

Índices pronósticos de mortalidad en cuidados intensivos pediátricos

S. Prieto Espuñes^a, J. López-Herce Cid^b, C. Rey Galán^a, A. Medina Villanueva^a, A. Concha Torre^a y P. Martínez Cambolor^c

^aUnidad de Cuidados Intensivos Pediátricos. Hospital Universitario Central de Asturias. Universidad de Oviedo. Asturias. España. ^bUnidad de Cuidados Intensivos Pediátricos. Hospital General Universitario Gregorio Marañón. Madrid. España. ^cUnidad Académica Multidisciplinar de Agronomía y Ciencias. Universidad Autónoma de Tamaulipas. México.

Objetivo

Validar el *Pediatric Risk of Mortality score* (PRISM), el *Pediatric Index of Mortality* (PIM) y el PIM-2 en dos unidades de cuidados intensivos pediátricos españolas.

Pacientes y métodos

Se incluyeron en el estudio 241 niños críticos ingresados consecutivamente durante 6 meses. El funcionamiento general de los índices se evaluó con la Razón de Mortalidad Estandarizada (RME), comparando la mortalidad observada con la esperada por cada índice. La discriminación o capacidad del modelo para diferenciar entre pacientes que sobreviven y pacientes que fallecen se estimó por el área bajo la curva de rendimiento diagnóstico (curva ROC). La calibración o exactitud de las predicciones de riesgo se calculó con la prueba de bondad de ajuste de Hosmer-Lemeshow, en la que la calibración se evidencia por una $p > 0,05$.

Resultados

La tasa de mortalidad fue del 4,1%. El PRISM sobreestimó el número de muertes, con una RME de 0,44. La discriminación fue mejor para el PRISM y el PIM-2 que para el PIM (áreas bajo las curvas ROC: 0,883, 0,871 y 0,800, respectivamente), sin diferencias significativas. Finalmente, la calibración fue buena para el PIM-2 ($\chi^2(8) = 4,8730$; $p = 0,8461$) y para el PIM ($\chi^2(8) = 8,0876$; $p = 0,5174$), no encontrándose calibración estadística para el PRISM ($\chi^2(8) = 15,0281$; $p = 0,0133$).

Conclusiones

El PIM y el PIM-2 parecen ser más fiables, en términos de discriminación y calibración, que el PRISM en una muestra heterogénea de niños críticos españoles, aunque estos resultados deben ser confirmados en un estudio más amplio.

Palabras clave:

Pediatría. Medicina crítica. Unidades de cuidados intensivos pediátricos. Índices de gravedad. Modelos de predicción de mortalidad.

PROGNOSTIC INDEXES OF MORTALITY IN PEDIATRIC INTENSIVE CARE UNITS

Objective

To assess the validity of the Pediatric Risk of Mortality score (PRISM), the Pediatric Index of Mortality (PIM) and the PIM 2 in two Spanish pediatric intensive care units.

Patients and methods

We prospectively studied 241 critically ill children consecutively admitted over a 6-month period. The overall performance of the scoring systems was assessed by the Standardized Mortality Ratio (SMR), comparing observed deaths with expected deaths by each index. Discrimination (the ability of the model to distinguish between patients who live and those who die) was quantified by calculating the area under the receiver operating characteristic (ROC) curve. Calibration (the accuracy of mortality risk predictions) was calculated with the Hosmer-Lemeshow goodness-of-fit test, in which statistical calibration is evidenced by $p > 0.05$.

Results

The mortality rate was 4.1%. PRISM overestimated mortality (SMR = 0.44). Discrimination was better for PRISM and PIM 2 than for PIM (areas under ROC curves: 0.883, 0.871, and 0.800 respectively), with no significant differences. Finally, calibration was acceptable for PIM 2 ($\chi^2(8) = 4.8730$, $p = 0.8461$) and for PIM ($\chi^2(8) = 8.0876$, $p = 0.5174$), no statistical calibration was found for PRISM ($\chi^2(8) = 15.0281$, $p = 0.0133$).

Correspondencia: Dra. S. Prieto Espuñes.
Unidad de Cuidados Intensivos Pediátricos.
Hospital Universitario Central de Asturias.
Celestino Villamil, s/n. 33006 Oviedo. España.
Correo electrónico: msoledadp@yahoo.es.

Recibido en mayo de 2006.
Aceptado para su publicación en diciembre de 2006.

0.5174), but no statistical calibration was found for PRISM ($\chi^2(8) = 15.0281$, $p = 0.0133$).

Conclusions

PIM and PIM 2 showed better discrimination and calibration than PRISM in a heterogeneous group of children in Spanish critical care units. However, these results should be confirmed in a larger study.

Key words:

Pediatrics. Critical care. Pediatric intensive care units. Severity of illness index. Mortality. Prediction model.

INTRODUCCIÓN

Los índices pronósticos de mortalidad han sido desarrollados para cuantificar objetivamente la gravedad del paciente crítico, estimando la probabilidad de muerte que presenta según su estado clínico¹. Actualmente existen varios índices pronósticos de mortalidad disponibles para las unidades de cuidados intensivos pediátricos (UCIP). El más utilizado es el *Pediatric Risk of Mortality score* (PRISM), publicado en 1988² y que constituye el patrón de referencia de los sistemas de valoración de gravedad pediátricos³. Incluye 14 variables con 32 rangos, recogiendo el peor valor de cada una durante las primeras 24 h de cuidados intensivos. El riesgo de mortalidad se calcula mediante una ecuación de regresión logística en la que el valor del PRISM es una de las variables, junto con la edad y la existencia o no de intervención quirúrgica previa al ingreso. En 1996 se publicó una versión más sencilla del PRISM, el PRISM-III, que incluye 17 variables con 14 rangos⁴ y recoge el peor valor de cada una durante las primeras 12 o 24 h de cuidados intensivos. Esta versión se actualiza periódicamente mediante el reajuste de los coeficientes de su ecuación de probabilidad de muerte, pero para su utilización es necesaria una licencia renovable anualmente.

Pueden argumentarse ciertas objeciones al PRISM debido a su planteamiento, ya que en los pacientes más graves, que fallecen en las primeras 24 h, el peor valor de cada variable puede estar diagnosticando la muerte más que cuantificando la gravedad de la enfermedad⁵. Además el PRISM no discrimina la calidad de los cuidados intensivos administrados en las primeras 24 h; así, un paciente que reciba un tratamiento adecuado se recuperará más rápidamente y tendrá un PRISM menor que si recibiera un tratamiento menos correcto, lo cual empeoraría los parámetros valorados por el PRISM y éste indicaría aparentemente una mayor gravedad de su enfermedad⁵. Por otro lado, el PRISM mide el estado del paciente sin tener en cuenta si tiene alguna enfermedad de base, lo cual puede influir en su pronóstico³.

Para intentar corregir los problemas que origina el planteamiento del PRISM, en 1997 se publicó el *Paediatric Index of Mortality* (PIM)⁵. Este índice evalúa 8 variables, recogiendo el primer valor de cada una durante la

primera hora del ingreso. A cada uno de estos valores se le aplica un coeficiente para determinar la probabilidad de muerte. En 2003 el PIM fue actualizado para ajustarlo a la práctica intensiva más reciente⁶. El PIM-2 incluye 3 variables más, además de modificar ligeramente las variables del PIM y los coeficientes de la ecuación de probabilidad de muerte.

El PIM resulta más sencillo de aplicar que el PRISM, pero también presenta ciertos inconvenientes. En primer lugar, la primera medida de las constantes fisiológicas al ingreso en la UCIP puede ser muy variable y no reflejar la gravedad de la enfermedad, sino un estado transitorio relacionado con el traslado del paciente a la unidad^{7,8}. Además, la fórmula matemática aplicada para el cálculo de la probabilidad de muerte es muy compleja. Por último, la experiencia con este sistema es muy limitada, especialmente con el PIM-2.

El objetivo del presente estudio es comprobar la validez del PRISM, el PIM y el PIM-2 en una muestra heterogénea de niños críticos españoles.

PACIENTES Y MÉTODOS

Pacientes

El estudio se desarrolló en las UCIP de dos hospitales: la del Hospital General Universitario Gregorio Marañón de Madrid (HGUGM), unidad de nivel asistencial III que cuenta con 10 camas; y la del Hospital Universitario Central de Asturias (HUCA), unidad de nivel asistencial II que cuenta con 4 camas y en la que la cirugía cardiovascular pediátrica no está disponible. Ambas unidades admiten pacientes entre un mes y 16 años, aunque dependiendo de las necesidades asistenciales del hospital pueden ingresar excepcionalmente pacientes menores de un mes o mayores de 16 años, y están atendidas las 24 h del día por pediatras especializados en cuidados intensivos.

Se incluyeron en el estudio los pacientes críticamente enfermos ingresados de forma consecutiva desde el 1 de noviembre de 2003 hasta el 30 de junio de 2004. Se excluyeron los pacientes que cumplían alguno de los siguientes criterios: *a*) pacientes ingresados para control de sedación durante la realización de procedimientos diagnósticos o terapéuticos; *b*) pacientes con estancia menor de 2 h, y *c*) pacientes trasladados a otra UCIP para recibir cuidados que no se podían administrar en la UCIP estudiada, ya que en estos casos no se podía conocer la evolución del paciente.

Metodología

Se diseñó un estudio de cohortes prospectivo y observacional. De cada paciente se recogieron los siguientes datos: edad, sexo, diagnóstico de ingreso, tiempo de estancia, estado al alta de la UCIP (vivo, muerto), así como las variables del PRISM, del PIM y del PIM-2. Para definir

el diagnóstico de ingreso se tuvo en cuenta la disfunción que originó la necesidad de cuidados intensivos, según las Normas de la Sociedad de Cuidados Intensivos de la Academia Americana de Pediatría⁹, añadiendo un apartado para politraumatismos con o sin traumatismo craneoencefálico. Para calcular la probabilidad de muerte de los pacientes se aplicaron las ecuaciones descritas por Pollack et al² para el PRISM, por Shann et al⁵ para el PIM y por Slater et al⁶ para el PIM-2.

Análisis estadístico

Los datos fueron procesados mediante el programa informático SPSS (Statistical Program for Social Science), versión 11.0. Para expresar las características de los pacientes se han utilizado estadísticos descriptivos simples, como media \pm desviación estándar, medianas y porcentajes. Para las comparaciones se utilizaron pruebas no paramétricas por no presentar las variables una distribución normal, aceptando como nivel de significación un valor de $p < 0,05$. La comparación de las curvas ROC se realizó con el programa informático EpiDat (Análisis Epidemiológico de Datos Tabulados), versión 3.1.

Para evaluar los índices pronósticos de mortalidad se analizaron tres aspectos: el funcionamiento general, la discriminación y la calibración. El funcionamiento general se estimó a través de la Razón de Mortalidad Estandarizada (RME), que compara la mortalidad observada en la muestra con la esperada por cada modelo (RME = mortalidad observada/mortalidad esperada). La discriminación (capacidad del índice para diferenciar entre los pacientes que sobreviven y los que fallecen) se calculó midiendo el área bajo la curva de rendimiento diagnóstico (curva ROC)¹⁰, de forma que el mayor valor del área bajo la curva ROC se corresponde con la mejor capacidad de discriminación; se considera que un área de 0,70-0,79 es aceptable, de 0,80-0,89 es buena y de 0,90 o más es excelente. La calibración (grado de exactitud de las predicciones de riesgo realizadas por el modelo) se estimó con la prueba de bondad de ajuste de Hosmer-Lemeshow¹¹, que consiste en comparar el número de muertes observado con el número de muertes esperado y el número de supervivientes observado con el número de supervivientes esperado en 10 grupos de riesgo de mortalidad crecientes o deciles de riesgo; la calibración estadística se evidencia por un valor de $p > 0,05$, de forma que a mayor valor de p , mejor ajuste del modelo.

RESULTADOS

Desde el 1 de noviembre de 2003 hasta el 30 de junio de 2004 fueron ingresados en las unidades de estudio 245 pacientes, de los cuales se excluyeron 4 pacientes del HUCA que fueron trasladados a otra UCIP. Por tanto, se analizaron los datos de 241 pacientes, 140 procedentes del HGUGM y 101 procedentes del HUCA. La edad media fue de $62,6 \pm 61,8$ meses (mediana 41,3 meses), con una rela-

ción varón:mujer de 1,3:1. Los distintos diagnósticos por los que los pacientes precisaron ingreso aparecen reflejados en la figura 1. La estancia media fue $7,3 \pm 12,6$ días, con una estancia mediana de 4 días. La tasa global de mortalidad fue del 4,1% (10 pacientes: 9 en el HGUGM y uno en el HUCA); el 30% de los pacientes fallecidos lo fueron en las primeras 24 h del ingreso.

Los pacientes fallecidos presentaron una probabilidad de muerte calculada por el PRISM, el PIM y el PIM-2 significativamente mayor que los pacientes supervivientes (tabla 1). Pero al analizar el funcionamiento general de cada índice, se observaron diferencias entre los tres índices pronósticos de mortalidad. Así, la RME fue 0,44 para el PRISM (mortalidad esperada: 22,8 pacientes; intervalo de confianza del 95% [IC 95%]: 0,17-0,71), 0,86 para el PIM (mortalidad esperada: 11,7 pacientes; IC 95%: 0,33-1,39) y 0,91 para el PIM-2 (mortalidad esperada: 11 pacientes; IC 95%: 0,35-1,48). La discrimina-

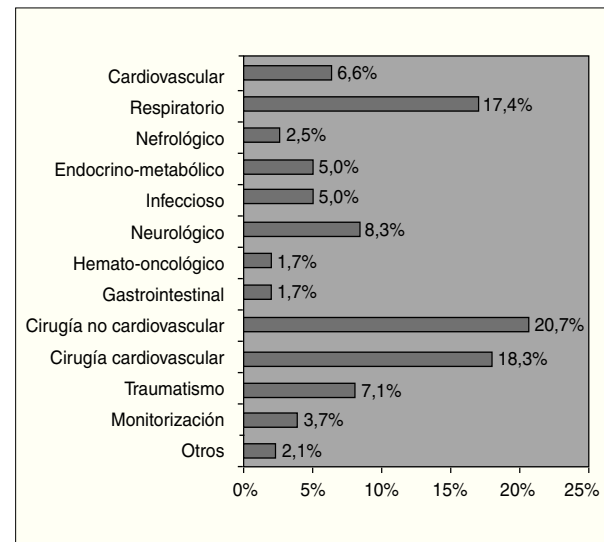


Figura 1. Distribución de los pacientes por grupo diagnóstico.

TABLA 1. Comparación de los distintos índices pronósticos de mortalidad entre pacientes supervivientes y fallecidos

		Supervivientes	Muertes	p
Riesgo de mortalidad PRISM	Media \pm DE	8,4 \pm 15,9	40,8 \pm 37,3	< 0,001
	Mediana	2,3	27,7	
	Rango	0,2-96,7	7,0-98,3	
Riesgo de mortalidad PIM	Media \pm DE	4,1 \pm 7,7	23,0 \pm 27,7	0,001
	Mediana	1,4	8,6	
	Rango	0,2-79,0	0,9-75,4	
Riesgo de mortalidad PIM-2	Media \pm DE	3,7 \pm 8,4	25,1 \pm 30,7	< 0,001
	Mediana	1,4	8,9	
	Rango	0,1-84,0	1,7-82,7	

DE: desviación estándar; p: nivel de significación estadística; PRISM: *Pediatric Risk of Mortality score*; PIM: *Pediatric Index of Mortality*.

ción medida por el área bajo la curva ROC fue buena para los tres índices (fig. 2), aunque no se encontraron diferencias estadísticamente significativas entre las

3 áreas ($\chi^2 (2) = 1,23$; $p = 0,27$). Por último, no se observó calibración estadística para el PRISM, pero sí para el PIM y el PIM-2; los datos se muestran en la tabla 2.

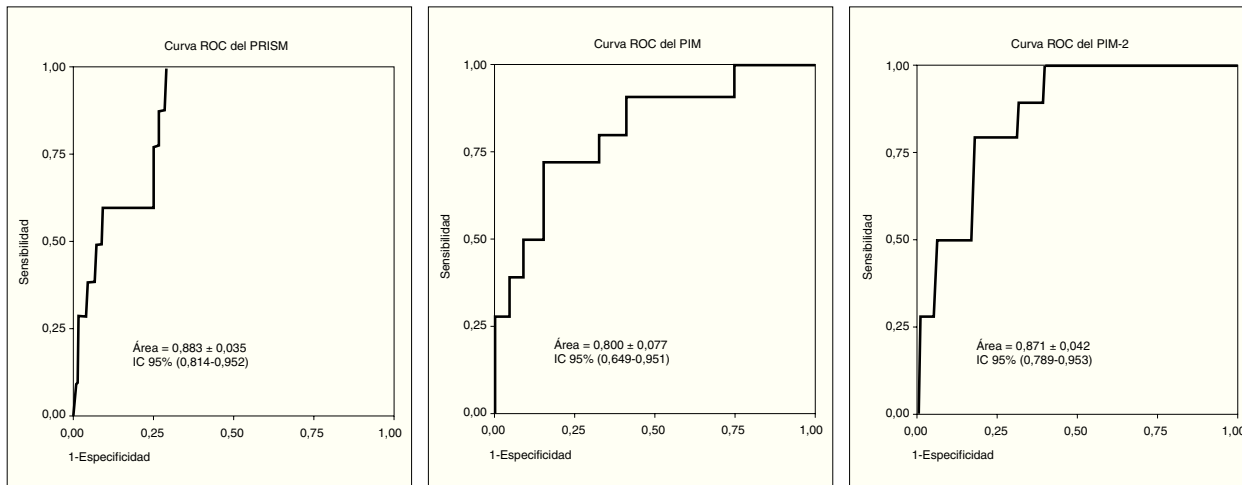


Figura 2. Curvas de rendimiento diagnóstico (curvas ROC) calculadas para cada índice pronóstico de mortalidad, especificando el área bajo la curva con el error estándar y el intervalo de confianza del 95 %.

TABLA 2. Tabla de Hosmer-Lemeshow comparando la mortalidad y la supervivencia observadas con las esperadas por cada modelo en los distintos deciles de riesgo

PRISM: Valor del estadístico de Hosmer-Lemeshow ($\chi^2 (8) = 15,0281$; $p = 0,0133$)										
Deciles de riesgo	1	2	3	4	5	6	7	8	9	10
Fallecimientos										
Observados	0	0	0	0	0	0	0	4	1	5
Esperados	0,10	0,16	0,24	0,27	0,62	0,72	1,12	1,96	4,37	13,22
Supervivientes										
Observados	25	23	23	18	29	23	22	20	23	18
Esperados	24,90	22,84	22,76	17,73	28,39	22,28	20,88	22,04	19,63	9,78
Total	25	23	23	18	29	23	22	24	24	23
PIM: Valor del estadístico de Hosmer-Lemeshow ($\chi^2 (8) = 8,0876$; $p = 0,5174$)										
Deciles de riesgo	1	2	3	4	5	6	7	8	9	10
Fallecimientos										
Observados	0	0	1	0	0	1	1	0	3	4
Esperados	0,06	0,19	0,21	0,27	0,33	0,42	0,60	1,09	1,90	6,62
Supervivientes										
Observados	24	24	22	25	24	23	22	25	21	20
Esperados	23,94	23,81	22,79	24,73	23,67	23,58	22,40	23,91	22,10	17,38
Total	24	24	23	25	24	24	23	25	24	24
PIM-2: Valor del estadístico de Hosmer-Lemeshow ($\chi^2 (8) = 4,8730$; $p = 0,8461$)										
Deciles de riesgo	1	2	3	4	5	6	7	8	9	10
Fallecimientos										
Observados	0	0	0	0	0	1	1	0	3	5
Esperados	0,04	0,07	0,15	0,23	0,30	0,37	0,54	0,97	1,72	6,58
Supervivientes										
Observados	24	24	24	24	24	23	23	24	21	19
Esperados	23,96	23,93	23,85	23,77	23,70	23,63	23,46	23,03	22,28	17,42
Total	24	24	24	24	24	24	24	24	24	24

DISCUSIÓN

Los índices pronósticos de mortalidad han sido diseñados para estimar la gravedad de los pacientes ingresados en las UCIP, pero también permiten evaluar la calidad médica de los cuidados administrados mediante la comparación de la mortalidad real con la esperada por la inestabilidad fisiológica de los enfermos¹²⁻¹⁴. Además, se emplean para estratificar a los pacientes en los estudios clínicos y pueden formar parte de las guías clínicas de tratamiento⁸.

Existen varios factores que pueden influir en la validez de los índices de mortalidad. Estos modelos estadísticos se desarrollan a partir de los datos recogidos en determinadas unidades durante determinado período de tiempo. Si la muestra de pacientes sobre la que se aplica un índice presenta características diferentes a la muestra original, el ajuste del modelo puede no ser adecuado y por tanto la mortalidad calculada no será fiable^{15,16}. Por otra parte, la evolución de los cuidados intensivos administrados ha mejorado el pronóstico de los pacientes¹⁷, por lo que un índice basado en datos de varios años atrás podría calcular una mortalidad excesiva para la calidad actual de los cuidados intensivos¹⁸. Por ello se deben validar los modelos de predicción del riesgo de mortalidad antes de poder utilizarlos en una población. En este estudio hemos intentado averiguar cuál de los índices pronósticos de mortalidad disponibles actualmente para los niños críticamente enfermos es el más útil para nuestros pacientes, aunque el análisis estadístico no tiene fuerza suficiente para extraer conclusiones definitivas debido al pequeño tamaño muestral y al escaso número de muertes observado. Estos dos factores también limitan el análisis de los índices en cada una de las UCIP incluidas y su comparación.

En el presente estudio, los tres índices pronósticos de mortalidad evaluados (PRISM, PIM y PIM-2) resultaron útiles para predecir la mortalidad, ya que la probabilidad de muerte calculada para los pacientes fallecidos fue significativamente superior a la calculada para los supervivientes con los tres sistemas. Pero el funcionamiento general estimado por la RME mostró que la mortalidad esperada por el PRISM fue dos veces superior al número de pacientes fallecidos, mientras que el PIM y el PIM-2 estimaron mejor el número de muertes. La RME realmente compara el número de muertes encontrado en el estudio con el número de muertes esperado si los mismos pacientes hubieran recibido los cuidados administrados en las unidades de las que se derivaron los índices en el momento en que éstos fueron desarrollados¹⁸. Para generar el PRISM se emplearon los datos recogidos en 4 UCIP estadounidenses desde 1980 hasta 1985². Para el PIM se utilizaron datos de 4 UCIP australianas desde 1988 hasta 1996⁵. Y para el PIM-2 los datos se recogieron en 7 UCIP, cuatro en Australia, dos en el Reino Unido y una en Nueva Zelanda, desde 1997 hasta 1999⁶. Así, la mortalidad

calculada por el PRISM en nuestra muestra es la que se obtendría con la calidad de cuidados intensivos estadounidenses de la década de 1980. La evolución de la medicina intensiva ha permitido un mejor tratamiento para los pacientes críticos, de forma que la supervivencia es mayor ahora que hace una década¹⁷. Por esta razón, para medir la gravedad de los pacientes críticamente enfermos se deberían utilizar modelos más recientes, como el PRISM-III, el PIM o el PIM-2.

La discriminación y la calibración son más importantes que la RME en la evaluación de los índices pronósticos de mortalidad. Estos parámetros miden aspectos diferentes de la validez de los modelos de predicción del riesgo de mortalidad: la discriminación examina la sensibilidad y la especificidad del sistema, mientras que la calibración determina la exactitud de las predicciones de riesgo realizadas por el modelo. En nuestro estudio, los tres índices presentaron una discriminación buena (fig. 2), pero el PIM mostró un área bajo la curva ROC menor que el PRISM y el PIM-2, sin ser esta diferencia estadísticamente significativa. En cuanto a la calibración, al aplicar la prueba de Hosmer-Lemeshow tanto el PIM como el PIM-2 presentaron un buen ajuste a la muestra de pacientes ($p = 0,5174$ y $p = 0,8461$, respectivamente), mientras que no se encontró calibración estadística para el PRISM ($p = 0,0133$). La prueba de Hosmer-Lemeshow fue diseñada para calcular la bondad del ajuste de los modelos estadísticos basados en la regresión logística¹¹, pero presenta limitaciones en las poblaciones pequeñas como la del presente estudio. En estos casos, el examen de los valores de la tabla de Hosmer-Lemeshow ofrece más información y esta es más fiable que el simple valor de la p ¹⁸. La tabla 2 expone los valores obtenidos en nuestro estudio. Las 5 columnas situadas a la izquierda representan los pacientes con menor probabilidad de muerte. Al agrupar los valores en estas categorías se observa que el PRISM sobreestima ligeramente la mortalidad observada, mientras que el PIM y el PIM-2 predicen mejor el número de muertes y supervivientes (mortalidad observada/mortalidad esperada: PRISM = 0/1,4; PIM = 1/1,1; PIM-2 = 0/0,8). En las 5 columnas situadas a la derecha, que representan a los pacientes con mayor probabilidad de muerte, la mortalidad calculada por el PRISM duplica a la mortalidad real, mientras que las predicciones realizadas por el PIM-2 son las más ajustadas a la realidad (mortalidad observada/mortalidad esperada: PRISM = 10/21,4; PIM = 9/10,6; PIM-2 = 10/10,2).

Así pues, en este estudio el PIM y el PIM-2 presentan una discriminación y una calibración buenas, mientras que el PRISM muestra una discriminación buena sin calibración estadística. Las discrepancias entre la calibración y la discriminación de los índices pronósticos de mortalidad ya han sido previamente descritas. En otros estudios que comparan el PRISM y el PIM, ambos modelos presentaron buena discriminación pero mala calibración^{19,20}.

La discriminación de un sistema depende de su capacidad de valoración de la inestabilidad fisiológica del paciente, mientras que la calibración depende de que la ecuación utilizada para calcular la probabilidad de muerte sea válida para la muestra estudiada. Si un sistema presenta una buena discriminación pero una mala calibración, ésta se podría mejorar cambiando los coeficientes de la ecuación de probabilidad de muerte para que ésta se ajuste mejor a la población estudiada^{21,23}. Aunque de esta forma se inutilizan los índices pronósticos de mortalidad para evaluar la calidad de los cuidados intensivos administrados o para realizar comparaciones entre distintas UCIP, el riesgo de mortalidad calculado sería absolutamente fiable.

En resumen, el PIM y el PIM-2 presentan un ajuste adecuado, en términos de discriminación y calibración, para estimar el riesgo de mortalidad de una muestra heterogénea de niños críticamente enfermos. Sin embargo, estos resultados deben ser confirmados en un estudio más amplio.

BIBLIOGRAFÍA

- Gunning K, Rowan K. ABC of intensive care: Outcome data and scoring systems. *BMJ*. 1999;319:241-4.
- Pollack MM, Ruttimann UE, Getson PR. Pediatric risk of mortality (PRISM) score. *Crit Care Med*. 1988;16:1110-6.
- Prieto Espuñes S, Rey Galán C, Medina Villanueva A. Sistemas de valoración de gravedad en cuidados intensivos pediátricos. En: López-Herce Cid J, Calvo Rey C, Lorente Acosta MJ, Baltodano Agüero A, editores. *Manual de cuidados intensivos pediátricos*. 2ª ed. Madrid: Publimed; 2004. p. 46-52.
- Pollack MM, Patel KM, Ruttimann UE. PRISM-III: An updated Pediatric Risk of Mortality score. *Crit Care Med*. 1996;24:743-52.
- Shann F, Pearson G, Slater A, Wilkinson K. Paediatric index of mortality (PIM): A mortality prediction model for children in intensive care. *Intensive Care Med*. 1997;23:201-7.
- Slater A, Shann F, Pearson G. PIM2: A revised version of the Paediatric Index of Mortality. *Intensive Care Med*. 2003;29: 278-85.
- García Sanz C, Rupérez Lucas M, López-Herce Cid J, Vigil Escribano D, Garrido Cantarero G. Valor pronóstico de la puntuación PIM (índice pediátrico de mortalidad) y del ácido láctico en niños críticamente enfermos. *An Esp Pediatr*. 2002;57: 394-400.
- Marcin JP, Pollack MM. Review of the methodologies and applications of scoring systems in neonatal and pediatric intensive care. *Pediatr Crit Care Med*. 2000;1:20-7.
- Guidelines for developing admission and discharge policies for the pediatric intensive care unit. American Academy of Pediatrics. Committee on Hospital Care and Section of Critical Care. Society of Critical Care Medicine. Pediatric Section Admission Criteria Task Force. *Pediatrics*. 1999;103:840-2.
- Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143:29-36.
- Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol*. 1982;115:92-106.
- Gemke RJ, Bonsel GJ, Van Vught AJ. Effectiveness and efficiency of a Dutch pediatric intensive care unit: Validity and application of the Pediatric Risk of Mortality score. *Crit Care Med*. 1994;22:1477-84.
- Martinot A, Leteurtre S, Grandbastien B, Duhamel A, Leclerc F, et le groupe francophone de Réanimation et urgences pédiatriques. Caractéristiques des patients et utilisation des ressources dans les services de réanimation pédiatrique français. *Arch Pediatr*. 1997;4:730-6.
- Prieto Espuñes S, Medina Villanueva A, Concha Torre A, Rey Galán C, Menéndez Cuervo S, Crespo Hernández M. Asistencia a los niños críticamente enfermos en Asturias: características y efectividad. *An Esp Pediatr*. 2002;57:22-8.
- Murphy-Filkins R, Teres D, Lemeshow S, Hosmer DW. Effect of changing patient mix on the performance of an intensive care unit severity-of-illness model: How to distinguish a general from a specialty intensive care unit. *Crit Care Med*. 1996; 24:1968-73.
- Moreno R, Apolone G, Miranda DR. Evaluation of the uniformity of fit of general outcome prediction models. *Intensive Care Med*. 1998;24:40-7.
- Tilford JM, Roberson PK, Lensing S, Fiser DH. Differences in pediatric ICU mortality risk over time. *Crit Care Med*. 1998;26: 1737-43.
- Shann F. Are we doing a good job: PRISM, PIM and all that. *Intensive Care Med*. 2002;28:105-7.
- Tibby SM, Taylor D, Festa M, Hanna S, Hatherill M, Jones G, et al. A comparison of three scoring systems for mortality risk among retrieved intensive care patients. *Arch Dis Child*. 2002; 87:421-5.
- Leteurtre S, Leclerc F, Wirth J, Noizet O, Magnenant E, Sadik A, et al. Can generic paediatric mortality scores calculated 4 hours after admission be used as inclusion criteria for clinical trials? *Crit Care* [en línea] 21 mayo 2004 [fecha de acceso 10 de agosto de 2004] 8: R185-R193. URL disponible en: <http://ccforum.com/content/8/4/R185>.
- Moreno R, Apolone G. Impact of different customization strategies in the performance of a general severity score. *Crit Care Med*. 1997;25:2001-8.
- Zhu BP, Lemeshow S, Hosmer DW, Klar J, Avrunin J, Teres D. Factors affecting the performance of the models in the Mortality Probability Model II system and strategies of customization: A simulation study. *Crit Care Med*. 1996;24:57-63.
- Livingston BM, MacKirdy FN, Howie JC, Jones R, Norrie JD. Assessment of the performance of five intensive care scoring models within a large Scottish database. *Crit Care Med*. 2000; 28:1820-7.