

## ORIGINAL

## DESARROLLO DE UN MODELO DE PREDICCIÓN DE RIESGO DE HOSPITALIZACIONES NO PROGRAMADAS EN EL PAÍS VASCO (\*)

Juan F Orueta Mendiá (1,3), Arturo García-Álvarez (2), Edurne Alonso-Morán (3) y Roberto Nuño-Solinis (3).

(1) Centro de Salud de Astrabudua.Osakidetza (Servicio Vasco de Salud). Erandio. Bizkaia. España.

(2) Kronikgune (Centro de Investigación en Cronicidad).

(3) O+berri. Instituto Vasco de Innovación Sanitaria. Sondika. Bizkaia. España.

Los autores declaran no presentar ningún conflicto de intereses.

(\*) Este trabajo forma parte de un estudio financiado por Kronikgune (Centro de Investigación en Cronicidad) mediante una beca (KRONIK011/035).

## RESUMEN

**Fundamentos:** La hospitalización son eventos indeseables que en ocasiones pueden ser evitados mediante intervenciones proactivas. El objetivo del estudio es determinar la capacidad de modelos basados en *Adjusted Clinical Groups* (ACGs) en nuestro medio para identificar a los pacientes que presentarán ingresos no programados en los meses siguientes a su clasificación, tanto en la población general como en subpoblaciones de enfermos crónicos (diabetes mellitus, enfermedad pulmonar obstructiva crónica e insuficiencia cardíaca).

**Métodos:** Estudio transversal que analizó información de un periodo de 2 años, de todos los residentes en Euskadi mayores de 14 años de edad ( $n=1.964.337$ ). Los datos del primer año (demográficos, índice de privación socioeconómica, diagnósticos, prescripciones, procedimientos, ingresos y otros contactos con el servicio de salud) sirvieron para construir las variables independientes. Las hospitalizaciones del segundo año, las dependientes. Se empleó el área bajo la curva ROC (AUC) para evaluar la capacidad de los modelos en discriminar a los pacientes con hospitalizaciones y se calculó el valor predictivo positivo y la sensibilidad en diferentes puntos de corte.

**Resultados:** En la población general, los modelos para predecir ingresos a los 6 y 12 meses así como hospitalizaciones prolongadas mostraron un comportamiento bueno ( $AUC>0,8$ ), mientras que fue aceptable ( $AUC 0,7-0,8$ ) en los grupos de pacientes crónicos.

**Conclusiones:** Un sistema de estratificación de riesgo de ingresos, basado en ACGs resulta válido y aplicable en nuestro medio. Estos modelos permiten clasificar a los pacientes en una escala de mayor a menor riesgo, lo cual hace posible la aplicación de las intervenciones preventivas más costosas solamente a un pequeño subgrupo de pacientes, mientras que otras menos intensivas pueden proporcionarse a grupos más amplios.

**Palabras clave:** Ajuste de riesgo. Modelos predictivos. Hospitalizaciones. Enfermedades crónicas. Sistemas de información sanitaria.

Correspondencia  
Juan F Orueta  
Centro de Salud de Astrabudua  
Calle Mezo, 35  
48950 Erandio, Bizkaia  
jon.orueta@osakidetza.net

## ABSTRACT

## Development of a Predictive Risk Model for Unplanned Admissions in the Basque Country

**Background:** Hospitalizations are undesirable events that can be avoided to some degree through proactive interventions. The objective of this study is to determine the capability of models based on Adjusted Clinical Groups (ACG), in our milieu, to identify patients who will present unplanned admissions in the following months to their classification, in both the general population and in subpopulations of chronically ill patients (diabetes mellitus, chronic obstructive pulmonary disease and heart failure).

**Methods:** Cross-sectional study which analyzes data from a two year period, of all residents over 14 years old in the Basque Country ( $N = 1.964.337$ ). Data from the first year (demographic, deprivation index, diagnoses, prescriptions, procedures, admissions and other contacts with the health service) were used to construct the independent variables; hospitalizations of the second year, the dependent ones. We used the area under the ROC curve (AUC) to evaluate the capability of the models to discriminate patients with hospitalizations and calculated the positive predictive value and sensitivity of different cutoffs.

**Results:** In the general population, models for predicting admission at 6 and 12 months, as well as long-term hospitalizations showed a good performance ( $AUC > 0.8$ ), while it was acceptable ( $AUC 0.7$  to  $0.8$ ) in the groups of chronic patients.

**Conclusion:** A hospitalization risk stratification system, based on ACG, is valid and applicable in our milieu. These models allow classifying the patients on a scale of high to low risk, which makes possible the implementation of the most expensive preventive interventions to only a small subset of patients, while other less intensive ones can be provided to larger groups.

**Keyword:** Risk adjustment. Predictive modeling. Hospitalizations. Chronic diseases. Health information systems.

## INTRODUCCIÓN

La predicción de las hospitalizaciones y la identificación de las personas que requerirán ingresos en el futuro es motivo de interés para las organizaciones sanitarias por diversas razones. En primer lugar, porque son eventos indeseables para los pacientes y sus familiares. En segundo lugar, porque las hospitalizaciones resultan especialmente costosas para los sistemas de salud, que deben dedicar a ellas una parte sustancial de su presupuesto. Y por último, porque existen evidencias de que en determinados casos pueden ser evitadas mediante la puesta en marcha de intervenciones oportunas<sup>1</sup>.

Por otra parte, una proporción muy importante de los ingresos hospitalarios se concentran en unas pocas personas. Así, en el País Vasco en el año 2011, el 1% de la población requirió 62.757 hospitalizaciones (que supusieron el 29,5% del número total y su coste alcanzó el 46,7% del generado por todos los ingresos en hospitales públicos de la comunidad autónoma). La identificación previa de las personas que necesitarán ser ingresadas en un futuro próximo, especialmente las que requerirán hospitalizaciones prolongadas o repetidas, es un elemento fundamental para la realización de programas proactivos, como los de gestión de casos u otros.

En España<sup>2</sup>, y en otros países, principalmente Estados Unidos<sup>3</sup> y el Reino Unido<sup>4</sup>, la mayor parte de los modelos de predicción de ingresos se han dirigido específicamente a las rehospitalizaciones. Sin embargo, la identificación de las personas en riesgo pero que todavía no han sido previamente ingresadas, ofrece oportunidades de intervención más amplias y puede permitir actuaciones más precoces que eviten un mayor deterioro en este segmento de la población. Los modelos predictivos de ajuste de riesgo son instrumentos que fueron diseñados para clasificar a las personas en función de la cantidad de cuidados sanitarios que requerirán en

el futuro, y pueden emplearse como método de cribado poblacional para identificar a las personas en mayor riesgo de hospitalización<sup>5</sup>.

Un aspecto a destacar es que el objetivo final de estos programas no consiste en la correcta predicción de qué pacientes serán ingresados, sino la selección del subgrupo formado por las personas que con mayor probabilidad obtendrán buenos resultados, como, por ejemplo, los que presentan determinados problemas de salud susceptibles de ser tratados con las citadas actividades proactivas<sup>6</sup>. En el País Vasco se están desarrollando intervenciones en ese sentido, cuyo fin es avanzar hacia un sistema sanitario más integrado y sostenible, que proporcione cuidados de mayor calidad a la población<sup>7</sup>. Una de las iniciativas fue la puesta en marcha del Programa de Estratificación Poblacional<sup>8</sup>, para el cual los ciudadanos son clasificados anualmente aplicando el sistema de *case-mix Adjusted Clinical Groups* (ACGs) de la Universidad Johns Hopkins<sup>9</sup> y otras variables.

El objetivo de este estudio fue determinar la capacidad de modelos basados en ACGs para identificar a pacientes que presentarán ingresos no programados en los periodos de 6 y 12 meses siguientes a su clasificación, así como aquellos que requerirán hospitalizaciones prolongadas. Un segundo objetivo es comprobar la validez de dichos modelos predictivos en determinadas subpoblaciones de enfermos con algunas enfermedades crónicas, que se consideran especialmente sensibles a una atención ambulatoria oportuna para evitar hospitalizaciones.

## MATERIAL Y MÉTODOS

Este estudio transversal fue llevado a cabo en el servicio vasco de salud, Osakidetza. La población de estudio estuvo formada por todas las personas de edad superior a 14 años que contaban con cobertura sanitaria por Osakidetza el 1 de septiembre

de 2008. El periodo de estudio corresponde a dos intervalos consecutivos de 12 meses. Los datos del primer año (de 1/09/2007 a 31/08/2008) se utilizaron para construir las variables explicativas, mientras que los del segundo año (de 1/09/2008 a 31/08/2009) para las variables respuesta. Se estableció un periodo mínimo de seguimiento en el primer año, incluyéndose en el estudio solamente a las personas que dispusieron de aseguramiento público durante al menos 6 meses, independientemente de que hubieran realizado una visita o tenido cualquier tipo de contacto con los servicios de Osakidetza (n=1.973.971). De ellos, 28.182 personas no completaron el segundo año, por defunción (n=18.548) o traslado fuera de la comunidad autónoma y otras causas (n=9.634). Se consideró sujetos de estudio a los ciudadanos que fallecieron en el segundo año, mientras que se dio de baja a los excluidos por otras razones. Por tanto, la población definitiva se compuso de 1.964.337 personas.

El estudio recibió aprobación del Comité ético de investigación clínica de Euskadi (CEIC-E). Se empleó un identificador opaco, para garantizar la confidencialidad de datos personales.

**Fuente de información.** Se empleó la base de datos del Programa de Estratificación Poblacional del País Vasco (PREST), la cual incluye registros individualizados de los ciudadanos que disponen de cobertura sanitaria pública por Osakidetza, esto es, la práctica totalidad de los residentes en la comunidad autónoma. Engloba información procedente de atención primaria, especializada y otras bases de datos del Departamento de Salud, conteniendo información demográfica (edad, sexo y área censal de residencia), códigos CIE-9-MC de todos los diagnósticos (de altas de hospital y de motivos de consulta al médico de primaria y a los servicios de urgencias), procedimientos relevantes, fármacos prescritos, contactos con los servicios de salud y coste de los mis-

mos. Como proxy de la posición socioeconómica individual, se empleó el índice de privación de área censal de residencia elaborado por el proyecto MEDEA<sup>10</sup> categorizado en quintiles. PREST no contiene más información que la registrada de forma rutinaria en las historias clínicas. Una descripción más detallada de esta base de datos puede obtenerse en otra publicación<sup>8</sup>.

**Sistema de clasificación de pacientes: Adjusted Clinical Groups predictive modelling (ACG-PM)<sup>9</sup>.** ACGs es un sistema de case-mix diseñado por investigadores de la Universidad Johns Hopkins. Para sus modelos predictivos emplea como variables explicativas información procedente de diagnósticos, prescripciones, algunos procedimientos relevantes, utilización de servicios sanitarios y su coste en el año previo. Para el presente estudio se empleó la versión 9.0 de ACG-PM.

Los ACGs son 94 categorías autoexcluyentes en las cuales se clasifica a cada persona en función de su edad, sexo y diagnósticos que se le hayan realizado durante los últimos 12 meses. Los diagnósticos son clasificados también en otros 3 modos distintos: en 264 *Expanded Diagnosis Clusters* (EDCs) en función de las características clínicas de este problema de salud, HOSDOM, que identifica patologías con alto riesgo de hospitalización en el año siguiente y un marcador de fragilidad.

Por otra parte, los fármacos se clasifican en 69 *Rx-defined Morbidity Groups* (Rx-MGs), según la patología que puede deducirse a partir de su prescripción. Un marcador de polifarmacia identifica a los pacientes que en un año han recibido más de 12 fármacos distintos.

No todos estos grupos son incluidos como variables en los modelos predictivos. El modelo final incorpora 140 procedentes de diagnósticos (34 categorías para ACGs, 101 EDCs, 4 categorías para el número de

HOSDOM y el marcador de fragilidad) y 65 para prescripciones (64 Rx-MGs y polifarmacia). A ellos se agregan 9 variables demográficas (sexo, más 8 grupos de edad), 20 referentes a coste sanitario previo total y en farmacia (generados a partir de la agrupación de los pacientes en bandas de percentiles) y 16 correspondientes a utilización de recursos (inclusión en programa de diálisis crónica, agrupaciones de número de hospitalizaciones, visitas a urgencias y a consultas externas).

El sistema ofrece predicciones de distintos riesgos de ingreso: en hospital de agudos en los siguientes 6 y 12 meses, en unidad de cuidados intensivos, estancia prolongada en hospital de agudos (más de 12 días ingresado como consecuencia de uno o más ingresos, durante los siguientes 12 meses) y hospitalización como consecuencia de una lesión traumática.

**Variables de estudio y modelos estadísticos.** Se construyeron modelos de regresión logística, empleando 3 variables respuesta diferentes: ingreso no programado en los 6 meses siguientes a la clasificación de cada persona, ingreso en los 12 meses y hospitalización prolongada. De dichas hospitalizaciones se excluyeron las generadas por causa obstétrica o lesiones traumáticas, pues nuestro objetivo fundamental era el diseño de un modelo para la identificación de pacientes (principalmente enfermos crónicos y pluripatológicos) que puedan beneficiarse de programas de gestión de casos.

Se emplearon dos conjuntos de variables independientes:

1.- Las predicciones (*out of the box*) que ofrece directamente el software ACGs. Se basan en la calibración de los modelos realizada por los diseñadores del case-mix en poblaciones americanas

2.- Recalibración de los modelos utilizando el conjunto completo de las 250 variables que emplea el agrupador más el índice de privación del área de residencia.

Para evitar problemas de sobreajuste se realizó validación cruzada, dividiéndose a la población en dos mitades de forma aleatoria. Al evaluar y comparar los modelos se empleó el área bajo la curva ROC (AUC)<sup>11</sup> calculando los intervalos de confianza mediante el método de *bootstrapping*, utilizando 1.000 submuestreos. Además, se calculó el valor predictivo positivo (VPP) y la sensibilidad en diferentes puntos de corte.

Tales análisis se repitieron en la población total (1.964.337) y en tres subpoblaciones de enfermos crónicos, identificados a partir de las agrupaciones de diagnósticos y prescripciones del sistema ACG: diabetes mellitus (100.342), insuficiencia cardiaca (64.191) y enfermedad pulmonar obstructiva crónica (29.516). En estas tres subpoblaciones se compararon las predicciones de los valores *out-of-the-box* con las de dos recalibraciones locales: una obtenida aplicando a cada subpoblación los coeficientes de regresión calculados para toda la población de País Vasco y otra resultante de realizar una nueva calibración (incluyendo validación cruzada) que empleó exclusivamente los datos de cada subpoblación.

Los análisis se realizaron utilizando el programa SAS (versión 9.2)<sup>12</sup> y las funciones de la librería ROCR<sup>13</sup> del entorno de programación R<sup>14</sup>.

**Cálculo de los promedios de coste observado en el año 2 por hospitalizaciones para los pacientes con mayor riesgo esperado.** A partir de las predicción de hospitalizaciones a 12 meses (con el modelo recalibrado localmente) se clasificó en percentiles a toda la población de estudio en función de la probabilidad de que se produjera dicho evento. Se emplearon diferentes puntos de corte y, de este modo, consideramos como “alto riesgo” a las personas cuyo percentil superaba dicho umbral. De este modo, se establecieron 6 bandas de riesgo que no son autoexcluyentes y van desde la más amplia, que incluye al 10% de toda la población ( $\geq$ Percentil 90%) hasta la más restrictiva ( $\geq$ Percentil 99,9%).

Posteriormente se estimaron los costes de las hospitalizaciones observadas en el año 2 a partir de la clasificación de todos los ingresos en Grupos Relacionados por el Diagnóstico (GRD) y sus valores estandarizados. Se calcularon los promedios de coste de hospitalización durante el año 2 para los grupos de pacientes incluidos en cada una de las bandas de riesgo antes citadas.

## RESULTADOS

En la **tabla 1** se muestran algunas características de la población estudiada. El 51% eran mujeres y el porcentaje de mayores de 65 años fue 22%. Casi el 79% de las personas tuvo algún contacto con los servicios de salud de Osakidetza en el primer año y los promedios por paciente de visitas a consultas, urgencias de hospital e ingresos (por causas diferentes a las obstétricas y traumáticas) fueron, respectivamente: 6,17; 0,26 y 0,08.

Durante el segundo año, el 2,97% de las personas requirió una o más hospitalizaciones no programadas durante los primeros 6 meses, más del 5% a lo largo de los 12 meses y el 1,4% de la población estuvo ingresado durante al menos 12 días en dicho periodo. De ellos, solo una minoría lo había estado también el año previo. Es decir, de los 101.384 pacientes con ingresos en el año 2, 75.093 (74%) estuvieron libres de ellos en el año 1, incluso de los 27.758 con estancias prolongadas en el segundo año 17.797 (64%) no presentaron ninguna hospitalización durante el primero. Los modelos mostraron una buena capacidad para discriminar a los pacientes que presentarían hospitalizaciones (**tabla 2**) alcanzando valores de AUC que oscilaron entre 0,796 y 0,870. Como cabría esperar, los modelos recalibrados con nuestros datos mostraron valores más altos (0,823 para predicción de hospitalización en 6 meses; 0,809 en 12 meses y 0,870 de hospitalización prolongada) que los que ofrece por defecto el software del sistema ACGs (0,811; 0,796 y 0,855 respec-

**Tabla 1**  
**Descripción de la población. Distribución en grupos de sexo y edad.**  
**Promedios anuales de contactos con los servicios de salud y número de personas hospitalizadas**

Sexo	n	%
Mujeres	1.006.793	51,3%
Varones	957.544	48,7%
Grupos Edad	N	%
14-24	204.655	10,4%
25-34	350.084	17,8%
35-44	381.613	19,4%
45-54	330.617	16,8%
55-64	274.649	14,0%
65-69	101.113	5,2%
70-74	101.126	5,2%
75-79	95.499	4,9%
80-84	67.116	3,4%
85+	57.865	3,0%
Contactos de los pacientes con Osakidetza (año-1)		
	Promedio	
Nº de Hospitalizaciones	0,08	
Nº de Visitas a urgencias	0,26	
Nº de Visitas a consultas primaria + especializada	6,17	
Pacientes con una o más hospitalizaciones no programadas en el año 2		
	n	%
En los 6 primeros meses	58.250	2,97%
A lo largo de 12 meses	101.384	5,16%
Hospitalizaciones de larga estancia	27.758	1,41%

- Las hospitalizaciones excluyen las producidas por causas obstétricas o lesiones traumáticas.
- Año-1 corresponde al periodo de 01/09/2007 a 31/08/2008.
- Año-2 corresponde al periodo de 01/09/2008 a 31/08/2009.

**Tabla 2**  
**Comparación de resultados de modelos predictivos basados en el sistema ACGs, sin y con recalibración local: área bajo la curva (AUC), valor predictivo positivo (VPP) y sensibilidad**

Predicción de hospitalizaciones en los 6 meses siguientes a la clasificación de los pacientes				
	Modelo Out-of-the-box 0,811 (0,809-0,813)		Modelo recalibrado en País Vasco 0,823 (0,821-0,825)	
Banda de Riesgo	VPP	Sensibilidad	VPP	Sensibilidad
Percentil 99,9	0,539	0,018	0,540	0,018
Percentil 99,5	0,421	0,071	0,432	0,073
Percentil 99	0,347	0,117	0,362	0,122
Percentil 95	0,192	0,324	0,201	0,338
Percentil 90%	0,140	0,472	0,145	0,491
Predicción de hospitalizaciones en los 12 meses siguientes a la clasificación de los pacientes				
AUC (IC 95%)	Modelo Out-of-the-box 0,796 (0,794-0,798)		Modelo recalibrado en País Vasco 0,809 (0,807-0,811)	
Banda de Riesgo	VPP	Sensibilidad	VPP	Sensibilidad
Percentil 99,9	0,663	0,013	0,641	0,012
Percentil 99,5	0,555	0,054	0,574	0,056
Percentil 99	0,481	0,093	0,502	0,097
Percentil 95	0,294	0,285	0,309	0,300
Percentil 90%	0,223	0,433	0,232	0,450
Predicción de hospitalizaciones de larga estancia (estancia superior o igual a 12 días) en los 12 meses siguientes a la clasificación de los pacientes				
AUC (IC 95%)	Modelo Out-of-the-box 0,855 (0,852-0,858)		Modelo recalibrado en País Vasco 0,870 (0,867-0,873)	
Banda de Riesgo	VPP	Sensibilidad	VPP	Sensibilidad
Percentil 99,9	0,368	0,026	0,396	0,028
Percentil 99,5	0,287	0,101	0,300	0,105
Percentil 99	0,232	0,163	0,248	0,174
Percentil 95	0,117	0,411	0,128	0,450
Percentil 90%	0,080	0,563	0,087	0,610

tivamente). Observando los VPP y la sensibilidad, los modelos que presentan unos resultados mejores son los de pacientes ingresados en 12 meses: así, por ejemplo, identificando al 1% de la población con una predicción de riesgo más alta, más del 50% de las personas seleccionadas necesitarán al menos un ingreso y en este grupo estarán incluidos casi el 10% de las personas que serán hospitalizadas, mientras que seleccionando al 5% de la población con mayor riesgo, la sensibilidad y VPP son superiores al 30%.

Como se puede comprobar en la [tabla 3](#), los valores de AUC para las subpoblaciones de DM, EPOC e ICC son ligeramente más bajos que los de la población total (por ejemplo, para predicción de hospitalizaciones a 12 meses, 0,744 en diabetes; 0,759 en EPOC y 0,715 en insuficiencia cardiaca), y la realización de modelos que emplearon estimadores específicos de esa poblaciones no supusieron ninguna ventaja sobre el uso de los estimadores de la población general (respectivamente: 0,741 ; 0,750; 0,715) .

**Tabla 3**

**Comparación de resultados en subpoblaciones de pacientes con diabetes mellitus (DM), enfermedad pulmonar obstructiva crónicas (EPOC) e insuficiencia cardiaca (ICC) de modelos predictivos basados en el sistema ACGs: sin recalibración local, con valores recalibración de toda la población del País Vasco y recalibraciones específicas para esas subpoblaciones. Area bajo la curva ROC (IC 95%)**

	Out-of-the-box	Recalibración población total	Recalibración subpoblación
Predicción de hospitalizaciones en los 6 meses siguientes a la clasificación de los pacientes			
DM	0,741 (0,734-0,748)	0,754 (0,748-0,762)	0,749 (0,742-0,755)
EPOC	0,755 (0,744-0,763)	0,766 (0,755-0,773)	0,756 (0,745-0,767)
ICC	0,709 (0,701-0,716)	0,724 (0,716-0,731)	0,723 (0,716-0,730)
Predicción de hospitalizaciones en los 12 meses siguientes a la clasificación de los pacientes			
DM	0,730 (0,723-0,736)	0,744 (0,738-0,749)	0,741 (0,736-0,747)
EPOC	0,749 (0,740-0,757)	0,759 (0,750-0,768)	0,750 (0,743-0,759)
ICC	0,698 (0,691-0,704)	0,715 (0,709-0,721)	0,715 (0,708-0,721)
Predicción de hospitalizaciones de larga estancia (estancia superior o igual a 12 días) en los 12 meses siguientes a la clasificación de los pacientes			
DM	0,750 (0,742-0,758)	0,774 (0,767-0,782)	0,769 (0,760-0,779)
EPOC	0,765 (0,753-0,776)	0,781 (0,769-0,791)	0,766 (0,755-0,777)
ICC	0,703 (0,692-0,711)	0,732 (0,723-0,740)	0,730 (0,722-0,739)

Los promedios de coste observados en hospitalizaciones para los pacientes de diferentes bandas de riesgo se presentan en la **tabla 4**. Se observa que los grupos con mayor riesgo supusieron un mayor coste. Así, mientras que los pacientes que en función de su predicción se encontra-

ron por encima del percentil 90 de riesgo, requirieron en promedio 2.181,9 € en la atención hospitalaria de los 12 meses posteriores, los que quedaron por encima del P99,9 consumieron 10.691,3 €. Este mismo gradiente se repitió en las subpoblaciones de enfermos crónicos, en las

**Tabla 4**

**Promedios de coste (en euros) en hospitalizaciones en el año 2 de los pacientes que presentan mayor riesgo según predicciones en año 1 clasificados en bandas**

Banda de riesgo	Coste promedio hospitalizaciones 12 meses siguientes				Coste promedio por hospitalizaciones larga estancia			
	Población total	DM	EPOC	ICC	Población total	DM	EPOC	ICC
Percentil 99,9	10.691,3	15.849,1	12.730,3	14.373,4	8.916,6	11.373,8	12.797,5	9.876,1
Percentil 99,7	9.234,8	14.451,6	13.915,1	14.446,7	6.851,7	11.530,3	11.939,4	10.593,0
Percentil 99,5	8.086,0	13.105,1	14.145,7	14.472,7	5.796,1	10.740,0	11.434,2	11.321,1
Percentil 99	6.525,9	11.641,2	13.627,2	13.168,7	4.559,4	8.950,1	12.800,5	11.575,7
Percentil 95	3.202,8	7.215,5	11.165,7	9.945,9	2.061,5	5.270,3	8.746,9	7.945,2
Percentil 90%	2.181,9	5.841,9	9.508,0	8.379,1	1.334,2	4.092,9	7.153,8	6.159,7

Las bandas no son autoexcluyentes (por ejemplo, los pacientes incluidos en "≥Percentil 99,9" lo están también en las siguientes).

que los costes oscilaron entre 5.841,9 y 15.849,1 (DM); 9508,0 y 12.730,3 (EPOC); 8379,1 y 14.373,4 (ICC).

## DISCUSIÓN

El sistema de predicción de riesgo de hospitalizaciones en el País Vasco, basado en la clasificación de la población mediante ACGs con la adición de variables socioeconómicas (índice de privación del área de residencia), muestra un comportamiento adecuado. El AUC es el método empleado habitualmente para evaluar la capacidad predictiva de este tipo de modelos estadísticos. Un AUC de 0,5 indicaría que la capacidad del modelo no es mejor que la que se obtendría por simple azar, mientras que un valor de 1 supondría que la sensibilidad y especificidad son óptimas. Con frecuencia se considera que un modelo alcanza una capacidad discriminativa aceptable si sus cifras de AUC están entre 0,7 - 0,8 y es buena si supera el valor de 0,8<sup>3</sup>. Según este criterio, los resultados descritos resultan satisfactorios en las predicciones para la población general y aceptables para los subgrupos de pacientes afectados de patologías crónicas (diabetes mellitus, enfermedad pulmonar obstructiva crónica e insuficiencia cardíaca).

También resultan suficientes los valores de VPP y sensibilidad que ofrecen. Aunque distan de ser perfectos, estos modelos permiten clasificar a los pacientes según una escala de mayor a menor riesgo, la cual se asocia con un gradiente de número y coste de hospitalizaciones<sup>3,15</sup>. Este hecho hace posible la aplicación de las intervenciones más costosas a solamente un pequeño subgrupo de pacientes, mientras que otras menos intensivas puede proporcionarse a grupos más amplios. La comparación del coste de tales actividades, con el previsible ahorro que pueda conseguirse, permite hacer una proyección sobre su coste-efectividad y servirá de ayuda a la toma de decisiones sobre su implantación.

Nuestro estudio presenta algunas limitaciones. En primer lugar, los datos analizados pro-

ceden de los registrados en los sistemas de información sanitaria del País Vasco de forma habitual y, como sucede siempre que se recurre al empleo de bases de datos administrativas o anotaciones de las historias clínicas informatizadas, pueden contener algunos registros incompletos o inexactos<sup>16</sup>. En cualquier caso, si bien la calidad de la información puede afectar a los resultados, nuestra base de datos se alimenta de registros de hospitales, atención especializada y primaria, cuyo empleo complementario puede paliar algunas de estas deficiencias<sup>17</sup>. Por otra parte, si bien en un sistema de salud como el nuestro, que proporciona asistencia universal y financiada por impuestos, existen menos barreras de acceso que en otros, en la medida que determinados subgrupos tengan menor accesibilidad, sus necesidades de atención quedarán peor registradas. Por último, la variable socioeconómica empleada (índice de privación) puede disminuir la contribución individual de las características socioeconómicas, dado su carácter ecológico.

Hasta la fecha se han descrito diversos sistemas para predicción de hospitalizaciones. Muchos de ellos combinan variables demográficas, diagnósticos e ingresos previos<sup>18-19</sup>. En algunos casos añaden información relativa a prescripciones<sup>5</sup>, datos de laboratorio<sup>20</sup>, factores de riesgo y factores socioeconómicos<sup>21</sup>. Los modelos basados exclusivamente en el número previo de hospitalizaciones tienen una escasa validez, por el fenómeno de regresión a la media<sup>22</sup>. Así mismo, la capacidad de los clínicos para identificar correctamente a los pacientes que volverán a ser ingresados resulta muy inferior a la de los modelos estadísticos<sup>23</sup>. También se han estudiado modelos muy simples que incluyen solamente el número de diagnósticos o fármacos como únicas variables para medir la morbilidad<sup>24</sup>. Sin embargo, independientemente de su menor capacidad explicativa, su utilidad se verá limitada, pues no ofrecen información clínica suficiente.

Una exigencia para la puesta en marcha de intervenciones sanitarias es que benefi-

cien a toda la población en riesgo. Un modelo predictivo que no tenga en cuenta variables socioeconómicas puede provocar que las necesidades de los grupos más desfavorecidos se vean peor reflejadas y generar una sutil discriminación contra ellos en cuanto a su inclusión en programas de salud<sup>25</sup>. Por ese motivo, en nuestro modelo predictivo se emplea el índice de privación.

Además de los factores citados existen otros que también condicionan el riesgo de ingreso. Algunos tienen relación con las personas (como su grado de apoyo social, acceso a los servicios sanitarios o estado funcional) o con las organizaciones sanitarias (entre otros, coordinación de la atención o disponibilidad de camas de hospital)<sup>3</sup>. Aunque la inclusión de tal información pudiera mejorar la capacidad predictiva de los modelos, la aplicabilidad en el mundo real se verá condicionada por la disponibilidad de tales datos, que habitualmente no son registrados<sup>5,21</sup>.

Nuestro estudio ha permitido comprobar la validez de un sistema de predicción de riesgo de ingresos que resulta aplicable en nuestro medio, tanto en población general como en pacientes crónicos. Sin embargo, aunque tales modelos pueden resultar aptos para la identificación de pacientes candidatos a beneficiarse de programas proactivos, son necesarias nuevas investigaciones que evalúen la utilidad de la incorporación de nuevas variables<sup>3,26</sup> y el desarrollo de modelos dirigidos específicamente a la discriminación de hospitalizaciones evitables<sup>6</sup>.

### AGRADECIMIENTOS

Los autores agradecen a Gonzalo López Arbeloa (Organización Central de Osakidetza) su ayuda en la obtención de datos para este estudio.

Al equipo del proyecto MEDEA del País Vasco, por calcular el índice de privación y

aportar los datos del índice, y en especial a Montse Calvo, por el trabajo de geocodificación de los datos.

### BIBLIOGRAFÍA

1. Lewis G, Curry N, Bardsley M. Choosing a predictive risk model: a guide for commissioners in England. London: The Nuffield Trust; 2011.
2. López-Aguilá S, Contel JC, Farré J, Campuzano JL, Rajmil L. Predictive model for emergency hospital admission and 6-month readmission. *Am J Manag Care*. 2011; 17(9):e348-57.
3. Kansagara D, Englander H, Salanitro A, Kagen D, Theobald C, Freeman M, et al. Risk prediction models for hospital readmission: a systematic review. *JAMA*. 2011; 306(15):1688-98.
4. Billings J, Dixon J, Mijanovich T, Wennberg D. Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients. *BMJ*. 2006; 333:327-30.
5. Lemke KW, Weiner JP, Clark JM. Development and validation of a model for predicting inpatient hospitalization. *Med Care*. 2012; 50(2):131-9.
6. Lewis GH. "Impactability models": identifying the subgroup of high-risk patients most amenable to hospital-avoidance programs. *Milbank Q*. 2010; 88(2):240-55.
7. Nuño-Solinis R, Orueta JF, Mateos M. An Answer to Chronicity in the Basque Country: Primary Care-Based Population Health Management. *J Ambulatory Care Manage*. 2012; 35(3):167-73.
8. Orueta JF, Mateos Del Pino M, Barrio Beraza I, Nuño Solinis R, Cuadrado Zubizarreta M, Sola Sarabia C. Estratificación de la población en el País Vasco: resultados en el primer año de implementación. *Aten Primaria*. 2013; 45(1):54-60.
9. Johns Hopkins University, School of Public Health: The Johns Hopkins University ACG Case-Mix System. Baltimore (MD): The Johns Hopkins University. Disponible en: [http://acg.jhsph.org/index.php?option=com\\_content&view=article&id=46&Itemid=61](http://acg.jhsph.org/index.php?option=com_content&view=article&id=46&Itemid=61)
10. Domínguez-Berjón MF, Borrell C, Cano-Serral G, Esnaola S, Nolasco A, Pasarín MI, et al. Construcción de un índice de privación a partir de datos censales en grandes ciudades españolas (Proyecto MEDEA). *Gac Sanit*. 2008; 22(3):179-87.

11. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982; 143:29-36.
12. SAS version 9.2. Cary NC: SAS Institute; 2008.
13. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCr: visualizing classifier performance in R. *Bioinformatics*. 2005; 21(20):3940-1.
14. The R Development Core Team. R: A language and environment for statistical computing. Version 2.11.1. Vienna: R Foundation for Statistical Computing ; 2010.
15. Billings J, Blunt I, Steventon A, Georghiou T, Lewis G, Bardsley M.. Development of a predictive model to identify inpatients at risk of re-admission within 30 days of discharge (PARR-30). *BMJ Open*. 2012; 00:e001667.
16. Orueta JF, Nuño-Solinis R, Mateos M, Vergara I, Grandes G, Esnaola S. Predictive risk modelling in the Spanish population: a cross-sectional study. *BMC Health Serv Res*. 2013; 13:269.
17. Orueta JF, Nuño-Solinis R, Mateos M, Vergara I, Grandes G, Esnaola S.. Monitoring the prevalence of chronic conditions: which data should we use? *BMC Health Serv Res*. 2012; 12:365.
18. McAna JF, Crawford AG, Novinger BW, Sidorov J, Din FM, Maio V, et al. A predictive model of hospitalization risk among disabled medicaid enrollees. *Am J Manag Care*. 2013; 19(5):e166-74.
19. Holman CD, Preen DB, Baynham NJ, Finn JC, Semmens JB. A multipurpose comorbidity scoring system performed better than the Charlson index. *J Clin Epidemiol*. 2005; 58(10):1006-14.
20. Hippisley-Cox J, Coupland C. Predicting risk of emergency admission to hospital using primary care data: derivation and validation of QAdmissions score. *BMJ Open* 2013; 3:e003482.
21. Billings J, Georghiou T, Blunt I, Bardsley M. Choosing a model to predict hospital admission: an observational study of new variants of predictive models for case finding. *BMJ Open*. 2013; 3:e003352.
22. Lewis G. Predictive modeling in action: How 'virtual wards' help high-risk patients receive hospital care at home. New York: The Commonwealth Fund; 2010.   
D i s p o n i b l e e n :   
[http://www.commonwealthfund.org/~media/Files/Publications/Issue%20Brief/2010/Aug/1430\\_Lewis\\_predictive\\_modeling\\_in\\_action\\_virtual\\_wards\\_intl\\_brief.pdf](http://www.commonwealthfund.org/~media/Files/Publications/Issue%20Brief/2010/Aug/1430_Lewis_predictive_modeling_in_action_virtual_wards_intl_brief.pdf)
23. Allaudeen N, Schnipper JL, Orav EJ, Wachter RM, Vidyarthi AR. Inability of providers to predict unplanned readmissions. *J Gen Intern Med*. 2011; 26(7):771-6.
24. Quail JM, Lix LM, Osman BA, Teare GF. Comparing comorbidity measures for predicting mortality and hospitalization in three population-based cohorts. *BMC Health Serv Res*. 2011; 11:146.
25. Panattoni LE, Vaithianathan R, Ashton T, Lewis GH. Predictive risk modelling in health: options for New Zealand and Australia. *Aust Health Rev*. 2011; 35(1):45-51.
26. Wharam JF, Weiner JP. The promise and peril of healthcare forecasting. *Am J Manag Care*. 2012; 18(3):e82-5.