



Desempeño de los sistemas de puntuación pronóstica en la Unidad de Terapia Intensiva del Hospital Ángeles Pedregal

Performance of the prognostic scoring systems in the Intensive Care Unit of the Hospital Ángeles Pedregal

Christopher Hernández Cortés,¹ Marlet Meyer Talón,¹ Claudia López López,¹ Alinne Lemus Cázares,¹ Silvia del Carmen Carrillo Ramírez,² Juan Gerardo Esponda Prado²

Resumen

Objetivo: Evaluar el desempeño de los sistemas de puntuación pronóstica SAPS III y SOFA en la Unidad de Terapia Intensiva del Hospital Ángeles Pedregal (UTI-HAP). **Material y métodos:** Estudio retrospectivo de cohorte realizado entre el 01 de enero de 2016 y el 30 de junio de 2017. Se calculó la mortalidad pronosticada con SAPS III y SOFA a su ingreso. El desempeño de ambas pruebas fue calificado evaluando su discriminación (habilidad para distinguir entre un paciente que vivirá y uno que morirá; área bajo la curva) y su poder de calibración (comparación entre la mortalidad observada y la presentada; prueba de bondad de ajuste Lemeshow-Hosmer). En el caso de la escala de SOFA, se realizó una evaluación de la sensibilidad y la especificidad. **Resultados:** Nuestra cohorte se conformó de 440 pacientes y presentó una mortalidad de 11.1%. SAPS III: tuvo un área bajo la curva de 0.852 ± 0.030 ; IC95% 0.793-0.912; $p = 0.00$. Prueba de Lemeshow-Hosmer: 0.852 ± 0.30 ; $p = 0.362$. SOFA: área bajo la curva de 0.859 ± 0.025 ; IC95% 0.810-0.908; $p = 0.00$. Sensibilidad de 75% y especificidad de 74%. **Conclusiones:** Tanto SAPS III como SOFA mostraron un desempeño adecuado en la valoración pronóstica de los pacientes de la UTI-HAP.

Palabras clave: Escalas de predicción, mortalidad, unidad de terapia intensiva, SOFA, SAPS III.

Summary

Objective: To evaluate the performance of the SAPS III and SOFA prognostic scoring systems in the Intensive Care Unit of the Ángeles Pedregal Hospital (UTI-HAP). **Material and methods:** Retrospective cohort study conducted between January 1, 2016 and June 30, 2017. Mortality predicted with SAPS III and SOFA was calculated upon admission. The performance of both tests was graded by evaluating their discrimination (ability to distinguish between a patient who will live and one who will die; area under the curve) and their calibration power (comparison between observed mortality versus presented one; Lemeshow-Hosmer goodness-of-fit test). In the case of the SOFA scale, an assessment of sensitivity and specificity was made. **Results:** Our cohort was composed of 440 patients, with a mortality of 11.1%. SAPS III: presented an area under the curve of 0.852 ± 0.030 ; IC95% 0.793-0.912; $p = 0.00$. Lemeshow-Hosmer test: 0.852 ± 0.30 ; $p = 0.362$. SOFA: area under the curve of 0.859 ± 0.025 ; IC95% 0.810-0.908; $p = 0.00$. Sensitivity of 75% and specificity of 74%. **Conclusions:** Both SAPS III and SOFA showed an adequate performance in the prognostic assessment of the patients of the UTI-HAP.

Keywords: Prediction scales, mortality, intensive care unit, SOFA, SAPS III.

INTRODUCCIÓN

Los sistemas de evaluación de severidad o de puntuación pronóstica de la enfermedad como APACHE (Acute Physiology and Chronic Health Evaluation),¹ SAPS (Simplified Acute Physiology Score),² MPM (Mortality Probability Models),³ y SOFA (Sequential Organ Failure Assessment)⁴ fueron elaborados con el propósito de evaluar el pronóstico individual de los pacientes y guiar su cuidado.

Su creación se realiza a través de la selección subjetiva de variables clínicas y fisiológicas, las cuales son adaptables a modelos de regresión logística con el objetivo

¹ Médico residente, Medicina del Enfermo Adulto en Estado Crítico. Facultad Mexicana de Medicina. Universidad La Salle.

² Médico adscrito a la Unidad de Terapia Intensiva.

Hospital Ángeles Pedregal.

Correspondencia:

Dr. Christopher Hernández Cortés
Correo electrónico: chcmex@hotmail.com

Aceptado: 17-11-2018.

Este artículo puede ser consultado en versión completa en:
www.medicgraphic.com/actamedica

de predecir variables significativas. Posteriormente, son expuestos a regresión logística múltiple para determinar las variables que predicen la supervivencia, y por último, son reexaminados de manera prospectiva para validar la exactitud de la predicción.⁵

Se considera que un adecuado sistema de puntuación pronóstica debe presentar las siguientes características:⁵

- 1) *Discriminación*: habilidad de un modelo para distinguir entre un paciente que vivirá y uno que morirá basándose en curvas ROC (verdaderos positivos versus 1-falsos positivos).
- 2) *Calibración*: comparación entre la mortalidad observada y la presentada dentro de los datos de severidad. Se toma como modelo matemático la "bondad de ajuste" de Hosmer-Lemeshow.
- 3) *Confiabilidad*: acuerdo entre y dentro del observador en el uso de cualquier escala de severidad, tomando en cuenta que entre mayor subjetividad, menor confiabilidad. Se utiliza como modelo matemático al índice κ (kappa).
- 4) *Validación del contenido*: una adecuada comprensión del modelo.
- 5) *Rigor metodológico*: con el objetivo de evitar el sesgo.

Con el objetivo de ampliar el panorama de estas escalas, se han utilizado con propósitos administrativos para evaluar el desempeño de las unidades de terapia intensiva y en estudios controlados aleatorizados en investigación médica.⁵ También se han pretendido emplear como criterio para ingreso, para adecuar el grado de monitorización y para decidir el alta de los pacientes. No obstante, no existen datos que demuestren su utilidad en estas indicaciones.⁶

Debido a que las escalas requieren de un mejoramiento continuo, ha sido necesario el desarrollo de múltiples versiones de estas. Esto se puede ver, por ejemplo, para la escala APACHE, con sus versiones II,⁷ III⁸ y IV,⁹ así como para MPM, con diversas variantes de sus versiones que incluyen distintos modelos con distintas mediciones a las cero, 24, 48 y 72 horas,^{3,10-13} y con SAPS, con sus versiones I,² II¹⁴ y III.¹⁵

Dentro de la Unidad de Terapia Intensiva del Hospital Ángeles Pedregal (UTI-HAP), son dos las escalas que se utilizan de manera cotidiana: SAPS III¹⁵ y SOFA.⁴

La escala SAPS en su tercera versión se realizó en 2003 y comenzó a aplicarse en 2005.⁵ Se formuló a partir de 19,577 pacientes de 307 unidades de terapia intensiva entre el 14 de octubre y el 15 de diciembre de 2002, y en su validación se incluyeron 35 países divididos en siete regiones: Austra-asia, Sudamérica y Centroamérica, Europa Central y Europa de Oriente, Europa del Este, Norteamérica, Europa del Norte y Europa del Sur y países del Mediterráneo.¹⁵ Se creó con la necesidad de incluir factores clínicos y no clínicos que se presentan en la unidad de terapia intensiva y que podrían tener un impacto importante

en el pronóstico del paciente.¹⁵ Se representa por la suma aritmética de tres puntuaciones parciales o cajas, que por medio de un modelo matemático logarítmico determinan la probabilidad de muerte al egreso hospitalario:

Caja I. Lo que sabemos de las características del paciente antes de su ingreso a la UTI. Está compuesta por cinco variables: edad, comorbilidades, uso de drogas vasoactivas previamente a la admisión a la UTI y tiempo de estancia en el hospital antes de la admisión a la UTI.¹⁵

Caja II. Lo que sabemos acerca de la circunstancias de admisión a la UTI. Está compuesta también de cinco variables: razón de ingreso a la UTI, admisión planeada o no planeada, estatus quirúrgico al momento de su admisión a la UTI, sitio anatómico de la cirugía y presencia de la infección antes del ingreso a la UTI.¹⁵

Caja III: Lo que sabemos de la presencia y grado de alteración fisiológica al ingreso a la UTI (hasta una hora antes de la admisión). Está compuesta de 10 variables: puntaje más bajo en la escala de Glasgow, frecuencia cardiaca más alta, presión sistólica más baja, bilirrubina más alta, temperatura corporal más alta, creatinina más alta, leucocitos más altos, niveles de plaquetas más bajos, menor concentración de iones de hidrógeno (pH) y presencia de soporte ventilatorio y oxigenación.¹⁵

En cuanto a la escala de evaluación SOFA, fue originalmente descrita como un continuo en la disfunción orgánica en pacientes críticos en el curso de su estancia en la UTI, considerando al fallo orgánico como un estado dinámico cuya descripción de la disfunción debe basarse en variables simples, específicas del órgano en cuestión y rutinariamente disponible en todas partes.³ Se compone de seis sistemas orgánicos (respiratorio, de coagulación, hepático, cardiovascular, neurológico y renal) graduados de cero a cuatro y fue planeada para describir la morbilidad de los pacientes;³ sin embargo, un estudio retrospectivo demostró una mejor relación con la mortalidad.¹⁶

El presente trabajo buscará evaluar el desempeño de los sistemas de puntuación pronóstica SAPS III y SOFA en la Unidad de Terapia Intensiva del Hospital Ángeles Pedregal.

MATERIAL Y MÉTODOS

La investigación se desarrolló como un estudio retrospectivo de cohorte. Se incluyó a todos los pacientes mayores de 18 años ingresados a la Unidad de Terapia Intensiva del Hospital Ángeles Pedregal en un periodo comprendido entre el 01 de enero de 2016 y el 30 de junio de 2017, en quienes se realizó una evaluación SAPS III y/o SOFA a su ingreso. Se excluyeron todos aquellos cuya estancia en la Unidad de Terapia Intensiva fue menor a 24 horas, aquellos ingresados a la Unidad de Cuidados Coronarios, quienes ingresaron a nuestra unidad como parte de una estrategia administrativa

de optimización de espacio y quienes para el 30 de junio de 2017 aún se encontraban dentro de la Unidad de Terapia Intensiva. Asimismo, se excluyó a aquellos que reingresaron a la UTI-HAP en la misma hospitalización y aquellos en quienes se limitaron los esfuerzos terapéuticos. Por último, eliminamos a los pacientes cuyos datos fueron considerados incompletos al momento de la recolección de los mismos por el equipo de investigación.

Al momento del ingreso de la Unidad de Terapia Intensiva del Hospital Ángeles Pedregal, se calculó la mortalidad pronosticada para SAPS III y el puntaje de SOFA. Con el objetivo de evitar el sesgo por operador, se efectuaron las siguientes intervenciones: primero, se estandarizó la fuente de ambas escalas. En el caso de SAPS III, se utilizó la "hoja de cálculo de puntuación SAPS 3" suministrada por el SAPS 3 Outcomes Research Group (SORG),¹⁷ mientras que en el caso del puntaje de SOFA, se realizó con base en la *Tabla 1*, referenciada en su artículo original.⁴

Posteriormente, se realizó una capacitación del personal a cargo de la evaluación del SAPS III y la escala SOFA, que en nuestra unidad son los médicos residentes y adscritos. Esta capacitación se llevó a cabo en un sistema tutorial persona a persona durante un mes. Más tarde, se efectuó una evaluación del tipo *two-point verification*, que consiste en la comprobación de la escala por medio de los residentes de mayor jerarquía o por el médico de guardia —en caso de que la primera valoración hubiera sido hecha por el médico residente de mayor jerarquía— (primer punto de verificación), y posteriormente, por el médico adscrito a cargo de la recepción de guardia, quien al momento de verificar la evaluación plasma su firma en el ingreso del paciente (segundo punto de verificación). Por último, el

cumplimiento de la evaluación de las dos escalas es calificado por el jefe de la unidad.

La validación de los sistemas fue probada evaluando la calibración y discriminación de las pruebas. La discriminación (habilidad de un modelo para distinguir entre un paciente que vivirá y uno que morirá) se realizó por medio de las curvas ROC y su área bajo la curva, con un intervalo de confianza de 95%. La calibración (comparación entre la mortalidad observada y la presentada) se realizó por medio de las curvas entre la mortalidad calculada y la mortalidad esperada, utilizando la prueba de bondad de ajuste Lemeshow-Hosmer para determinar si la discrepancia fue estadísticamente no significativa ($p > 0.05$). Debido a que no existe un estudio que relacione la mortalidad esperada en la escala SOFA, no se podría hacer una adecuada calibración sin recurrir a un sesgo importante. Por lo anterior, en su lugar, se decidió realizar una evaluación de la sensibilidad y la especificidad, con sus respectivos valores predictivos positivos y negativos.

Para el resto del análisis, se utilizó la estadística con medidas de tendencia central y de dispersión para describir las características generales de la cohorte, utilizando la prueba estadística Shapiro-Wilk para establecer la normalidad de cada una de nuestras variables cuantitativas. Se consideró como estadísticamente significativa a una $p < 0.05$. El análisis estadístico en su totalidad fue realizado con el programa SPSS v. 19.

RESULTADOS

Durante el periodo comprendido entre el 01 de enero de 2016 y el 30 de junio de 2017 se reportaron 1,071 ingresos a la Unidad de Terapia Intensiva del Hospital Ángeles.

Tabla 1: Análisis descriptivo de la cohorte (UTI-HAP 2016-2017).

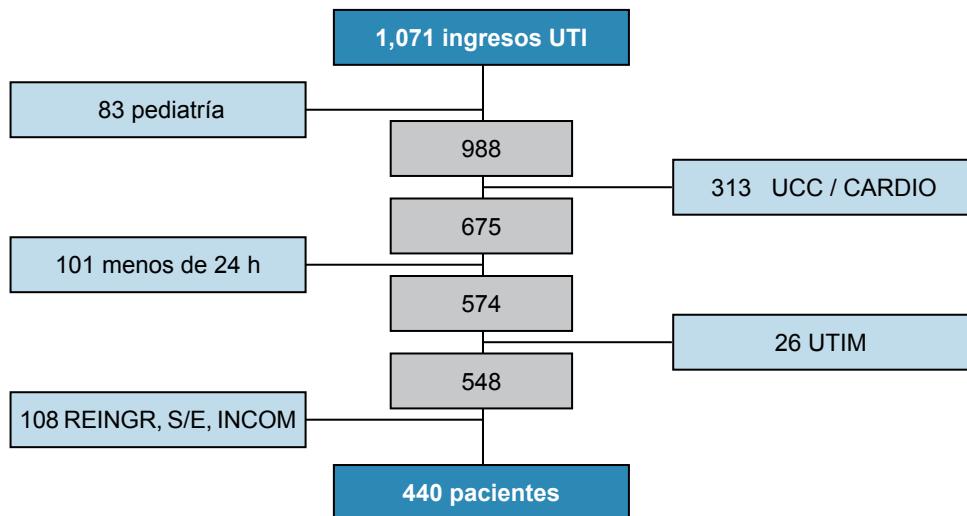
Variable	Frecuencia/mediana	Shapiro-Wilk	p
Sexo	Femenino 180 (40.9%) Masculino 260 (59.1%)		
Edad años	65 años* p25 = 46, p75 = 76, r = 79	0.960	0.00
Días de estancia	2 días* p25 = 1, p75 = 5, r = 73	0.495	0.00
Fuente de ingreso a la UTI	Urgencias 191 (43.4%) Hospitalización 55 (12.5%) UTIM 33 (7.5%) Quirófano 161 (36.6%)		
Condición clínica de ingreso	A. Falla orgánica 224 (50.9%) B. Riesgo de falla orgánica 69 (15.7%) C. Cuidados especiales 147 (33.4%)		

*Mediana. p25 = percentil 25, p75= percentil 75, r = rango.

UTI = Unidad de Terapia Intensiva,
 UCC = Unidad de Cuidados
 Coronarios,
 UTIM = Unidad de Terapia
 Intermedia, REINGR = Reingresos,
 S/E = Sin expediente,
 INCOM = Expedientes incompletos.

Figura 1:

Proceso de selección de pacientes de la cohorte (UTI-HAP 2016-2017).



geles Pedregal. Siguiendo nuestros criterios de inclusión, exclusión y eliminación, establecimos una cohorte de 440 individuos. La razón de pacientes que conformaron parte de la cohorte se demuestra en la *Figura 1*.

Análisis descriptivo de la cohorte

Nuestra cohorte se conformó de 440 pacientes. De ellos, 40.9% ($n = 180$) fueron del sexo femenino y 59.1% ($n = 260$) del sexo masculino. La mediana de edad presentada fue de 65 años, con un percentil 25 de 46 años y un percentil 75 de 76 años, con un rango de 79 años. La mediana de días de estancia en la Unidad de Terapia Intensiva fue de dos días, con un percentil 25 de un día y un percentil 75 de cinco días, con un rango de 73 días. Con respecto a su fuente de ingreso a la Unidad de Terapia Intensiva, 191 provinieron del Servicio de Urgencias, 55 del Área de Hospitalización, 33 de la Unidad de Terapia Intermedia y 161 del Área de Quirófanos. Respecto a la condición clínica de ingreso, estatificamos a los pacientes en condición clínica A o falla orgánica, 224 ingresos (50.9%); condición clínica B o en riesgo de falla orgánica, 69 ingresos (15.7%), y condición clínica C o cuidados especiales, 147 ingresos (33.4%). Este análisis descriptivo de nuestra cohorte, con sus respectivas pruebas de normalidad, se muestra en la *Tabla 1*.

Mortalidad

Se prosiguió a buscar la incidencia de mortalidad en nuestra cohorte; esta fue de 11.1% ($n = 49$).

SAPS III

En nuestro estudio, encontramos que la puntuación de SAPS III presentó una distribución anormal (Shapiro-Wilk

de 0.826, $p = 0.00$), con una mediana de 15 puntos, un percentil 25 de cinco puntos, percentil 75 de 36 puntos, con un mínimo de cero puntos, un máximo de 98 puntos (rango de 98 puntos) y una moda de un punto.

SAPS III. Discriminación

Posteriormente, proseguimos a evaluar la discriminación (habilidad de un modelo para distinguir entre un paciente que vivirá y uno que morirá) de la prueba, por lo que se calculó el área bajo la curva, la cual fue de 0.852 ± 0.030 , con un intervalo de confianza de 95% de 0.793 a 0.912, con una $p = 0.00$. Lo anterior se muestra en la *Figura 2*.

SAPS III. Calibración

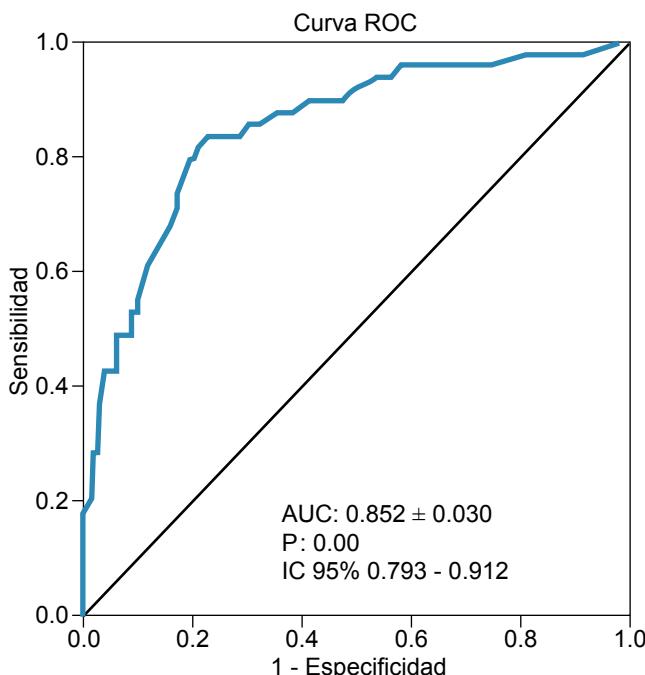
Para evaluar la calibración (comparación de la mortalidad observada contra la presentada dentro de los datos de severidad) del SAPS III en nuestra UTI-HAP se utilizó la prueba de bondad de ajuste Lemeshow-Hosmer, la cual se calculó de 0.852 ± 0.30 con una $p = 0.362$, lo cual traduce que no se encontró una diferencia estadísticamente significativa entre la mortalidad esperada y la presentada en la unidad.

SOFA

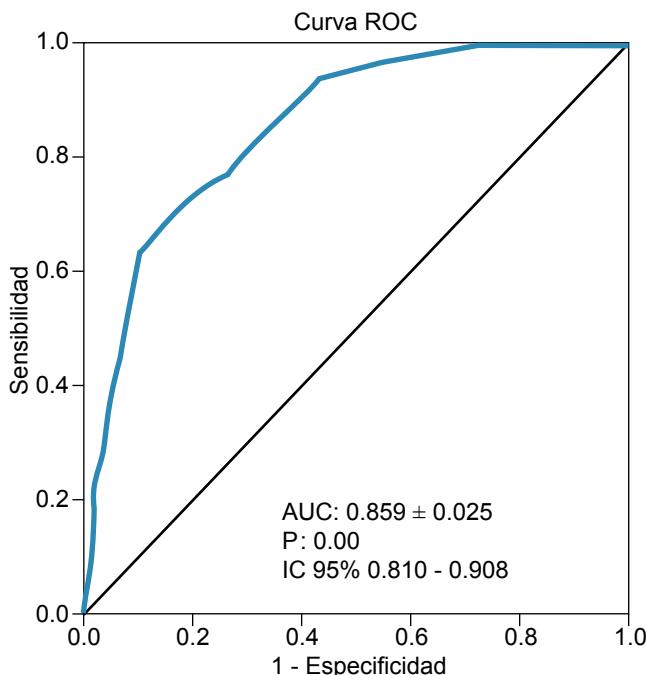
En nuestro estudio encontramos que la puntuación de SOFA presentó una distribución anormal (Shapiro-Wilk de 0.826, $p = 0.00$), con una mediana de dos puntos, un percentil 25 de un punto, percentil 75 de seis puntos. Un mínimo de cero puntos, un máximo de 17 puntos (rango de 17 puntos) y una moda de cero puntos.

Figura 2: Curva ROC para SAPS III (UTI-HAP 2016-2017).

Curva ROC para la cohorte general de SAPS III.
p significativa: < 0.05.

**Figura 3:** Curva ROC para SOFA (UTI-HAP 2016-2017).

Curva ROC para la cohorte general de SOFA.
p significativa: < 0.05.



SOFA. Discriminación

Al evaluar la discriminación (habilidad de un modelo para distinguir entre un paciente que vivirá y uno que morirá) de la prueba, encontramos un área bajo la curva de 0.859 ± 0.025 , con un intervalo de confianza de 0.810 a 0.908, con una $p = 0.00$. Lo anterior se muestra en la [Figura 3](#).

SOFA. Calibración

Tal y como se explicó en la metodología, debido a que no existe un estudio que relacione la mortalidad esperada en la escala de SOFA, no se podría hacer una adecuada calibración sin recurrir a un sesgo importante, por lo que se buscó la sensibilidad y especificidad de la prueba. Se determinó como el punto de mayor discriminación (sensibilidad de 0.776 y 1-especificidad de 0.263) a cinco, con lo que se encontró una sensibilidad de 75% y una especificidad de 74%, así como un valor predictivo positivo de 26% y un valor predictivo negativo de 96% ([Tabla 2](#)).

DISCUSIÓN DE RESULTADOS

Para realizar una evaluación objetiva de SAPS III y SOFA como sistema de puntuación en la Unidad de Terapia In-

Tabla 2: Distribución de la cohorte según el puntaje de SOFA obtenido y la mortalidad presentada (UTI-HAP 2016-2017).

	Defunción	No defunción	Total
SOFA ≥ 5	37	103	140
SOFA < 5	12	288	300
Total	49	391	440

Sensibilidad: $37/49 = 0.755$, Especificidad: $288/391 = 0.736$.

tensiva del Hospital Ángeles Pedregal, se buscó que estas presentaran una adecuada discriminación y calibración.

Respecto a la discriminación, encontramos en este estudio para SAPS III un área bajo la curva de 0.852 ± 0.030 (IC 95% 0.793-0.912; $p = 0.00$), y para SOFA, un área bajo la curva de 0.859 ± 0.025 (IC 95% 0.810 a 0.908; $p = 0.00$). Estos valores no solo muestran una adecuada discriminación, sino que son similares a lo encontrado previamente; podemos ver un ejemplo de esto en el estudio realizado por López-Calero y colaboradores¹⁸ en el que se buscó validar a SAPS III, por lo que se reunió una muestra

de 2,171 pacientes, de la que se obtuvo una mortalidad del 11.6%. Otro estudio mostró un área bajo la curva de 0.90 (IC 95%: 0.89-0.92), pero para SAPS II, y de 0.86 (IC 95%: 0.84-0.88) para SOFA.¹⁹

Respecto a la calibración, se logró demostrar una adecuada calibración para SAPS III en nuestra unidad, siendo esta por bondad de ajuste Lemeshow-Hosmer de 0.852 ± 0.30, con p de 0.362; esto confirma que no existe diferencia significativa entre la mortalidad esperada y la presentada. Curiosamente, en el estudio de López-Caler y su grupo,¹⁸ aun cuando se encontró una adecuada discriminación, no se halló una adecuada calibración por bondad de ajuste de Lemeshow-Hosmer: 31.71, con p < 0.05, lo que demuestra una diferencia estadísticamente significativa entre la mortalidad esperada y la presentada.

Respecto a la calibración de SOFA, creemos que aplicar una bondad de ajuste por Lemeshow-Hosmer podría implicar un sesgo debido a que en la mayoría de los estudios de SOFA, al momento en que se busca relacionar el puntaje con la mortalidad, se presentan los resultados en intervalos, tanto en el puntaje mostrado como en la mortalidad relacionada con ese puntaje. Así se puede ver en el estudio de Afshan Shabir y Muzaffar Maqbool,²⁰ en donde con una población de 190 pacientes, se mostró que aquellos con puntaje de cero a seis presentaban un 0% de mortalidad, mientras que puntajes de seis a nueve tenían 14%, y puntajes mayores a nueve, 85%. Por lo anterior, en nuestro trabajo se decidió utilizar mejor la sensibilidad y especificidad de la prueba, tomando el valor de cinco como el punto de mayor sensibilidad y especificidad (sensibilidad: 0.776 y 1-especificidad: 0.263); con esto se obtuvo una sensibilidad de 75% y una especificidad de 74%.

Respecto a la confiabilidad de la prueba, se ha demostrado que para SOFA es buena entre los médicos, presentando índices kappa que van de casi perfectos para las subclases de coagulación, hepático, circulatorio y renal de la escala (con índices κ—kappa— de 0.851-0.953) a bueno y moderado para la subclase respiratorio y neurológico (con índices κ de 0.552-0.634).²¹ En nuestro estudio no se realizó un estudio de confiabilidad debido a las intervenciones efectuadas en la unidad para mantener el rigor metodológico, que incluyeron tanto la capacitación tutorial persona a persona durante un mes como la evaluación del tipo *two-point verification*. Aun así, creemos que es una buena oportunidad para el futuro.

CONCLUSIONES

Tanto SAPS III como SOFA mostraron un rendimiento adecuado en la valoración pronóstica de los pacientes de la UTI-HAP.

REFERENCIAS

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. 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 (11): 975-977.
3. 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 (7): 519-525.
4. Vincent JL, de Mendonça A, Cantraine F, Moreno R, Takala J, Suter PM et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Working group on "sepsis-related problems" of the European Society of Intensive Care Medicine. *Crit Care Med.* 1998; 26 (11): 1793-800.
5. Russell JA. Assessment of Severity of Illness. In: Hall JB, Schmidt GA, Wood LD. *Principles of critical care*. Part 1. An overview of the approach to and organization of critical care. Chapter 13. 4th ed. McGraw-Hill Inc.; 2015. pp 83-96..
6. Mata VJ. Escalas pronósticas en la unidad de terapia intensiva. *Rev Asoc Mex Med Crit y Ter Int.* 2012; 26 (4): 234-241.
7. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. Apache II: a severity of disease classification system. *Crit Care Med.* 1985; 13 (10): 818-829.
8. 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 (6): 1619-1636.
9. Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Crit Care Med.* 2010; 34 (5): 1297-1310.
10. Teres D, Lemeshow S, Avrunin JS, Pastides H. Validation of the mortality prediction model for ICU patients. *Crit Care Med.* 1987; 15 (3): 208-213.
11. 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 (5): 470-477.
12. 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 (20): 2,478-2,486.
13. 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 (9): 1,351-1,358.
14. 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): 2,957-2,963.
15. Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RAK et al. SAPS 3-From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission. *Intensive Care Med.* 2005; 31 (10): 1345-1355.

16. Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL. Serial evaluation of the SOFA score to predict outcome in critically ill patients. *JAMA*. 2001; 286 (14): 1754-1758.
17. *The SAPS III Outcomes Research Group (SORG)*. [Internet]. Portugal: [actualizado n.d.; citado 25 octubre 2018]. SAPS 3 Admission Score Sheet (MS Excel). Disponible en: <http://www.saps3.org/resources-downloads/user-agreement/downloads/>.
18. López-Calero C, García-Delgado M, Carpio-Sanz J, Álvarez-Rodríguez J, Aguilar-Alonso E, Castillo-Lorente E et al. External validation of the Simplified Acute Physiology Score (SAPS) 3 in Spain. *Med Intensiva*. 2013; 38 (5): 288-296.
19. Williams JM, Greenslade JH, Chu K, Brown AF, Lipman J. Severity scores in emergency department patients with presumed infection: a prospective validation study. *Crit Care Med*. 2016; 44 (3): 539-547.
20. Shabir A, Maqbool M. Accuracy of SOFA score in predicting outcome in medical patients with various diagnosis in intensive care unit in a tertiary care hospital in Northern India. *Int J Contemp Med Res*. 2017; 4 (1): 168-172.
21. Arts DG, de Keizer NF, Vroom MB, de Jonge E. Reliability and accuracy of Sequential Organ Failure Assessment (SOFA) scoring. *Crit Care Med*. 2005; 33 (9): 1988-1993.