

Desempeño de los puntajes APACHE II y SAPS II para calcular la razón de mortalidad estandarizada en una Unidad de Cuidados Intensivos polivalente de la Argentina

NICOLÁS S. ROCCHETTI,* DANIEL H. BAGILET,** CLAUDIO J. SETTECASE,#
MARTA QUAGLINO†

* Carrera de Postgrado de Especialización en Clínica Médica

** Jefe, Unidad de Cuidados Intensivos, Hospital Escuela "Eva Perón",
Granadero Baigorria, Santa Fe

Subjefe, Unidad de Cuidados Intensivos, Hospital Escuela "Eva Perón"

† Profesora Titular de la Licenciatura en Estadística, Facultad de Ciencias
Económicas y Estadística, Universidad Nacional de Rosario, Santa Fe

Correspondencia:

Dr. Nicolás Rocchetti
nicolasrocchetti@hotmail.com

Los autores declaran que no han recibido financiamiento parcial o total para este trabajo.

Palabras clave

- APACHE
- Mortalidad
- Terapia Intensiva
- Indicadores de calidad

Resumen

Objetivo: Estudiar el desempeño del *Acute Physiology and Chronic Health Evaluation II* (APACHE II) y el *Simplified Acute Physiologic Score* (SAPS II) para calcular la razón de mortalidad estandarizada en nuestra Unidad de Cuidados Intensivos.

Diseño: Estudio retrospectivo y de observación entre enero de 2007 y diciembre de 2013.

Ámbito: Unidad de Cuidados Intensivos polivalente.

Pacientes: De ambos sexos, >18 años, internados, como mínimo, 24 horas.

Intervenciones: Ninguna.

Variables de interés: APACHE II, SAPS II, mortalidad, razón de mortalidad estandarizada.

Resultados: Se analizaron 2641 pacientes que reunieron los datos requeridos. Edad promedio: 50.14 (\pm 16.95) años, varones: 67%, promedio de internación: 6.79 (\pm 9.14) días, mortalidad: 23,82%. Mortalidad promedio estimada por APACHE II 21,8% (\pm 21,4) y por SAPS II 27,4% (\pm 29,1) ($p = 0,000$). La razón de mortalidad estandarizada para la población estudiada con APACHE II fue de 1,09 (IC95%: 1,01-1,18) y con SAPS II fue de 0,87 (IC95%: 0,80-0,93). El puntaje APACHE II tuvo una adecuada calibración, pero no así el SAPS II (test de Hosmer-Lemeshow: $p = 0,4042$ y $p = 0,0024$, respectivamente). La discriminación de los puntajes APACHE II y SAPS II fue buena (ABC = $0,830 \pm 0,010$; $p < 0,001$; IC95% 0,811-0,849 y ABC = $0,836 \pm 0,001$; $p < 0,001$; IC95% 0,818-0,854, respectivamente).

Conclusiones: El puntaje APACHE II tuvo una discriminación y una calibración buenas, y el SAPS II tuvo buena discriminación y mala calibración. Por lo tanto, en nuestra Unidad, sería razonable emplear el puntaje APACHE II para calcular la razón de mortalidad estandarizada.

Key words

- APACHE
- Mortality
- Intensive Care
- Quality indicators

Abstract

Objective: To study the performance of APACHE II and SAPS II for the estimation of the standardized mortality ratio (SMR) in our Unit.

Design: Retrospective observational study performed from January 2007 to December 2013.

Setting: An Intensive Care Unit in a tertiary hospital.

Patients: Both sexes, >18 years old, hospitalized for at least 24 hours.

Interventions: None.

Variables of interest: APACHE II, SAPS II, mortality, standardized mortality ratio.

Results: Data from 2,641 patients who gathered the required information, were analyzed. The average age was 50.14 (\pm 16.95) years, 67% were male, the average time of hospitalization was 6.79 (\pm 9.14) days and mortality was 23.82%. The average mortality was 21.8% estimated with APACHE II score, and 27.4% with SAPS II (p 0.000). The standardized mortality ratio for the studied population using APACHE II was 1.09 (95%CI: 1.01-1.18), and 0.87 (95%CI: 0.80-0.93) with SAPS II. APACHE II had adequate calibration but this not occurred with SAPS II (Hosmer-Lemeshow test: $p=0.4042$ and $p=0.0024$, respectively). Discrimination of APACHE II and SAPS II scores was good (AUC: 0.830 ± 0.010 , $p < 0.001$, 95%CI 0.811-0.849, and AUC: 0.836 ± 0.001 , $p < 0.001$, 95%CI 0.818-0.854, respectively).

Conclusions: We conclude that APACHE II score had good discrimination and calibration, while SAPS II score had good discrimination and bad calibration. Therefore, in our Intensive Care Unit it would be reasonable to use the APACHE II score to calculate the standardized mortality ratio.

Introducción

Los puntajes pronósticos reflejan las características poblacionales de los países en donde se originan. El desarrollo del puntaje APACHE II (*Acute Physiology and Chronic Health Evaluation II*) se basó en una cohorte de pacientes de los Estados Unidos, y es utilizado en muchas Unidades de Cuidados Intensivos (UCI) alrededor del mundo.^{1,2} El puntaje SAPS II (*Simplified Acute Physiology Score II*) fue desarrollado en Europa y, también, se emplea en todo el mundo.^{3,4} Ninguno de estos dos sistemas incluyó poblaciones de América del Sur en su confección. Sin embargo, son los más difundidos en nuestro país.

Los modelos pronósticos deben personalizarse para adaptarse al *case-mix* de la población donde son utilizados, ya que los resultados en las bases de datos originales y el período en el que los modelos se origi-

naron, son diferentes de los de las bases de datos de las UCI donde se emplean.^{5,6}

Aunque de los puntajes mencionados se reportó una buena adaptación a diferentes períodos y *case-mix*, pocos estudios han evaluado formalmente la precisión de los modelos predictivos cuando se aplican a nuevas poblaciones de otras instituciones o países.⁷⁻⁹

En la Argentina, en 2002, el Comité de EscORES de la Sociedad Argentina de Terapia Intensiva (SATI) realizó la validación externa de los puntajes APACHE II y SAPS II en 36 Centros de 11 provincias, y halló que ninguno de los dos puntajes estudiados validaban en la población de nuestro país.¹⁰

La comparación entre la mortalidad observada y la esperada empleando un puntaje predictivo, es decir, la razón de mortalidad estandarizada (RME), es un indicador de calidad de resultados de la atención médica dentro de la UCI; y sirve para la mejora continua de los servicios de salud prestados. Si bien los resultados

- Desempeño de los puntajes APACHE II y SAPS II para calcular la razón de mortalidad estandarizada en una Unidad de Cuidados Intensivos polivalente de la Argentina

de la RME deben ser interpretados con precaución y no es conveniente utilizarlos para auditoría comparativa entre Servicios, es una herramienta sumamente útil para evaluar el desempeño en el tiempo de una UCI en particular.^{11,12}

La probabilidad de muerte utilizada para calcular la RME puede ser estimada con diferentes puntajes, entre ellos el APACHE II y el SAPS II. Estos últimos son los más difundidos en nuestro país, aunque es importante señalar que ambos subestiman el riesgo de muerte según lo reportado por la SATI.¹⁰ En nuestro caso, observamos que la probabilidad de muerte predicha por el puntaje SAPS II es mayor que la calculada por el APACHE II en más de la mitad de los casos (55,43%). Este dato, sumado a la escasa información sobre el tema en nuestro medio, nos motivó a estudiar el desempeño de dichos puntajes en el cálculo de la RME.

Materiales y Métodos

Diseño, criterios de inclusión y de exclusión

Este estudio retrospectivo y de observación se llevó a cabo en la UCI polivalente de complejidad 1, con capacidad docente, del Hospital Escuela “Eva Perón”. Se trata de un efector público general dependiente del Ministerio de Salud y Medio Ambiente de la Provincia, ubicado en la ciudad de Granadero Baigorria (Gran Rosario), provincia de Santa Fe, Argentina. Funciona como Hospital de referencia de nueve centros asistenciales dentro del Área Programática I del Nudo Rosario de Salud. Tiene 137 camas disponibles para la atención de pacientes adultos con patología aguda, 14 de las cuales pertenecen a la UCI. En la Institución, no se realizan procedimientos quirúrgicos cardiovasculares centrales, está considerado Hospital Donante, pero no cuenta con Unidad de Trasplante de Órganos.

En este estudio, se analizaron las siguientes variables: edad, sexo, patología de ingreso: médica, quirúrgica, trauma, puntajes APACHE II y SAPS II, y mortalidad en la UCI de los pacientes ingresados, al menos, 24 horas, de 18 años de edad o más, entre el 1 de enero de 2007 y el 31 de diciembre de 2013.

Se excluyeron del análisis los registros de los pacientes a los cuales les faltaban los datos necesarios para el cálculo de los puntajes, las variables demográficas, la patología de ingreso o el resultado.

No se excluyó del análisis de mortalidad a los pacientes con limitación del esfuerzo terapéutico, aunque su número fue insignificante.

Instrumento de registro de datos

Se utilizó el programa SATI-Q como instrumento de registro de datos y para el cálculo automático de los puntajes. El programa SATI-Q es una herramienta

informática usada para el registro de datos referidos a estándares de calidad, auspiciado por la SATI para aquellas UCI participantes del programa *Quality Benchmarking*. La carga de datos se realizó en tiempo real y estuvo a cargo de personal médico, fisiátrico y de enfermería debidamente capacitado. Los datos de los puntajes APACHE II y SAPS II fueron cargados exclusivamente por médicos.¹³

No se realizó ningún tipo de intervención relacionada con este estudio.

Análisis estadísticos

Las variables continuas fueron resumidas como media (\pm desviación estándar) y las cualitativas, como número de casos y porcentajes. Para evaluar las hipótesis comparativas se utilizaron pruebas t bilaterales o de ji al cuadrado, respectivamente. En todas las pruebas, el nivel de significación empleado fue alfa = 0,05. Para evaluar la capacidad predictiva de los puntajes de gravedad (discriminación) se analizaron los perfiles de las curvas ROC (*Receiver Operating Characteristic*) y las AUC (*Area Under Curve*), con sus intervalos de confianza del 95% (IC95%). Para evaluar la calibración se empleó el test de bondad de ajuste de Hosmer-Lemeshow.

Los análisis se realizaron con el programa SPSS PASW Statistics.

Aspectos éticos

Este estudio fue aprobado por el Comité de Docencia e Investigación del Hospital Escuela “Eva Perón”. Para proteger la confidencialidad de los pacientes se reemplazó el nombre y apellido por un código alfanumérico. Esta información sólo fue manejada por los autores y, por ningún motivo, estuvo en conocimiento de personas ajenas al estudio.

Resultados

Durante el período de observación, 3396 pacientes ingresaron en el Servicio y estuvieron internados, como mínimo, 24 horas. De ellos, 2641 (77,8%) cumplieron con los criterios de inclusión y constituyeron la población del estudio.

La edad promedio fue de 50.14 (\pm 16.95) años, el 67% eran varones, el promedio de internación fue de 6.79 (\pm 9.14) días y el 23,82% falleció dentro de la UCI.

Las características de los pacientes sobrevivientes y fallecidos se detallan en la Tabla 1.

El puntaje APACHE II predijo una mortalidad promedio para la población estudiada del 21,8% (576 muertes), mientras que la mortalidad predicha por el puntaje SAPS II fue del 27,4% (724 muertes) ($p = 0,000$).

El valor de la RME utilizando el puntaje APACHE II como predictivo de mortalidad esperada fue de 1,09

(IC95% 1,01-1,18), mientras que, con el puntaje SAPS II, fue de 0,87 (IC95% 0,8-0,93) (p = 0,02).

En cuanto a la calibración (Tabla 2), se evidencia que el puntaje APACHE calibra adecuadamente, pero no así el SAPS II (test de Hosmer-Lemeshow: p = 0,4042 y p = 0,0024, respectivamente). La sobrevaloración de la mortalidad esperada por el puntaje SAPS II se da especialmente en pacientes con riesgos de moderados a graves.

Las curvas ROC para predecir la mortalidad con los puntajes APACHE II y SAPS II se muestran en la Figura y las AUC con sus respectivos intervalos de confianza, en la Tabla 3. Ambos modelos mostraron buena capacidad de discriminación. El que empleó el puntaje APACHE II evidenció una muy ligera inferioridad, aunque no significativa, en la discriminación (AUC = 0,830 ± 0,010; p < 0,001; IC95% 0,811-0,849) con respecto al que utilizó el puntaje SAPS II (AUC = 0,836 ± 0,001; p < 0,001; IC95% 0,818-0,854).

TABLA 1
Características de los pacientes sobrevivientes y fallecidos

Características	Sobrevivientes (n = 2012)	Fallecidos (n = 629)	p
Edad, años (± DE)	48.45 (± 18)	55.54 (± 16,3)	0,003
Sexo masculino – n (%)	1343 (66,7)	410 (65,2)	0,483
Motivos de ingreso – n (%)			
Patología médica	1379 (68,5)	519 (82,5)	<0,001
Politraumatismos	166 (8,3)	33 (5,3)	0,011
Cirugía programada	206 (10,2)	22 (3,5)	<0,001
Cirugía de urgencia	261 (13)	55 (8,7)	0,004
Comorbilidades graves – n (%)			
Hepática	54 (2,7)	32 (5,1)	0,003
Renal	12 (0,6)	15 (2,4)	<0,001
Respiratoria	51 (2,5)	29 (4,6)	0,006
Cardiovascular	45 (2,2)	23 (3,7)	0,040
Inmunodeficiencia	27 (1,3)	30 (4,8)	<0,001
Ventilación mecánica – n (%)	735 (36,5)	459 (72,9)	<0,001
Días de internación, días (± DE)	6.45 (± 8,9)	7.9 (± 9,1)	0,485
APACHE II, puntaje (± DE)	11,74 (± 6,7)	22,03 (± 8,1)	<0,001
SAPS II, puntaje (± DE)	31,20 (± 15,6)	56,63 (± 20,7)	<0,001

DE = desviación estándar.

TABLA 2
Calibración de los puntajes APACHE II y SAPS II

Decil de riesgo	APACHE II				SAPS II			
	n	Muertes observadas	Muertes esperadas	RME	n	Muertes observadas	Muertes esperadas	RME
0,01-0,1	1130	61	60,580	1,007	1112	56	47,118	1,189
0,11-0,2	541	99	80,869	1,224	417	67	61,407	1,091
0,21-0,3	295	73	74,499	0,980	233	51	56,922	0,896
0,31-0,4	194	81	68,280	1,186	158	48	54,606	0,879
0,41-0,5	151	72	68,251	1,055	127	39	56,341	0,692
0,51-0,6	126	81	70,080	1,156	141	60	77,031	0,779
0,61-0,7	80	59	52,850	1,116	84	46	54,656	0,842
0,71-0,8	68	55	51,350	1,071	117	71	88,108	0,806
0,81-0,9	42	35	35,730	0,980	119	74	101,571	0,729
0,91-1,0	14	13	13,110	0,992	133	117	126,514	0,925
ji al cuadrado = 9,3661; DF = 9; valor p = 0,4042					ji al cuadrado = 25,59; DF = 9; valor p = 0,0024			

RME = razón de mortalidad estandarizada; DF = *degrees of freedom* (grados de libertad).

- Desempeño de los puntajes APACHE II y SAPS II para calcular la razón de mortalidad estandarizada en una Unidad de Cuidados Intensivos polivalente de la Argentina

TABLA 3
Áreas bajo la curva con sus respectivos intervalos de confianza

Variables de resultado de contraste	Área	Error estándar ^a	Sig. asintótica ^b	Intervalo de confianza asintótico al 95%	
				Límite inferior	Límite superior
PM APACHE*	0,830	0,010	0,000	0,811	0,849
PM SAPS**	0,836	0,009	0,000	0,818	0,854

^a Bajo el supuesto no paramétrico.

^b Hipótesis nula: área verdadera = 0,5.

* PM APACHE: Probabilidad de muerte predicha por el puntaje APACHE II.

** PM SAPS: Probabilidad de muerte predicha por el puntaje SAPS II.

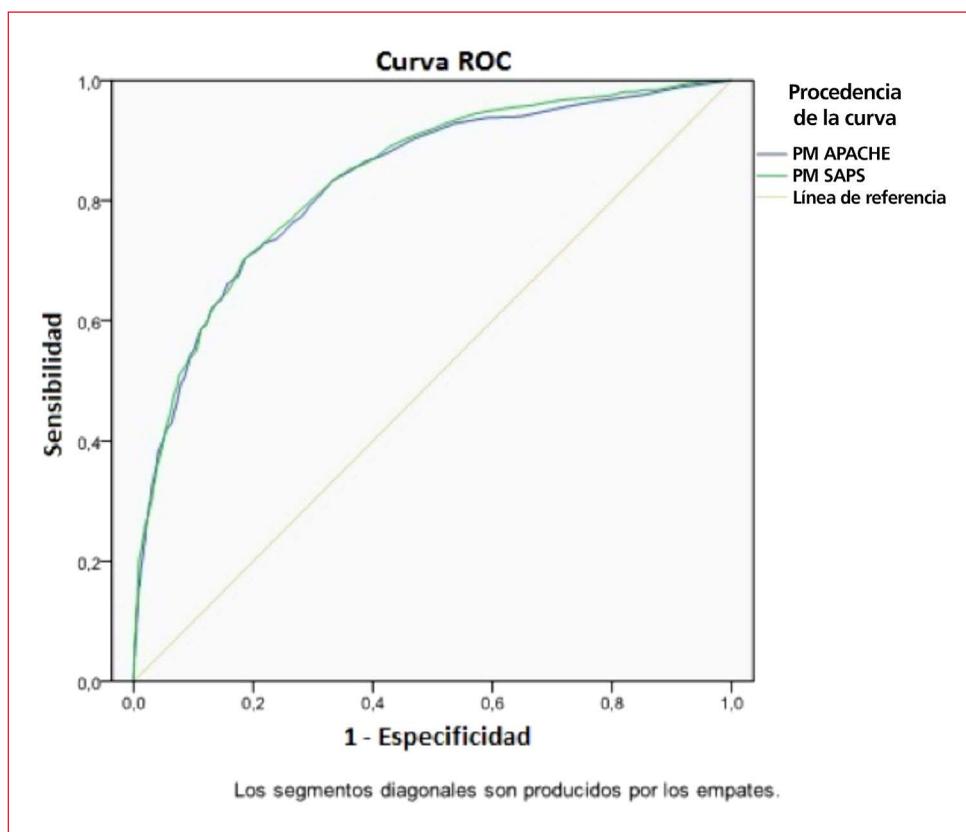


Figura. Curvas ROC para predecir la mortalidad con los puntajes APACHE II y SAPS II.

Discusión

El desempeño de los modelos pronósticos y puntualmente de los puntajes de gravedad depende de la capacidad de predecir, con exactitud, la probabilidad de muerte (calibración) y la de diferenciar entre los que sobrevivirán y los que no lo harán (discriminación).

Es imposible que un modelo tenga simultáneamente perfecta calibración y discriminación. La ca-

libración es la más evaluada y utilizada en la mayoría de los ensayos clínicos y las comparaciones de indicadores de calidad en la UCI. Sin embargo, desde el punto de vista práctico, la discriminación es la de mayor relevancia.¹⁴⁻¹⁶

En nuestro estudio, la calibración del puntaje APACHE II fue superior a la del SAPS II (χ^2 al cuadrado 9,36; $p = 0,404$ y χ^2 al cuadrado 25,59; $p = 0,002$, respectivamente).

Ambos modelos tuvieron buena capacidad de discriminación como se observa en las curvas ROC y AUC, ya que estas últimas se encuentran por encima de 0,80 y no muestran diferencias en los intervalos de confianza.

En nuestro trabajo, el puntaje APACHE II subestimó la mortalidad (21,8% vs. 23,82%), mientras que el puntaje SAPS II la sobrestimó (27,4% vs. 23,82%), por lo que el valor de la RME calculada por APACHE II fue de 1,09 y por SAPS II de 0,87.

Sakr et al evaluaron 1851 pacientes de una UCI quirúrgica de Alemania, comparando, entre otros, el desempeño de los puntajes APACHE II y SAPS II, y hallaron una AUC-ROC de 0,80 y 0,83, respectivamente, pero con pobre calibración para ambos modelos.¹⁷

En Arabia Saudita, Arabi et al estudiaron casi 1000 pacientes de una UCI polivalente y encontraron que los puntajes APACHE II y SAPS II tenían una discriminación aceptable (AUC-ROC de 0,83 y 0,79), pero una mala calibración que no mejoró con el ajuste regional.¹⁸

Livingston, a su vez, estudió el desempeño de cinco modelos predictivos (APACHE II y III, APACHE II ajustado, SAPS II y el *Mortality Probability Model II*) en 10.393 pacientes de 22 Unidades de Escocia. Los resultados fueron uniformes entre todos ellos. Logran una discriminación razonable, pero la calibración es inadecuada sin lograr la validación de ninguno de los puntajes utilizados.¹⁹

En los trabajos antes mencionados, así como en la mayoría de los estudios de este tipo, se observa que la discriminación supera a la calibración, debido a que se aplica un puntaje en una población distinta de la población en la que fue generado.

Castella et al compararon los mismos puntajes que Livingston en un estudio multicéntrico, multinacional (12 países de Europa y los Estados Unidos) y hallaron diferencias entre el APACHE II y el SAPS II a favor de este último. Sin embargo, es importante señalar que más del 70% de los pacientes incluidos en este trabajo pertenecían a los centros de desarrollo del puntaje SAPS, lo que explica su superioridad.²⁰

Del Bufalo et al compararon los puntajes APACHE II y SAPS II para predecir la mortalidad en una UCI respiratoria. El puntaje APACHE II fue ligeramente mejor predictor de resultados que el SAPS II (RME de 0,86 y 0,83, respectivamente).²¹

Un grupo coreano comparó el desempeño de los puntajes APACHE II, IV, SAPS III y un modelo de SAPS III ajustado en 1314 pacientes de una UCI quirúrgica. Obtuvieron una buena discriminación de los puntajes, pero con una pobre calibración.²²

En un estudio colombiano, se comparó el desempeño de los puntajes APACHE II y el SAPS III (original y ajustado) en 2523 pacientes de una UCI polivalente y se encontró una aceptable discriminación para todos

los modelos, pero una pobre calibración para el puntaje APACHE.²³

En Brasil, Serpa Neto et al estudiaron 3333 pacientes de una UCI polivalente y hallaron que el puntaje APACHE II presentaba menor poder de discriminación y calibración con respecto al puntaje SAPS III.²⁴

En 2002, el Comité de Escores de la SATI realizó la validación externa de los puntajes APACHE II y SAPS II en 36 Centros de 11 provincias de la Argentina. Con este fin, empleó la calibración para constatar si la mortalidad predicha por el modelo se correspondía con la mortalidad observada y la discriminación para valorar si el modelo propuesto tenía la capacidad de diferenciar vivos de muertos en la nueva muestra.¹⁰ En dicho estudio, se observó que ninguno de los dos puntajes predecía, en forma adecuada, la mortalidad en la población de nuestro país.

Evaluada la calibración se observó que ambos subestimaban el riesgo de morir, de manera global y más precisamente en los estratos intermedios. En los extremos de gravedad, ambos puntajes sobrestimaban la mortalidad.

Sin embargo, el puntaje APACHE II estratificó adecuadamente a los pacientes de acuerdo con su gravedad, por lo que podría utilizarse en la investigación clínica.

En cuanto a la discriminación, tanto el puntaje APACHE II como el SAPS II tuvieron un buen desempeño similar a los resultados de nuestro trabajo.

En el estudio anterior, la mortalidad real fue del 28%. La probabilidad de muerte promedio predicha fue del 24,6% por el puntaje APACHE II y del 19,3% por el SAPS II. La RME calculada con los resultados mencionados es de 1,13 y 1,45, respectivamente, lo que habla de la mejor estratificación de la gravedad de los pacientes a través del puntaje APACHE II.

En resumen, los trabajos antes citados muestran, en general, una buena discriminación para los puntajes APACHE II y SAPS II, similar a los resultados de nuestro estudio. Sin embargo, hallaron una pobre calibración para ambos puntajes a diferencia de nuestro estudio en el que encontramos una buena calibración para el puntaje APACHE II y una mala para el SAPS II.

Las fortalezas de nuestro trabajo están dadas por el tamaño de la muestra, que dio suficiente poder al estudio, y la calidad de los datos registrados por personal calificado en un programa confiable (SATI-Q).

La primera y más importante debilidad es que el estudio fue realizado en un único Centro y, por lo tanto, no refleja el comportamiento global de dichos puntajes para la Argentina y Latinoamérica. La segunda debilidad es la imposibilidad de reproducir con certeza la composición y "case-mix" de la cohorte expresada en la heterogeneidad de las patologías de ingreso.

- Desempeño de los puntajes APACHE II y SAPS II para calcular la razón de mortalidad estandarizada en una Unidad de Cuidados Intensivos polivalente de la Argentina

Conclusiones

A partir de estos resultados, podemos concluir que el puntaje APACHE II tuvo una discriminación y una calibración buenas, mientras que el puntaje SAPS II tuvo buena discriminación y mala calibración. Por lo tanto, en nuestra UCI, sería razonable el empleo del puntaje APACHE II para calcular la RME.

Bibliografía

1. 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(8): 591-597.
2. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med* 1985; 13(10): 818-829.
3. Le Gall JR, Loirat P, Alperovitch A, et al. A simplified acute physiology score for ICU patients. *Crit Care Med* 1984; 12(11): 975-977.
4. 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(24): 2957-2963.
5. Sociedad Española de Medicina Intensiva, Crítica y Unidades Coronarias. Indicadores de calidad en el enfermo crítico. Actualización 2011, 1-185. Disponible en: www.semicyuc.org/temas/calidad/indicadores-de-calidad.
6. Metnitz PG, Valentin A, Vesely H, et al. Prognostic performance and customization of the SAPS II: results of a multicenter Austrian study. *Simplified Acute Physiology Score. Intensive Care Med* 1999; 25(2): 192-197.
7. Beck DH, Smith GB, Pappachan JV, Millar B. External validation of the SAPS II, APACHE II and APACHE III prognostic models in South England: a multicentre study. *Intensive Care Med* 2003; 29(2): 249-256.
8. Wong DT, Crofts SL, Gomez M, et al. Evaluation of predictive ability of APACHE II system and hospital outcome in Canadian intensive care unit patients. *Crit Care Med* 1995; 23(7): 1177-1183.
9. Markgraf R, Deutschinoff G, Pientka L, Scholten T. Comparison of acute physiology and chronic health evaluations II and III and simplified acute physiology score II: a prospective cohort study evaluating these methods to predict outcome in a German interdisciplinary intensive care unit. *Crit Care Med* 2000; 28(1): 26-33.
10. Comité de Escores de SATI e Investigadores. Validación del APACHE II y SAPS II en la República Argentina. *Medicina Intensiva* 2002; 19(2): 18-28.
11. Settecase CJ, Bagilet DH. Escores pronósticos. En: Bagilet DH, Settecase CJ. *Medicina Crítica. Etiología, fisiopatología, clínica y tratamiento*, Juris; 2011: 1093-1104.
12. Donabedian A. Evaluating the quality of medical care. *Milbank Q* 2005; 83(4): 691-729.
13. Programa SATI-Q: <http://www.hardineros.com.ar/satiq/>.
14. Serrano N. Calibration strategies to validate predictive models: is new always better? *Intensive Care Med* 2012; 38(8): 1246-1248.
15. Diamond G. What price perfection? Calibration and discrimination of clinical prediction models. *J Clin Epidemiol* 1992; 45(1): 85-89.
16. Altman DG, Royston P. What do we mean by validating a prognostic model? *Statistics in Medicine* 2000; 19(4): 453-473.
17. Sakr Y, Krauss C, Amaral AC, et al. Comparison of the performance of SAPS II, SAPS 3, APACHE II, and their customized prognostic models in a surgical intensive care unit. *Br J Anaesth* 2008; 101(6): 798-803.
18. Arabi Y, Haddad S, Goraj R, Al-Shimemeri A, Al-Malik S. Assessment of performance of four mortality prediction systems in a Saudi Arabian intensive care unit. *Crit Care* 2002; 6(2): 166-174.
19. 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(6): 1820-1827.
20. 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(8): 1327-1335.
21. Del Bufalo C, Morelli A, Bassein L, et al. Severity scores in respiratory intensive care: APACHE II predicted mortality better than SAPS II. *Respir Care* 1995; 40(10): 1042-1047.
22. Lee H, Shon YJ, Kim H, Paik H, Park HP. Validation of the APACHE IV model and its comparison with the APACHE II, SAPS 3, and Korean SAPS 3 models for the prediction of hospital mortality in a Korean surgical intensive care unit. *Korean J Anesthesiol* 2014; 67(2): 115-122.
23. Giraldo N, Toro JM, Cadavid C, Zapata F, Jaimes F. Performance of APACHE II and SAPS 3 Regional adaptation in a population of critically ill patients in Colombia. *Acta Med Colomb* 2014; 39: 148-158.
24. Serpa Neto A, Assunção MS, Pardini A, Silva E. Feasibility of transitioning from APACHE II to SAPS III as prognostic model in a Brazilian general intensive care unit. A retrospective study. *Sao Paulo Med J* 2015; 133(3): 199-205.