el efecto de diseño del muestreo por conglomerados. Para estudiar las diferencias se utilizó el test de  $\chi^2$ .

## Resultados

Los 2.268 participantes tenían historia clínica electrónica activa en AP-Madrid. La **tabla 1** muestra las características sociodemográficas de la población de estudio. En AP-Madrid se identificaron 181 casos de diabetes registrada, de los cuales 126 (69,6%) tenían registrado episodio y tratamiento farmacológico, 46 (25,4%) sólo episodio y 9 (5%) únicamente tratamiento farmacológico. De PREDIMERC se obtuvieron 215 casos de diabetes, de los cuales 170 (79,1%) eran conocidos.

La **tabla 2** muestra los índices de validez y concordancia de AP-Madrid en relación con PREDIMERC. La sensibilidad de AP-Madrid fue del 74%, la especificidad del 98,9%, el VPP del 87,9% y el VPN del 97,3%; el índice kappa fue de 0,78. Tomando como referencia la diabetes conocida de PREDIMERC, la sensibilidad de AP-Madrid alcanza el 83,5% y disminuye ligeramente el VPP, aunque las diferencias no fueron estadísticamente significativas.

El análisis por sexos muestra valores muy similares de concordancia e índices de validez. Por edad se observan grandes diferencias: la sensibilidad de AP-Madrid en el grupo de 60-74 años de edad alcanza valores  $>80\%$  y una concordancia  $>0,85$ , tanto respecto a la diabetes poblacional como a la diabetes conocida.

En AP-Madrid la prevalencia de la diabetes registrada fue del 6,7%, frente al 8,1% de diabetes poblacional y el 6,3% de diabetes conocida en PREDIMERC. Estas diferencias se mantienen por edad y sexo (**tabla 2**).

## Discusión

Se han obtenido buenos indicadores de validez y concordancia de AP-Madrid, y la prevalencia de diabetes registrada en AP-MADRID, aunque inferior a la prevalencia de diabetes de PREDIMERC, es igual a la diabetes conocida.

La diabetes registrada en AP-Madrid presenta una sensibilidad del 74% y una especificidad del 98,9% con respecto a la diabetes poblacional, con una concordancia de 0,78. Al tomar como referencia la diabetes conocida, la sensibilidad aumenta hasta el 83,5%, probablemente porque los diabéticos conocidos son los que refieren que un médico o una enfermera les han dicho que tienen diabetes. La diabetes no registrada puede explicarse por un infraregistro del médico de atención primaria, o porque sean personas que no utilizan la atención primaria. Por otro lado, el que haya personas que no refieren ser diabéticos y tienen registro de diabetes en AP-Madrid podría deberse a un sesgo de memoria o a que son pacientes recién diagnosticados sin tratamiento farmacológico.

**Tabla 2**

Prevalencia, índices de validez y concordancia de la diabetes registrada en AP-Madrid y la diabetes y diabetes conocida en el estudio PREDIMERC

AP-MADRID	Sí	No	Diabetes -PREDIMERC					Prevalencia <sup>a</sup> diabetes	
			S	E	VPP	VPN	K	Registrada	PREDIMERC
<b>Diabetes registrada</b>									
<i>Total</i>									
Sí	n	n	(IC95%)	(IC 95%)	(IC95%)	(IC95%)	(IC95%)	(IC95%)	(IC95%)
No	159	22	74,0	98,9	87,9	97,3	0,78	6,7	8,1
	56	2031	(67,5-79,6)	(98,4-99,3)	(82,0-92,1)	(96,5-97,9)	(0,74-0,83)	(5,7-7,8)	(6,9-9,2)
<i>Sexo</i>									
<i>Hombres</i>									
Sí	101	9	74,3	99,1	91,8	96,4	0,80	8,5	10,5
No	35	940	(65,9-81,2)	(98,1-99,5)	(84,6-96,0)	(95,0-97,5)	(0,74-0,86)	(7,0-10,4)	(8,8-12,6)
<i>Mujeres</i>									
Sí	58	13	73,4	98,8	81,7	98,1	0,76	5,0	5,6
No	21	1091	(62,1-82,4)	(97,9-99,3)	(70,4-89,5)	(97,1-98,8)	(0,68-0,84)	(3,9-6,5)	(4,4-7,1)
<i>Edad</i>									
<i>30-59 años</i>									
Sí	42	14	57,5	99,1	75,0	98,0	0,64	2,9	3,8
No	31	1521	(45,4-68,8)	(98,4-99,5)	(61,4-85,2)	(97,1-98,6)	(0,53-0,74)	(2,2-3,9)	(3,0-4,9)
<i>60-74 años</i>									
Sí	117	8	82,4	98,5	93,6	95,3	0,85	15,9	18,1
No	25	510	(74,9-88,1)	(96,9-99,3)	(87,4-97,0)	(93,1-96,9)	(0,79-0,90)	(13,3-19,0)	(15,3-21,3)
<b>Diabetes conocida-PREDIMERC</b>									
<b>Diabetes registrada</b>									
<i>Total</i>									
Sí	142	39	83,5	98,1	78,5	98,7	0,79	6,7	6,3
No	28	2059	(76,9-88,6)	(97,4-98,7)	(71,6-84,1)	(98,0-99,1)	(0,74-0,84)	(5,7-7,8)	(5,4-7,4)
<i>Sexo</i>									
<i>Hombres</i>									
Sí	86	29	85,1	97,6	78,2	98,5	0,80	8,5	7,8
No	15	960	(76,4-91,2)	(96,3-98,4)	(69,1-85,3)	(97,4-99,1)	(0,73-0,86)	(7,0-10,4)	(6,3-9,6)
<i>Mujeres</i>									
Sí	56	15	81,2	98,7	78,9	98,8	0,79	5,0	4,9
No	13	1099	(69,6-89,2)	(97,7-99,2)	(67,3-87,3)	(98,0-99,3)	(0,71-0,87)	(3,9-6,5)	(3,8-6,3)
<i>Edad</i>									
<i>30-59 años</i>									
Sí	35	21	64,8	98,6	62,5	98,8	0,62	2,9	2,8
No	19	1533	(50,6-77,0)	(97,9-99,1)	(48,5-74,8)	(98,1-99,2)	(0,51-0,74)	(2,2-3,9)	(2,2-3,9)
<i>60-74 años</i>									
Sí	107	18	92,2	96,7	85,6	98,3	0,86	15,9	14,8
No	9	526	(85,4-96,2)	(94,7-98,0)	(77,9-91,0)	(96,7-99,2)	(0,81-0,91)	(13,3-19,0)	(12,2-17,8)

S: sensibilidad; E: especificidad; VPP: valor predictivo positivo; VPN: valor predictivo negativo; K: índice kappa; IC95%: intervalo de confianza del 95%.

<sup>a</sup> Prevalencia ponderada e IC95%.

Se observa que la prevalencia de diabetes registrada (6,7%) es menor que la de diabetes de PREDIMERC (8,1%). Esta diferencia se debe fundamentalmente a la diabetes no conocida, que según el estudio PREDIMERC representa el 22% del total<sup>2</sup>. Las prevalencias de diabetes registrada y de diabetes conocida (6,3%) coinciden, lo que era de esperar puesto que las personas que refieren tener diabetes es más probable que lo tengan registrado en su historia clínica electrónica.

Estudios recientes han demostrado la validez de la historia clínica electrónica para la realización de estudios epidemiológicos de diabetes<sup>7-10</sup>. Nuestros datos de concordancia son similares a los encontrados por Baena-Díez et al.<sup>11</sup> (kappa = 0,79) en la validación del cuestionario MONICA con información recogida en la historia clínica, aunque la sensibilidad es claramente inferior (96,8%).

La prevalencia de diabetes registrada en nuestro estudio difiere de la estimada por otros estudios realizados en la Comunidad de Madrid también a partir de datos de la historia clínica electrónica de atención primaria, es ligeramente superior a la encontrada por Burgos-Lunar et al.<sup>12</sup> (5,02%) e inferior a la estimada por Martín-Martínez et al.<sup>13</sup> (8% en 2006 y 10,4% en 2010), si bien estos autores sólo incluyen las antiguas áreas 4 y 7, respectivamente, lo que podría explicar las diferencias. También es menor que la obtenida

en Cataluña por Vinagre et al. (7,6%) a partir de datos de SIDIAP, pero este estudio incluye población de 31-90 años de edad<sup>14</sup>.

Una de las limitaciones del estudio es que los datos de AP-Madrid se refieren únicamente a personas que utilizan el sistema público; no obstante, la cobertura sanitaria pública en España es superior al 95%. La mejora progresiva de la calidad y la exhaustividad de los registros incidirá positivamente en la validez de AP-Madrid. Otra limitación es la utilización de una sola determinación de glucemia basal en PREDIMERC, pues en la práctica clínica el criterio diagnóstico requiere tener dos determinaciones de la glucemia basal  $\geq 126$  mg/dl, aunque en estudios poblacionales está aceptado emplear sólo una. Finalmente, debido a que la muestra de diabéticos de 30-59 años de edad es pequeña, los indicadores en este grupo no son concordantes.

En conclusión, AP-Madrid, a través del indicador de diabetes registrada, es una fuente de información válida y oportuna para la vigilancia epidemiológica de la diabetes en la población. Las historias clínicas electrónicas ofrecen una valiosa fuente de información para la realización de estudios epidemiológicos y para obtener datos poblacionales de una manera menos costosa<sup>6,7,13</sup>. Este estudio forma parte de un proyecto de vigilancia epidemiológica de la diabetes a través de AP-Madrid que permitirá estudiar la

evolución de la prevalencia y la incidencia, así como el control de la enfermedad.

#### ¿Qué se sabe sobre el tema?

La historia clínica electrónica constituye un potencial como fuente de información para realizar estudios epidemiológicos. Una cuestión fundamental es evaluar si estos datos generados con fines asistenciales pueden ser válidos como fuente de información epidemiológica.

#### ¿Qué añade el estudio realizado a la literatura?

Los resultados de este estudio muestran una buena validez y concordancia de AP-Madrid para ser utilizada como fuente de información en estudios epidemiológicos sobre diabetes. La validación de AP-Madrid es el punto de partida para crear un sistema de vigilancia de la diabetes a través de la historia clínica electrónica que permita estudiar amplias muestras de población.

#### Editora responsable del artículo

Glòria Pérez.

#### Contribuciones de autoría

E. Gil Montalbán, H. Ortiz Marrón y D. López-Gay Lucio-Villegas han contribuido a la concepción y diseño del trabajo, analizaron e interpretaron los datos, y escribieron la primera versión del artículo. B. Zorrilla Torrás, F. Arrieta Blanco y P. Nogales Aguado han participado en la interpretación de los datos y realizaron una revisión crítica con importantes contribuciones al manuscrito. Todos los autores han revisado y aprobado la versión final para su publicación.

#### Financiación

Este trabajo ha sido financiado parcialmente por el Fondo de Investigación Sanitaria (FIS PI07/1213) 2007-2009.

#### Conflicto de intereses

Ninguno.

#### Bibliografía

- Whiting DR, Guariguata L, Weil C, et al. IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. *Diabetes Res Clin Pract.* 2011;94:311-21.
- Gil Montalbán E, Zorrilla Torrás B, Ortiz Marrón H, et al. Prevalencia de diabetes mellitus y factores de riesgo cardiovascular en población adulta de la Comunidad de Madrid. *Gac Sanit.* 2010;24:233-40.
- Bellón Saameno JA. Grandes bases de datos poblacionales para la investigación en atención primaria: jugamos con ventaja. *Aten Primaria.* 2010;42:549-51.
- Walley T, Mantgani A. The UK General Practice Research Database. *Lancet.* 1997;350:1097-9.
- Salvador Rosa A, Moreno Perez JC, Sonogo D, et al. El Proyecto BIFAP: Base de datos para la Investigación Farmacoepidemiológica en Atención Primaria. *Aten Primaria.* 2002;30:655-61.
- Bolibar B, Fina Avilés F, Morros R, et al. Base de datos SIDIAP: la historia clínica informatizada de atención primaria como fuente de información para investigación epidemiológica. *Med Clin (Barc).* 2012;138:617-21.
- Desai JR, Wu P, Nichols GA, et al. Diabetes and asthma case identification, validation, and representativeness when using electronic health data to construct registries for comparative effectiveness and epidemiologic research. *Med Care.* 2012;50 (Supl 1):30-5.
- Herrett E, Thomas SL, Schoonen WM, et al. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol.* 2010;69:4-14.
- Ramos R, Ballo E, Marrugat J, et al. Validez del Sistema de Información para el Desarrollo de la Investigación en Atención Primaria (SIDIAP) en el estudio de enfermedades vasculares: estudio EMMA. *Rev Esp Cardiol.* 2012;65:29-37.
- Nichols GA, Desai J, Elston-Lafata J, et al. Construction of a multisite DataLink using electronic health records for the identification, surveillance, prevention, and management of diabetes mellitus: the SUPREME-DM project. *Prev Chronic Dis.* 2012;9:e110.
- Baena-Díez JM, Alzamora-Sas MT, Grau M, et al. Validez del cuestionario cardiovascular MONICA comparado con la historia clínica. *Gac Sanit.* 2009;23:519-25.
- De Burgos-Lunar C, Salinero-Fort MA, Cardenas-Valladolid J, et al. Validation of diabetes mellitus and hypertension diagnosis in computerized medical records in primary health care. *BMC Med Res Methodol.* 2011;11:146-57.
- Martín Martínez MA, Carmona Alférez R, Prado Galbarro FJ, et al. Incidencia y prevalencia de diabetes en una población adulta de Madrid: estudio mediante la historia clínica informatizada en atención primaria. *Gac Sanit.* 2013;27:284-5.
- Vinagre I, Mata-Cases M, Hermosilla E, et al. Control of glycemia and cardiovascular risk factors in patients with diabetes type 2 in primary care in Catalonia (Spain). *Diabetes Care.* 2012;35:774-9.