

Validación de los *Mortality Probability Models II* al ingreso (MPM II-0), a las 24 horas (MPM II-24), y a las 48 horas (MPM II-48) comparados con las predicciones de mortalidad hospitalaria de APACHE II y SAPS II realizadas en los días 1 y 2 de estancia en UCI*

N. SERRANO HERNÁNDEZ*, A. GARCÍA DE LORENZO Y MATEOS**, M.L. MORA QUINTERO*
Y J. FEDRIANI GORRÍA***

Servicio de Medicina Intensiva. *Hospital Universitario de Canarias. La Laguna. Tenerife.
Hospital Universitario La Paz. Madrid. *Hospital General de Albacete, Albacete.

Fundamento. Comparar la certeza de las predicciones de mortalidad hospitalaria de la secuencia de modelos MPM II al ingreso, a las 24, y a las 48 horas de ingreso con la de las predicciones obtenidas a partir de los modelos APACHE II y SAPS II medidos consecutivamente los días primero y segundo de ingreso en UCI en pacientes polivalentes.

Métodos. Determinación prospectiva de los valores de probabilidad individual de muerte definida por los modelos MPM II-0, MPM II-24, y MPM II-48; y APACHE II y SAPS II a las 24 horas y a las 48 horas de ingreso en UCI, y de la evolución final, expresada como mortalidad o supervivencia al alta hospitalaria. Análisis de calibración y discriminación mediante pruebas de bondad del ajuste y medida del estadístico χ^2 de Lemeshow-Hosmer, y tablas de clasificación y medida del área bajo la curva ROC (AUC) de cada modelo, respectivamente.

Resultados. Fueron estudiados 202 pacientes críticos en el primer día de estancia y 149 pacientes restantes en el segundo día. La mortalidad al alta hospitalaria fue del 25,7% del total, siendo el 19,8% intra-UCI. Las predicciones de MPM II-0, MPM II-24, APACHE II-24 y SAPS II-24 sobre 202 pacientes tuvieron unos valores de χ^2 : 5,54; 10,77; 8 y 11,7, con unos valores de p: 0,69; 0,21; 0,40 y 0,16 respectivamente, y los valores de AUC para los mismos índices fueron: 0,762; 0,791; 0,763; y 0,798. Las predicciones de MPM II-48, APACHE II-48 y SAPS II-48 sobre 149 pacientes a las 48 horas de ingreso tuvieron unos valores de χ^2 : 4,94; 13,4; y 5,8, con unos valores de p: 0,76; 0,09; y 0,66 respectivamente, y los valores de AUC para los mismos índices fueron 0,759; 0,728; y 0,788.

Conclusiones. La secuencia de modelos MPM II medidos al ingreso, a las 24 y a las 48 horas proporciona unas predicciones de mortalidad hospitalaria ligeramente superiores a las de APACHE II y SAPS II por separado usados secuencialmente a las 24 y 48 horas. Sin embargo, la utilización combinada de APACHE II en las primeras 24 horas y del SAPS II en el segundo día podría proporcionar resultados similares a los obtenidos con la secuencia de modelos MPM II-0, MPM II-24, y MPM II-48.

*Este estudio ha sido financiado a través de una beca de Investigación de la Consejería de Sanidad de la Junta de Comunidades de Castilla-La Mancha.

Correspondencia: Dr. N. Serrano Hernández.
Servicio de Medicina Intensiva.
Proyecto EPICURE (www.epicure.org)
Hospital Universitario de Canarias.
Buzón 273.
38320 La Laguna. Tenerife.

Manuscrito aceptado el 9-XII-1999.

PALABRAS CLAVE: mortalidad, pronóstico, bondad del ajuste, APACHE, SAPS, MPM.

VALIDATION OF MORTALITY PROBABILITY MODELS II (MPM II) AT ADMISSION (MPM II-0), AT 24 HOURS (MPM II-24), AND AT 48 HOURS (MPM II-48) COMPARED WITH THE HOSPITAL MORTALITY PREDICTIONS FROM APACHE II AND SAPS II MEASURED IN THE FIRST AND SECOND DAYS OF ICU STAY

Objective. To compare the accuracy of the hospital mortality predictions from the MPM II models at ICU admission, at 24, and at 48 hours of ICU admission with those obtained from the APACHE II and SAPS II models measured in the first and the second days of ICU stay after admission of multidisciplinary critically ill patients.

Methods. Prospective calculations of individual probabilities of death defined by the models MPM II-0, MPM II-24, and MPM II-48; and also by the APACHE II and SAPS II models at 24 and 48 hours of ICU admission. Actual outcomes were recorded as death or survival at hospital discharge. Calibration and discrimination analysis by Lemeshow-Hosmer's goodness-of-fit chi-square statistics, classification tables, and measurements of the area under the ROC curve (AUC), were performed for each model.

Results. Two hundred and two (202) consecutive critically ill patients were studied in the first ICU day, and 149 remaining patients were studied in the 2nd day. Mortality rate at hospital discharge was 25.7%, while in-ICU mortality was 19.8%. Mortality predictions from MPM II-0, MPM II-24, APACHE II-24 and SAPS II-24 on 202 patient had the following chi-square values: 5.54; 10.77; 8 and 11.7, with p values: 0.69; 0.21; 0.40 and 0.16 respectively, and the AUC values in the same order were: 0.762; 0.791; 0.763; and 0.798.

Mortality predictions from MPM II-48, APACHE II-48 and SAPS II-48 on 149 patients at 48 hours of ICU admission had the following chi-square values: 4.94; 13.4; and 5.8, with p-values: 0.76; 0.09; and 0.66 respectively, and the AUC values in the same order were 0.759; 0.728; and 0.788.

Conclusions. The sequence of MPM II models measured at admission, at 24 hours, and at 48 hours of ICU stay provides slightly superior hospital mortality predictions than either APACHE II or SAPS II used at 24 and 48 hours. However, combined utilization of APACHE II in the first 24 hours with SAPS II in the second day might provide similar results than those obtained by the MPM II-0, MPM II-24, and MPM II-48 models.

KEY WORDS: Mortality, prognosis, outcome goodness-of-fit, APACHE, SAPS, MPM.

(*Med Intensiva* 2000; 24: 49-60)

INTRODUCCIÓN

En los últimos años se han desarrollado modelos predictivos de mortalidad hospitalaria en el paciente crítico que establecen sus pronósticos de forma dinámica a lo largo de la evolución clínica¹. Representando estas tendencias se encuentran los mo-

delos denominados *Mortality Probability Models* (MPM)²⁻¹⁰. Frente a ellos, los modelos tradicionales APACHE¹¹⁻¹³ y SAPS^{14,15} representan los planteamientos clásicos, en los que el establecimiento del pronóstico se realiza a partir de los datos recogidos en las primeras 24 horas de ingreso en Unidad de Cuidados Intensivos (UCI), aunque en algún caso también ha sido propuesto su uso para más allá de las primeras 24 horas, con intención de optimizar su capacidad predictiva¹⁶.

Desde su presentación en 1985⁷, se han publicado cuatro variantes de MPM bajo la misma denominación, por lo que se ha producido durante años una notable confusión en relación con este modelo. Por este motivo, nos ha parecido apropiado incluir un anexo (Apéndice I) con una breve descripción de la evolución histórica de éste y también del resto de los modelos predictivos que incluiremos en este estudio.

Nuestro objetivo final ha sido comparar la capacidad predictiva de mortalidad hospitalaria de la secuencia de modelos MPM II al ingreso, a las 24, y a las 48 horas de ingreso, con la de las predicciones obtenidas de los modelos APACHE II y SAPS II medidos consecutivamente los días primero y segundo de ingreso en UCI, y el potencial grado de refinamiento alcanzado por las predicciones hechas por cada modelo a lo largo del tiempo.

MATERIAL Y MÉTODOS

La recogida de datos se llevó a cabo de forma prospectiva durante doce meses en la Unidad de Cuidados Intensivos Polivalentes del Hospital General de Albacete. No se incluyeron en el estudio los pacientes de causa cardiológica (coronaria, marcapasos, etc.) atendidos en la Unidad de Cuidados Coronarios. Los pacientes fueron estudiados desde su ingreso en la UCI hasta el alta hospitalaria. La variable dependiente en todos los casos fue la evolución expresada en términos de mortalidad o supervivencia al alta hospitalaria. El conjunto de variables demográficas, clínicas y de laboratorio necesarias para el cálculo de los diferentes índices predictivos incluidos en el estudio fueron recogidas para cada paciente al ingreso y en las primeras 24 y 48 horas de estancia en UCI. A todos los pacientes de la cohorte en estudio les fueron determinados:

1. Los valores de probabilidad individual de muerte definida por los modelos MPM II-0³, MPM II-24³, y MPM II-48².
2. Los valores de probabilidad individual de mortalidad hospitalaria definida y por APACHE II y SAPS II a las 24 horas según la metodología descrita por sus respectivos autores^{12,15}.
3. Los valores de probabilidad individual de mortalidad hospitalaria calculados mediante APACHE II y SAPS II a las 48 horas con una metodología idéntica a la utilizada en las primeras 24 horas, asumiendo que esta variación no está recogida en el diseño original de estos modelos, aunque ha sido aplicada previamente por otros^{16,17}.

4. La evolución final, expresada como mortalidad o supervivencia al alta hospitalaria. Durante el episodio en UCI la evolución fue registrada para cada paciente, y se realizó un seguimiento de los pacientes supervivientes hasta el final del episodio hospitalario.

El análisis estadístico de los datos se ha llevado a cabo mediante el programa SPSS 8.0 para Windows. Dado que no existe una única medida estándar para describir la bondad del ajuste o capacidad predictiva de estos modelos de regresión logística^{18,19}, la *calibración* (exactitud de las predicciones de riesgo hechas por los modelos) fue comparada usando los estadísticos χ^2 propuestos por Lemeshow y Hosmer¹⁸, y la *discriminación* (capacidad de los modelos para discriminar entre pacientes que sobreviven y pacientes que fallecen) fue comparada usando tablas de clasificación en el punto de corte del 50% y también comparando el área bajo la curva ROC de cada modelo.

Las pruebas de bondad del ajuste están consideradas como la herramienta más certera para valorar la idoneidad de un modelo predictivo⁵, y bajo esta premisa han sido adoptadas como un criterio mayor en nuestro estudio comparativo. Las probabilidades individuales de mortalidad obtenidas por los modelos a las 0, 24 y 48 horas fueron ordenadas según valores crecientes y agrupadas en estratos de riesgo de mortalidad predicha con un número similar de pacientes cada uno de ellos (tablas 1 y 2). El número de fallecimientos esperados en cada estrato o decil corresponde a la suma de las probabilidades individuales de muerte de los pacientes en dicho estrato. El número de supervivientes esperados se calcula por diferencia. El estadístico χ^2 de Lemeshow-Hosmer es la suma, para todos los estratos, de los cocientes de valores $[\text{Observados}-\text{Esperados}]^2/\text{Esperados}$.

Los grados de libertad son $n - k - 1$, siendo n = número de estratos (generalmente se adoptan 10, como en este caso), y k = número de variables, en este caso sólo una: el resultado final del episodio hospitalario en términos de mortalidad o supervivencia¹⁸. La prueba de bondad de ajuste de Lemeshow-Hosmer que resulte con menor valor χ^2 indicará mejor calibración para ese índice¹⁸, y un valor de p superior a 0,05 validará el modelo al indicar que no existen diferencias estadísticamente significativas entre lo observado y lo predicho¹⁸. Como el valor del estadístico χ^2 depende del número de sujetos de la muestra²⁰, para la comparación de resultados entre los días primero y segundo, con distinto número de pacientes, se prefirió comparar los valores del coeficiente de significación p del estadístico de Hosmer-Lemeshow, que ya no depende directamente del tamaño de la muestra.

El análisis de la capacidad de discriminación comparó los valores calculados para el área bajo la curva ROC de cada modelo (AUC). El mayor valor de AUC se corresponde con la mejor capacidad de discriminación. También, tomando un criterio de decisión del 50%, fueron comparados sensibilidad, es-

TABLA 1. Prueba de la bondad del ajuste de Hosmer-Lemeshow para las predicciones de mortalidad de 202 pacientes según los índices MPM II-0, MPM II-24, y APACHE II y SAPS II

Modelo predictivo	N.º pacientes	Supervivientes		No supervivientes	
		Observados	Esperados	Observados	Esperados
MPM II-0	20	19	18,055	1	1,945
	21	20	18,781	1	2,219
	21	18	18,521	3	2,479
	20	16	17,315	4	2,685
	20	18	16,830	2	3,170
	21	17	16,739	4	4,261
	20	12	14,677	8	5,323
	20	12	12,749	8	7,251
	20	12	9,896	8	10,104
	19	6	6,437	13	12,563
MPM II-24	20	20	18,398	0	1,602
	20	17	18,279	3	1,721
	20	20	18,067	0	1,933
	20	18	17,713	2	2,287
	21	15	18,044	6	2,956
	20	15	16,335	5	3,665
	20	14	15,013	6	4,987
	20	14	13,007	6	6,993
	20	12	9,720	8	10,280
	21	5	5,424	16	15,576
APACHE II-24	20	20	17,924	0	2,076
	20	19	17,730	1	2,270
	20	18	17,462	2	2,538
	20	14	17,188	6	2,812
	20	17	16,762	3	3,238
	21	16	16,755	5	4,245
	20	15	14,963	5	5,037
	21	13	14,134	8	6,866
	20	11	10,951	9	9,049
	20	7	6,131	13	13,869
SAPS II-24	21	21	19,549	0	1,451
	21	18	19,305	3	1,695
	22	20	19,797	2	2,203
	19	17	16,686	2	2,314
	22	20	18,823	2	3,177
	21	13	17,003	8	3,997
	20	17	14,805	3	5,195
	20	12	11,799	8	8,201
	21	10	8,103	11	12,897
	15	2	4,131	13	10,869

pecificidad, y cociente de clasificación correcta total (CCT) de los distintos índices predictivos.

Este estudio fue aprobado por la Comisión Local de Ética e Investigación y las predicciones obtenidas del empleo de los diferentes índices pronóstico nunca fueron utilizadas para modificar o influenciar las decisiones clínicas durante el período del estudio.

RESULTADOS

Fueron estudiados 202 pacientes críticos, 122 hombres y 80 mujeres, con una edad media de 50,9 (desviación estándar [DE] 19,4) años. En el segundo día de estancia, el estudio quedó limitado a 149 pacientes, pues entre las 24 y 48 horas se registraron 17 fallecimientos y 36 pacientes fueron evacuados a planta.

TABLA 2. Prueba de la bondad del ajuste de Hosmer-Lemeshow para las predicciones de mortalidad de 149 pacientes medidas a las 48 horas según los índices MPM II-48, APACHE II y SAPS II

Modelo predictivo	N.º pacientes	Supervivientes		No supervivientes	
		Observados	Esperados	Observados	Esperados
MPM II-48	15	14	13,629	1	1,371
	15	15	13,467	0	1,533
	15	14	13,320	1	1,680
	15	12	13,108	3	1,892
	15	12	12,801	3	2,199
	15	11	12,276	4	2,724
	15	11	11,676	4	3,324
	15	10	10,472	5	4,528
	15	10	8,326	5	6,674
	14	5	4,924	9	9,076
APACHE II-48	15	15	13,333	0	1,667
	15	14	13,166	1	1,834
	16	13	13,836	3	2,164
	15	13	12,791	2	2,209
	16	12	13,341	4	2,659
	15	13	12,038	2	2,962
	15	10	11,544	5	3,456
	15	7	10,779	8	4,221
	15	13	9,247	2	5,753
	12	4	3,925	8	8,075
SAPS II-48	15	15	13,943	0	1,057
	16	15	14,673	1	1,327
	17	16	15,340	1	1,660
	15	12	13,255	3	1,745
	16	12	13,756	4	2,244
	15	13	12,172	2	2,828
	16	12	12,201	4	3,799
	14	9	9,506	5	4,494
	16	9	7,242	7	8,758
	9	1	1,911	8	7,089

La mortalidad observada al alta hospitalaria fue del 25,7% (52 pacientes). El 19,8% (40 pacientes) fallecieron durante su estancia intra-UCI. Las categorías diagnósticas quedan reflejadas en la tabla 3.

Las tablas 1, 2, 4 y 5 muestran los resultados de las pruebas de bondad del ajuste de Hosmer-Lemeshow y los valores de calibración y discriminación para las predicciones realizadas sobre los pacientes; a) al ingreso y en el primer día, y b) en el segundo día, respectivamente.

La tabla 6 representa la comparación cualitativa de los resultados numéricos de las pruebas de calibración y discriminación recogidas en las tablas 1, 2, 4 y 5.

DISCUSIÓN

El primer estudio diseñado específicamente para comparar tres métodos de predicción de la mortalidad en pacientes críticos fue presentado por Lemeshow y colaboradores en 1987⁵. Se trataba entonces de comparar el *Acute Physiology Score* (APS) del APACHE I, el también inicial SAPS I, y la versión del primitivo MPM al ingreso que incluía la posibilidad de parada cardiorrespiratoria. Los tres sistemas se comportaron de modo similar con respecto a su

TABLA 3. Descripción de las categorías diagnósticas al ingreso

Cardiocirculatoria	63
Respiratoria	72
Gastrointestinal	11
Neurológica	23
Traumatológica	12
Metabólica/farmacológica	19
Fallo renal	2
Total pacientes	202

capacidad de discriminación. Las probabilidades expresadas por el MPM fueron las que más se aproximaron a la evolución observada, y esto fue atribuido a que se trata de un método cuyas variables son independientes del tratamiento dispensado en la UCI⁵.

En un estudio posterior en 1990, Schäfer et al²¹ compararon la capacidad predictiva de APACHE II con las versiones iniciales de SAPS I y MPM-0. Como en el estudio de Lemeshow, los tres sistemas se comportaron entre sí de modo similar con respecto a su sensibilidad, especificidad y cociente de clasificación correcta total, y los valores fueron bastante parecidos a los que hemos obtenido en nuestro estudio utilizando versiones posteriores de estos mismos modelos predictivos²¹.

Las categorías diagnósticas de nuestro estudio configuran una casuística comparable a las de otros estudios llevados a cabo en UCI polivalentes^{21,22}, en los que pacientes tales como quemados, trasplantados o postoperatorio de cirugía cardíaca no se han incluido^{23,24}.

El número de pacientes de nuestro trabajo, aunque reducido, es asimismo comparable con el tamaño de la muestra de otros estudios de validación no multicéntricos^{17,25,26}. La mortalidad hospitalaria observada en nuestra serie es también semejante a la encontrada en estudios multicéntricos recientes²⁷⁻²⁹.

Las predicciones de mortalidad de los modelos MPM II se realizaron de acuerdo con la metodología diseñada por sus autores, tal y como fue expuesta en los respectivos trabajos originales^{2,3}, y adoptado en cada índice el modelo original con objeto de hacer nuestros resultados comparables con los de otros investigadores. Con respecto a precedentes encontrados en la literatura en relación con la utilización de los índices predictivos APACHE II y SAPS II como escores dinámicos, Bion et al³⁰ en 1988 demostraron sobre 128 pacientes que el uso longitudinal de una variante del APACHE II, el sistema *sickness scoring*, fue el mejor predictor de la evolución³⁰. También ese año, Chang et al¹⁷ propusieron el uso diario del APACHE II corregido con el número de órganos o sistemas en fallo para el pronóstico individual de pacientes. Con posterioridad, en 1994 Douglas Wagner et al¹⁶, miembros del mismo equipo que desarrolló el APACHE II utilizaron secuencialmente este modelo, y sugirieron que su uso más allá de las primeras 24 horas podría proporcionar un refinamiento de sus predicciones de mortalidad.

TABLA 4. Comparación de predicciones realizadas sobre 202 pacientes al ingreso y a las 24 horas

	Calibración		Discriminación			
	\hat{C}	p	Sens	Espec	CCT	AUC
Al ingreso						
MPM II-0	5,54	0,69	34,62	92,67	77,72	0,762
A las 24 horas						
MPM II-24	10,77	0,21	42,31	93,33	80,2	0,791
APACHE II-24	8	0,40	30,77	94,00	77,7	0,763
SAPS II-24	11,7	0,16	48,08	91,33	80,2	0,798

\hat{C} : estadístico Chi cuadrado de Hosmer-Lemeshow; p: significación estadística de C; Sens: sensibilidad; Espec: especificidad; CCT: cociente de clasificación correcta total; AUC: área bajo la curva ROC.

Dentro de los objetivos de nuestro trabajo hemos considerado de interés investigar si este refinamiento del APACHE II se daría también en el modelo SAPS II frente a los modelos MPM II, que han sido diseñados específicamente para su utilización longitudinal en el tiempo como escores dinámicos.

Hemos encontrado que las predicciones sobre los 202 pacientes al ingreso con MPM-0 compararon favorablemente en cuanto a calibración con las realizadas por cualquiera de los modelos APACHE II, SAPS II o MPM II a las 24 horas (tabla 4). El MPM II-0 presentó superior calibración y similar discriminación que el APACHE II-24, que es el modelo que resultó con mejor calibración a las 24 horas, frente a MPM II-24 y SAPS II-24, que le superaron ligeramente en capacidad de discriminación. Estos dos últimos modelos tuvieron valores muy parecidos de calibración y discriminación, pero en conjunto no superaron a APACHE II-24 ni a MPM II-0.

En la comparación de predicciones realizadas a las 48 horas sobre los 149 pacientes restantes, las predicciones de MPM-48 compararon favorablemente con el resto de las realizadas a las 48 horas (tabla 5), presentando una calibración ligeramente superior a la del SAPS II-48, aunque este último tuvo la mejor discriminación. La calibración de APACHE II-48 resultó muy pobre frente a los otros dos modelos.

Nuestro estudio sugiere que, con excepción del APACHE II, todos los modelos refinan sus predicciones al avanzar en el seguimiento de la evolución del paciente. Es muy posible que el peso estadístico de la categoría diagnóstica, que tan buenos resulta-

TABLA 5. Comparación de predicciones realizadas sobre 149 pacientes a las 48 horas por los modelos MPM II-48 y APACHE II y SAPS II

	Calibración		Discriminación			
	\hat{C}	p	Sens	Espec	CCT	AUC
A las 48 horas						
MPM II-48	4,94	0,76	28,57	92,11	77,18	0,759
APACHE II-48	13,4	0,09	20,00	96,49	78,5	0,728
SAPS II-48	5,8	0,66	37,14	95,61	81,88	0,788

\hat{C} : estadístico Chi cuadrado de Hosmer-Lemeshow; p: significación estadística de C; Sens: sensibilidad; Espec: especificidad; CCT: cociente de clasificación correcta total; AUC: área bajo la curva ROC.

TABLA 6. Comparación cualitativa de la calibración (izquierda de la barra inclinada) y la discriminación (cruces a la derecha) de los modelos predictivos al ingreso, a las 24, y a las 48 horas de ingreso en UCI

MPM II-0	MPM II-24	MPM II-48
++++++/+++	+++/>+++++	++++++/++
	APACHE II-24	APACHE II-48
	++++/++++	+/+
	SAPS II-24	SAPS II-48
	++/>+++++	++++/+++++

dos otorga a APACHE II en las primeras 24 horas, le conduce a un mal ajuste de sus predicciones a las 48 horas frente a modelos, en los que no se incluye de forma explícita este parámetro. Con respecto a la comparación de la bondad del ajuste de las predicciones (tablas 1, 2, 4 y 5) entre las 24 y 48 horas con diferente número de pacientes en cada grupo, debemos recordar que los valores del estadístico χ^2 de Hosmer-Lemeshow dependen del número de sujetos de la muestra¹⁹. Así pues, hemos preferido realizar esta comparación basándonos en los valores del coeficiente p de significación estadística, ya que proporciona una idea muy intuitiva del acercamiento entre lo observado y lo predicho, y no depende, al menos directamente, del número de pacientes, evitando la comparación directa de los valores del estadístico χ^2 .

Para comparar la capacidad de discriminación de los modelos hemos elegido el área bajo la curva ROC como primer criterio, por encima de sensibilidad, especificidad y tanto por ciento de pacientes bien clasificados en el punto de corte del 50%. De estudios previos⁵, se conoce que usando un punto de corte del 50% virtualmente todos los modelos predictivos propuestos tendrán, al menos, un cociente de clasificación falsa del 10% al 15%. Esto ha sido interpretado por algunos autores como que todos los sistemas predictivos son aproximadamente equivalentes y ninguno es suficientemente bueno para la predicción individual²³. En comparación con nuestros resultados, Chang et al²⁵ en 1986, sobre una muestra de 210 pacientes, clasificaron correctamente con el APACHE II al 84,3% con una especificidad del 95,5%, considerando también un punto de corte del cincuenta por ciento²⁵. En otro estudio, Schäfer et al²¹ en 1990 clasificaron correctamente con el APACHE II el 68,3% de 593 pacientes polivalentes, con una especificidad del 74,9% en el mismo punto de corte. Finalmente, Castilla et al³¹ en 1991 clasificaron correctamente con el APACHE II el 83,73% de una muestra de 332 pacientes, con un 95,45% de especificidad³¹. En nuestra opinión, la comparación del área bajo la curva ROC cubriría todo el espectro de valores de discriminación tomando infinitos puntos de corte, y por tanto, nos ha parecido la prueba global más fiable a efectos de comparar la capacidad discriminativa de los modelos.

Nuestro análisis ha puesto de manifiesto la comparación favorable de la buena capacidad del sistema MPM-II para ser utilizado de forma secuencial, longitudinal en el tiempo, que superaría a APACHE II y SAPS II tomados por separado. La sencillez en la recogida de datos es mayor en MPM II y requiere un menor consumo de tiempo. Su principal inconveniente es la falta de implantación real del modelo más allá de su uso experimental dentro de la comunidad médica intensivista, donde el uso de los sistemas APACHE y/o SAPS se encuentra muy consolidado. Dado que APACHE II y SAPS II comparten un buen número de variables, una alternativa al uso secuencial del MPM II podría ser la determinación de probabilidades de mortalidad hospitalaria utilizando ambos modelos APACHE II y SAPS II conjuntamente. Aunque el número de pacientes de nuestro estudio es pequeño, y necesitaría ser validado con una muestra más amplia, nuestros resultados parecen sugerir que el uso del APACHE II sería recomendable en las primeras 24 horas, dentro de cuyo plazo tiene una excelente capacidad predictiva, con una calibración superior a la del SAPS II en ese período, y a continuación, el SAPS II, que presenta buen funcionamiento predictivo a las 48 horas, podría ser utilizado para este segundo período. La determinación secuencial de probabilidades de mortalidad hospitalaria usando APACHE II y SAPS II, según el esquema propuesto, posiblemente optimizaría los valores de calibración y discriminación en el uso combinado de estos modelos como escores dinámicos.

A modo de conclusión, en nuestro estudio el modelo MPM II presenta la mejor capacidad predictiva de la mortalidad hospitalaria cuando establece predicciones secuencialmente al ingreso, a las 24 y a las 48 horas en pacientes críticos polivalentes. Serán necesarios más estudios para confirmar esta aparente superioridad de los modelos MPM II sobre los modelos clásicos, pero mientras tanto, el uso combinado de APACHE II y SAPS II, de forma secuencial en los días primero y segundo de ingreso respectivamente, podría ser una alternativa válida de utilización de estos modelos predictivos en forma dinámica más allá de las primeras 24 horas de ingreso en UCI.

APÉNDICE I

Mortality Probability Models (MPM): Desarrollo histórico y estado actual

El sistema *Mortality Probability Model* (MPM) viene siendo desarrollado por Lemeshow, Teres et al²⁻¹⁰ y se trata de un modelo generado estadísticamente en su totalidad. Mediante la técnica de regresión logística múltiple sus autores seleccionaron las variables con mayor capacidad predictiva de la mortalidad hospitalaria³², determinaron los coeficientes de ponderación asociados a cada una de las variables seleccionadas, y también efectuaron el posterior refinamiento del peso aritmético de estos coeficientes⁹.

La primera versión de 1985 constaba de dos modelos: el modelo MPM-0, con siete variables medidas al ingreso, que fue obtenido a partir de datos de 737 pacientes⁷, y el modelo MPM-24, medido a las 24 horas, también con siete variables, y obtenido a partir de los datos de 458 pacientes⁷. De esta forma, los modelos proporcionaban las probabilidades de mortalidad hospitalaria estimadas al ingreso y a las 24 horas respectivamente. Los estudios de validación que siguieron a la publicación de esta primera versión del sistema MPM atribuyeron al modelo al ingreso, MPM-0, mayor capacidad predictiva que al MPM-24^{5,8}.

Una segunda versión de MPM fue presentada en 1987 coincidiendo con las validaciones anteriores⁸. Se propusieron dos modelos diferentes de MPM-0, también con siete variables. En el denominado MPM-0_[OSF] se incluyó como variable el número de órganos o sistemas en fallo [OSF= *organ system failure*]. Este modelo fue desestimado por sus propios autores por considerar que incluía una variable compleja (en realidad una variable por cada órgano o sistema) que no podía ser recogida de forma fiable, ya que podría depender de la percepción individual de cada observador. En el otro modelo MPM-0 se incluyó como variable la existencia de parada cardiorrespiratoria previa a la admisión, y se denominó MPM-0_[CPR]⁸.

La tercera versión de modelos MPM apareció en 1988⁴. Utilizando la base de datos almacenada de los estudios precedentes los autores diseñaron unos nuevos MPM: el MPM-0 medido al ingreso con 11 variables; el MPM-24 medido a las 24 horas con 14 variables; el MPM-48 medido al segundo día con 11 variables y, adicionalmente, el modelo llamado MPM-OT ("*over-time*") como una combinación de los tres anteriores, para ser usado en pacientes con más de 48 horas de estancia en la UCI⁴.

En 1993 apareció la cuarta versión del sistema MPM revisada y actualizada que es la que ha sido denominada MPM II³. El MPM II fue desarrollado, refinado y validado sobre 19.124 pacientes críticos de 137 hospitales de doce países de Europa y Norteamérica. Inicialmente constó de dos modelos, MPM II-0 y MPM II-24, para estimar la probabilidad de supervivencia al alta hospitalaria en el momento de ingreso en la UCI y a las 24 horas, respectivamente³. En 1994 se desarrollaron dos modelos adicionales MPM II-48 y MPM II-72, para las 48 y las 72 horas².

El MPM II-0 proporciona una estimación de la probabilidad de mortalidad hospitalaria antes de que comience el tratamiento en la UCI, y ello es útil para evaluar el rendimiento de la UCI y para estratificar los pacientes previamente a la randomización en ensayos clínicos. Consta de quince variables fácilmente obtenibles (tabla 7).

El MPM II-24 consta de trece variables, de las que cinco ya se encuentran incluidas en el MPM-0, y por tanto han sido recogidas en el momento del ingreso. Requiere la recogida de sólo ocho variables adicionales al cabo de 24 horas (tabla 8).

TABLA 7. Cálculo de probabilidades de mortalidad del MPM II-0

	β	x	βx
Término constante β_0	-5,46836		-5,46836
Diagnósticos fisiológicos			
Coma o estupor profundo	1,48592	0	0
Frecuencia cardíaca ≥ 150 /min	0,45603	0	0
Tensión arterial sistólica ≤ 90 mmHg	1,06127	1	1,06127
Diagnósticos crónicos			
Insuficiencia renal crónica	0,91906	0	0
Cirrosis	1,13681	1	1,13681
Carcinoma metastásico	1,19979	0	0
Diagnósticos agudos			
Fracaso renal agudo	1,48210	0	0
Arritmia cardíaca	0,28095	0	0
Accidente cerebrovascular	0,21338	0	0
Hemorragia gastrointestinal	0,39653	1	0,39653
Efecto masa intracraneal	0,86533	0	0
Otros			
Edad en años	0,03057	60	1,83420
Resucitación cardiopulmonar previa al ingreso	0,56995	0	0
Ventilación mecánica	0,79105	0	0
Causa médica o cirugía no electiva	1,19098	1	1,19098

Para el cálculo de la probabilidad individual de mortalidad en cualquier modelo MPM, cada una de las variables X_i queda expresada como valor absoluto (por ejemplo edad), o como términos de una condición dicotómica (valor 1, presencia; 0, ausencia). Este valor es corregido al multiplicarlo por un coeficiente de ponderación β_i , obtenido mediante regresión logística múltiple a partir de la cohorte de incepción del estudio original^{2,3}. El valor de la suma de este polinomio o *logit* se sustituye en la fórmula general del cálculo de probabilidad de un modelo logístico $Pr = e^{\text{logit}} / (1 + e^{\text{logit}})$, donde e es la constante matemática 2,7182818 que representa la base de los logaritmos neperianos y Pr es la probabilidad individual de mortalidad hospitalaria de ese paciente.

Ejemplo 1. En la tabla 7 para el cálculo de las probabilidades de mortalidad hospitalaria según MPM-0 supondremos el cálculo del $\text{logit} = \sum \beta_i X_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \dots + \beta_{15} X_{15}$ para un pa-

ciente de 60 años con antecedentes de cirrosis alcohólica, ingresado por causa médica en la UCI con el diagnóstico de hemorragia gastrointestinal, y que presenta hipotensión arterial con sistólica menor de 90 mmHg, y con frecuencia cardíaca al ingreso de 130 latidos/minuto sería según la suma de términos de la columna derecha:

$$\begin{aligned} \text{logit} &= -5,46836 + (1 \times 1,06127) + (1 \times 1,13681) \\ &+ (1 \times 0,39653) + (60 \times 0,0357) + (1 \times 1,19098) \\ &= 0,15143 \end{aligned}$$

y la probabilidad de mortalidad hospitalaria Pr , expresada en tanto por uno, sería:

$$Pr = e^{\text{logit}} / (1 + e^{\text{logit}}) = e^{0,15143} / (1 + e^{0,15143}) = 0,54$$

Esta probabilidad es una estimación o expectativa basada en la mortalidad observada en un grupo numeroso de pacientes con características similares, y representa la proporción de pacientes que se espera

TABLA 8. Cálculo de probabilidades de mortalidad del MPM II a las 24 y 48 horas

	β	x	βx
Término constante β_0			
A las 24 horas	-5,64592		-5,64592
A las 48 horas	-5,39153		-5,39153
Variables obtenidas al ingreso			
Edad, años	0,03268	60	1,96080
Cirrosis	1,08745	0	0
Efecto masa intracraneal	0,91314	0	0
Neoplasia metastásica	1,16109	0	0
Causa médica o cirugía no electiva	0,83404	1	0,83404
Variables evaluadas a intervalos de 24 horas			
Coma o estupor profundo a las 24 horas	1,68790	0	0
Creatinina > 2,0 mg/dl	0,72283	1	0,72283
Infección confirmada	0,49742	1	0,49742
Ventilación mecánica	0,80845	1	0,80845
$P_aO_2 < 60$ mmHg	0,46677	0	0
Tiempo de protrombina > 3 seg por encima del control	0,55352	0	0
Diuresis < 150 ml en 8 horas	0,82286	1	0,82286
Fármacos vasoactivos durante ≥ 1 hora	0,71628	0	0

que muestren la misma evolución. Así, se esperaba que falleciera el 54% de los pacientes que tuvieran esta probabilidad, mientras que se esperaría la supervivencia del 46% de los pacientes.

Los modelos MPM II-48 y MPM II-72² contienen las mismas 13 variables y coeficientes que el MPM II-24. La diferencia entre ellos estriba sólo en los incrementos del término constante β_0 , que, como veremos, aumenta su valor al pasar de MPM II-24 a MPM II-48, y de éste a MPM II-72, de forma que repercute en el aumento del polinomio $\sum \beta_i X_i$. Aun cuando no existiesen cambios diarios en ninguna variable, el valor de la probabilidad de mortalidad hospitalaria aumentaría conforme aumentase la estancia en la UCI, reflejando que si el paciente no mejora en el tiempo es que en realidad está tendiendo a empeorar².

Ejemplo 2. Supongamos ahora el cálculo de las probabilidades de mortalidad hospitalaria a las 24 y a las 48 horas según los modelos MPM II-24 y MPM II-48 (tabla 8) para un paciente de 60 años con perforación de un absceso diverticular (infección confirmada) que presenta un fracaso renal agudo (reducción de la diuresis a < 150 ml en ocho horas y aumento de la creatinina > 2,0 mg/dl) e insuficiencia respiratoria con necesidad de ventilación mecánica. En este ejemplo, asumiremos que no hay cambios en el perfil del paciente entre las 24 y las 48 horas de estancia en UCI. El cálculo del *logit* = $\sum \beta_i X_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \dots + \beta_{13} X_{13}$ se realizaría de acuerdo con los valores de la columna de la derecha de la tabla 8. En cada caso, el *logit* se calcula como la suma del término constante (β_0) más los productos de cada una de las trece variables del modelo por el valor de sus respectivos coeficientes β .

Los valores del *logit* y la probabilidad de mortalidad hospitalaria asociada para este paciente a las 24 y a las 48 horas serán:

$$\begin{aligned} \text{logit a las 24 horas: } & 0,00048 \\ \text{logit a las 48 horas: } & 0,25487 \end{aligned}$$

Este *logit* se transforma en una probabilidad de mortalidad hospitalaria: $\text{Pr} = e^{\text{logit}} / (1 + e^{\text{logit}})$, que a las 24 horas tendrá un valor:

$$\text{Pr} = e^{0,00048} / (1 + e^{0,00048}) = 0,50$$

De forma similar, la probabilidad de mortalidad hospitalaria a las 48 horas será:

$$\text{Pr} = e^{0,25487} / (1 + e^{0,25487}) = 0,56$$

La ausencia de cambios en el paciente entre las 24 y las 48 horas no sería un buen signo. En términos cuantitativos, esto se refleja en que si el perfil del paciente no cambia, el pronóstico no sólo no continúa siendo el mismo sino que realmente empeora.

Los sistemas APACHE y SAPS

1. El sistema APACHE

El conjunto de modelos *Acute Physiology and Chronic Health Evaluation* (APACHE) viene siendo diseñado y actualizado por Knaus et al desde finales de los años 70¹¹⁻¹³.

APACHE I. Introducido en 1981, constaba de 34 variables seleccionadas por un panel de siete expertos representando tres especialidades troncales: anestesia, medicina interna y cirugía. Se escogieron variables disponibles al ingreso en UCI, o que pudieran ser obtenidas durante las primeras 32 horas de estancia¹¹. El número de variables se juzgó excesivo en el primer estudio multicéntrico en el que se utilizó³³, y aunque cayó rápidamente en desuso, su impacto sobre la evolución y el desarrollo de nuevas medidas de la gravedad sigue todavía vigente.

APACHE II. En esta versión de 1985, el número de determinaciones se redujo hasta doce variables fisiológicas más la edad y el estado de salud previo¹². Se divide en dos componentes. El primero, llamado APS o *Acute Physiology Score* puntúa las variables fisiológicas. Para la determinación se escogen los peores valores de las primeras 24 horas de ingreso, y se les asignan valores de 0 a 4 puntos. Como excepción el *Glasgow Coma Scale* (GCS) se puntúa restando de 15 su valor para el paciente en estudio (tabla 9). La suma de las puntuaciones de estas variables proporcionará este primer componente APS del APACHE II, que se considera una medida de la gravedad de la enfermedad aguda del paciente.

El segundo componente, denominado *Chronic Health Evaluation*, puntúa la edad y el estado de salud previo (presencia de enfermedad crónica definida de los sistemas cardiovascular, respiratorio, hepático, renal e inmunológico).

La suma de ambas escalas constituye la puntuación *Acute Physiology And Chronic Health Evaluation II* o APACHE II. La puntuación máxima posible del sistema APACHE II es 71, aunque apenas existe supervivencia sobrepasando los 55 puntos. La puntuación APACHE II tiene, como expresión de gravedad, significado en sí misma. Además, es posible establecer la predicción individual de la mortalidad hospitalaria (la variable dependiente), tomando como variables independientes la propia puntuación APACHE II, si el paciente había o no recibido cirugía de urgencia, y el coeficiente de categoría diagnóstica, es decir, el coeficiente de ponderación asignado a la categoría diagnóstica del paciente por la regresión logística múltiple realizada inicialmente por Knaus et al¹². La probabilidad de muerte hospitalaria, Pr, viene dada por la misma fórmula general vista anteriormente para el cálculo de probabilidad en un modelo de regresión logística múltiple:

$$\text{Pr} = e^{\text{logit}} / (1 + e^{\text{logit}})$$

TABLA 9. Sistema de puntuación APACHE II (*Acute Physiology And Chronic Health Evaluation II*)

4	3	2	1	0	Puntuación	1	2	3	4
≥ 41,0	39,0-40,9		38,5-38,9	36,0-38,4	Temperatura central (°C)	34,0-35,9	32,0-33,9	30,0-31,9	≤ 29,9
≥ 160	130-159	110-129		70-109	Presión arterial media (mmHg)		50-69		≤ 49
≥ 180	140-170	110-139		70-109	Ritmo cardíaco (latidos/minuto)		55-69	40-54	≤ 39
≥ 50	35-49		25-34	12-24	Frecuencia respiratoria (con o sin VM)	10-11	6-9		≤ 5
≥ 500	350-499	200-349		< 200 > 70	Oxigenación* (mmHg) A-aDO ₂ PaO ₂	61-70		55-60	< 55
≥ 7,70	7,60-7,69		7,50-7,59	7,33-7,49	pH arterial		7,25-7,32	7,15-7,24	< 7,15
≥ 180	160-179	155-159	150-154	130-149	Sodio (mMol/l)		120-129	111-119	≤ 110
≥ 7,0	6,0-6,9		5,5-5,9	3,5-5,4	Potasio (mMol/l)	3,0-3,4	2,5-2,9		< 2,5
≥ 3,5	2-3,4	1,5-1,9		0,6-1,4	Creatinina** (mg/100 ml)		< 0,6		
≥ 60		50-59,9	46-49,9	30-45,9	Hematócrito (%)		20-29,9		< 20
≥ 40		20-39,9	15-19,9	3-14,9	Leucocitos (×10 ⁹ /l)		1-2,9		< 1

*Si la fracción inspirado de oxígeno (FiO₂) es ≥ 0,5, se asignan puntos al gradiente alveolo-arterial (A-aDO₂). Si la fracción inspirada de oxígeno es < 0,5, se asignarán puntos a la presión parcial de oxígeno arterial (PaO₂).

**La creatinina tendrá doble puntuación en presencia de fracaso renal agudo.

VM: ventilación mecánica.

Otros puntos

Glasgow Coma Scale: El GCS del paciente se restará de 15, y el valor de la diferencia 15-GCS se consignará como puntos.

Edad:	Años	Puntos
	≤ 44	= 0
	45-54	= 2
	55-64	= 3
	65-74	= 5
	≥ 75	= 6

Estado de salud crónico: Si el paciente, antes del ingreso en el hospital, tiene historia de insuficiencia severa de órganos o de compromiso inmunitario encuadrado en las definiciones siguientes, se asignarán como sigue:

para pacientes no quirúrgicos o postoperatorio de cirugía urgente: 5 puntos

para postoperatorio de cirugía electiva: 2 puntos

La insuficiencia de órganos o el compromiso inmunitario deben haber sido evidentes previamente al ingreso hospitalario actual, y conforme a los siguientes criterios:

Hígado: cirrosis diagnosticada por biopsia e hipertensión portal documentada; o episodios anteriores de sangrado gastrointestinal atribuidos a hipertensión portal; o episodios previos de fallo hepático, encefalopatía o coma.

Cardiovascular: Clase IV de la *New York Heart Association*.

Respiratorio: enfermedad respiratoria crónica restrictiva, obstructiva, o vascular, con limitación severa al ejercicio, por ejemplo incapacidad para subir escaleras o realizar las tareas domésticas; o situación crónica documentada de hipoxemia, hipercapnia, policitemia secundaria, hipertensión pulmonar severa (> 40 mmHg), o dependencia de respirador.

Renal: enfermedad renal dependiente de diálisis crónica.

Compromiso inmunitario: el paciente ha recibido tratamiento que suprime la resistencia a la infección, por ejemplo, inmunosupresión, radiación, quimioterapia, altas dosis de esteroides recientes o en tratamiento a largo plazo; o el paciente tiene una enfermedad que está suficientemente avanzada como para suprimir la resistencia a la infección, por ejemplo, leucemia, linfoma, o SIDA.

El valor del *logit* para el modelo APACHE II se calcula mediante la siguiente ecuación:

$$\text{logit} = -3,517 + \text{puntuación APACHE II} \times 0,146 + 0,603 \text{ (sólo si cirugía de urgencia)} + \text{coeficiente categoría diagnóstica.}$$

El valor de esta probabilidad individual de mortalidad se considera más descriptivo que la propia puntuación APACHE II, ya que ésta podría resultar idéntica para dos pacientes distintos con diferente categoría diagnóstica, y por tanto con diferente pronóstico³⁴.

APACHE III. La investigación de métodos para la optimización del sistema APACHE ha permanecido abierta, y su tercera versión apareció en 1991 con la novedad de un formato en paquete de software¹³. El sistema consta de dos partes: una puntuación, que permite la estratificación de la gravedad de los pacientes críticos dentro de grupos definidos de pacientes, y una ecuación predictiva, que proporciona el riesgo estimado de mortalidad hospitalaria en pacientes individuales. El APACHE III incluye variables muy parecidas a las de su versión anterior, pero el cálculo de la predicción de mortalidad y el manejo del producto se encuentran bajo licencia de utilización de APACHE *Medical Systems* (AMS), Inc, y

por tanto es preciso compararlo para poder acceder a los cálculos. La información disponible en la literatura sólo permite el cálculo de la puntuación aguda, denominada *APACHE III Physiologic Scoring* o *APS III*, que incluye alteraciones neurológicas, ácido-base, de los signos vitales y pruebas de laboratorio.

El *APACHE III* no ha conseguido hasta ahora desplazar el uso de su antecesor, el *APACHE II*. Por una parte, al quedar la utilización libre del *APACHE III* restringida al uso de la puntuación *APS III*, la metodología publicada en la literatura sólo ha permitido puntuar el grado de alteración aguda en el paciente crítico, y por tanto ha limitado su expansión. Por otra parte, la fiabilidad de la ecuación predictiva del *APACHE III* ha sufrido un duro revés tras la publicación en 1998 de un estudio de validación llevado a cabo en 285 UCI de 161 hospitales norteamericanos sobre 37.668 pacientes, en donde el modelo ha mostrado una notable falta de calibración³⁵.

2. El sistema SAPS

SAPS I. La primera versión del *Simplified Acute Physiology Score* (*SAPS*) fue propuesta en 1984 por Le Gall et al¹⁴ como simplificación del *APACHE I*. El *SAPS I* constaba de trece variables clínicas y la edad. A su gran difusión posiblemente contribuyó su mayor facilidad de aplicación, aunque cayó en desuso tras la aparición del *APACHE II*.

SAPS II. En 1993 el mismo equipo de investigadores publicó una nueva versión, el *SAPS II*¹⁵, desarrollado, refinado y validado usando datos de 12.997 pacientes de 137 UCI de Norteamérica y Europa. Consta de doce variables fisiológicas, tres variables de estado de salud previo, la edad, y una variable indicando la causa de admisión en la UCI (tabla 10).

Una vez obtenida la puntuación *SAPS II*, y al igual que en el *APACHE II*, el cálculo de la probabilidad individual de mortalidad hospitalaria se establece a partir del *logit* de cada paciente, calculado mediante la ecuación:

$$\text{logit} = -7,7631 + 0,0737 (\text{puntuación SAPS II}) + 0,9971 [\ln(\text{puntuación SAPS II} + 1)].$$

El *logit* es convertido a probabilidad de mortalidad hospitalaria mediante la misma ecuación general $\text{Pr} = e^{\text{logit}} / (1 + e^{\text{logit}})$ que hemos visto en los modelos anteriores.

BIBLIOGRAFÍA

1. Lemeshow S, Le Gall JR. Modeling the severity of illness of ICU patients. A systems update. *JAMA* 1994; 272: 1.049-1.055.
2. Lemeshow S, Klar J, Teres D, Avrunin JS, Gehlbach SH, Rapoport J, et al. Mortality probability models for patients in the intensive care unit for 48 or 72 hours: a prospective, multicenter study. *Crit Care Med* 1994; 22: 1.351-1.358.

3. Lemeshow S, Teres D, Klar J, Avrunin JS, Gehlbach SH, Rapoport J. Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients. *JAMA* 1993; 270: 2.478-2.486.

4. Lemeshow S, Teres D, Avrunin JS, Gage RW. Refining intensive care unit outcome prediction by using changing probabilities of mortality. *Crit Care Med* 1988; 16: 470-477.

5. Lemeshow S, Teres D, Avrunin JS, Pastides H. A comparison of methods to predict mortality of intensive care unit patients. *Crit Care Med* 1987; 15: 715-722.

6. Lemeshow S, Pastides H, Avrunin JS, Teres D. Applications of microcomputer spreadsheet packages as adjuncts to multiple logistic regression analysis. *Am J Epidemiol* 1985; 122: 710-716.

7. Lemeshow S, Teres D, Pastides H, Avrunin JS, Steingrub JS. A method for predicting survival and mortality of ICU patients using objectively derived weights. *Crit Care Med* 1985; 13: 519-525.

8. Teres D, Lemeshow S, Avrunin JS, Pastides H. Validation of the mortality prediction model for ICU patients. *Crit Care Med* 1987; 15: 208-213.

9. Lemeshow S, Hosmer DWJ. Estimating odds ratios with categorically scaled covariates in multiple logistic regression analysis. *Am J Epidemiol* 1984; 119: 147-151.

10. Teres D, Brown RB, Lemeshow S, Parsells JL. A comparison of mortality and charges in two differently staffed intensive care units. *Inquiry* 1983; 20: 282-289.

11. Knaus WA, Zimmerman JE, Wagner DP, Draper EA, Lawrence DE. *APACHE-acute physiology and chronic health evaluation: A physiologically based classification system*. *Crit Care Med* 1981; 9: 591-597.

12. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. *APACHE II: a severity of disease classification system*. *Crit Care Med* 1985; 13: 818-829.

13. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, et al. The *APACHE III* prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 1991; 100: 1.619-1.636.

14. Le Gall JR, Loirat P, Alperovitch A, Glaser P, Granthil C, Mathieu D, et al. A simplified acute physiology score for ICU patients. *Crit Care Med* 1984; 12: 975-977.

15. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (*SAPS II*) based on a European/North American multicenter study. *JAMA* 1993; 270: 2.957-2.963.

16. Wagner DP, Knaus WA, Harrell FE, Zimmerman JE, Watts C. Daily prognostic estimates for critically ill adults in intensive care units: results from a prospective, multicenter, inception cohort analysis. *Crit Care Med* 1994; 22: 1.359-1.372.

17. Chang RW, Jacobs S, Lee B. Predicting outcome among intensive care unit patients using computerized trend analysis of daily *Apache II* scores corrected for organ system failure. *Intensive Care Med* 1988; 14: 558-566.

18. Lemeshow S, Hosmer DWJ. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982; 115: 92-106.

19. Hosmer DW, Taber S, Lemeshow S. The importance of assessing the fit of logistic regression models: a case study. *Am J Public Health* 1991; 81: 1.630-1.635.

20. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med* 1997; 16: 965-980.

21. Schafer JH, Maurer A, Jochimsen F, Emde C, Wegscheider K, Arntz HR, et al. Outcome prediction models on admission in a medical intensive care unit: do they predict individual outcome? *Crit Care Med* 1990; 18: 1.111-1.118.

22. McClish DK, Powell SH. How well can physicians estimate mortality in a medical intensive care unit? *Med Decis Making* 1989; 9: 125-132.

23. Schuster DP. Predicting outcome after ICU admission. The art and science of assessing risk. *Chest* 1992; 102: 1.861-1.870.

24. Castella X, Artigas A, Bion J, Kari A. A comparison of severity of illness scoring systems for intensive care unit patients: results of a multicenter, multinational study. The European/North American Severity Study Group. *Crit Care Med* 1995; 23: 1.327-1.335.

25. Chang RW, Jacobs S, Lee B. Use of APACHE II severity of disease classification to identify intensive-care-unit patients who would not benefit from total parenteral nutrition. *Lancet* 1986; 1: 1.483-1.487.
26. Chang RW, Jacobs S, Lee B, Pace N. Predicting deaths among intensive care unit patients. *Crit Care Med* 1988; 16: 34-42.
27. Knaus WA, Draper EA, Wagner DP. Utilizing findings from the APACHE III research to develop operational information system for the ICU: the APACHE III ICU Management System. *Proc Annu Symp Comput Appl Med Care* 1991; 987-989.
28. Rowan KM, Kerr JH, Major E, McPherson K, Short A, Vessey MP. Intensive Care Society's APACHE II study in Britain and Ireland (I): Variations in case mix of adult admissions to general intensive care units and impact on outcome. *BMJ* 1993; 307: 972-977.
29. Rowan KM, Kerr JH, Major E, McPherson K, Short A, Vessey MP. Intensive Care Society's APACHE II study in Britain and Ireland (II): Outcome comparisons of intensive care units after adjustment for case mix by the American APACHE II method. *BMJ* 1993; 307: 977-981.
30. Bion JF, Aitchison TC, Edlin SA, Ledingham IM. Sickness scoring and response to treatment as predictors of outcome from critical illness. *Intensive Care Med* 1988; 14: 167-172.
31. Castella X, Gilabert J, Torner F, Torres C. Mortality prediction models in intensive care: acute physiology and chronic health evaluation II and mortality prediction model compared. *Crit Care Med* 1991; 19: 191-197.
32. Teres D, Brown RB, Lemeshow S. Predicting mortality of intensive care unit patients. The importance of coma. *Crit Care Med* 1982; 10: 86-95.
33. Knaus WA, Draper EA, Wagner DP, Zimmerman JE, Birnbaum ML, Cullen DJ, et al. Evaluating outcome from intensive care: a preliminary multihospital comparison. *Crit Care Med* 1982; 10: 491-496.
34. Civetta JM, Hudson-Civetta JA, Kirton O, Aragon C, Salas C. Further appraisal of APACHE II limitations and potential. *Surg Gynecol Obstet* 1992; 175: 195-203.
35. Zimmerman JE, Wagner DP, Draper EA, Wright L, Alzola C, Knaus WA. Evaluation of acute physiology and chronic health evaluation III predictions of hospital mortality in an independent database. *Crit Care Med* 1998; 26: 1.317-1.326.