

Capacidad discriminativa y costo de los sistemas de calificación de la gravedad de la enfermedad en la Unidad de Terapia Intensiva

Dr. Luis David Sánchez-Velázquez*

RESUMEN

Objetivo. Determinar la capacidad discriminatoria y costo de siete sistemas de severidad (TISS, APACHE II, APACHE III, SAPS I SAPS II) y el Modelo de Probabilidad de Muerte (MPM II).

Diseño. Estudio de cohorte.

Lugar. UCI de un hospital privado de la Ciudad de México.

Intervenciones. Análisis estadístico al ingreso y después de 24 horas.

Mediciones y principales resultados. Los sistemas con mayor capacidad discriminatoria en orden descendente fueron: APACHE III ($85.9 \pm 2.7\%$), SAPS II ($85.6 \pm 2.7\%$), APACHE II ($83.4 \pm 2.8\%$) y MPM II ($81.5 \pm 3.7\%$). Los sistemas más costosos fueron en orden decreciente: APACHE III 64.6, SAPS II 59.8, SAPS I 55 y APACHE II 54.3 en US dólares, respectivamente.

Conclusión. El sistema APACHE II tuvo la mejor capacidad discriminatoria al más bajo costo.

Palabras clave: Sistema de severidad de enfermedades, paciente crítico, pronóstico, costo.

Las Unidades de Terapia Intensiva (UTI) son las áreas hospitalarias donde se generan los más altos costos y se obtienen mayores cifras de morbilidad y mortalidad.¹ Dicha concepción se origina por el hecho de que los pacientes en estado crítico requieren un importante apoyo vital de alta tecnología que no siempre se justifica en todos los casos, además, de las elevadas tasas de infección nosocomial y morbilidad, y la calidad de vida alcanzada por los sobrevivientes. Los sistemas de calificación de la gravedad de la enfermedad (SCGE) se elaboran con parámetros clínicos y paraclínicos y, se ha pro-

SUMMARY

Objective. To determine the discriminative capacity and cost of seven severity systems (TISS, APACHE II, APACHE III, SAPS I and SAPS II) and Mortality Probability Model (MPM II).

Design. Cohort study.

Setting. ICU of a private hospital of Mexico City.

Interventions. Statistical analysis at admission and after 24 hours.

Measurements and main results. The systems with greater discriminatory capacity in descending order were: APACHE III ($85.9 \pm 2.7\%$), SAPS II ($85.6 \pm 2.7\%$), APACHE II ($83.4 \pm 2.8\%$) and MPM II ($81.5 \pm 3.7\%$). The most expensive systems were in decreasing order: APACHE III 64.6, SAPS II 59.8, SAPS I 55 and APACHE II 54.3 US dollars, respectively.

Conclusion. APACHE II system has the better discriminatory capacity at the lowest cost.

Key words: Disease severity systems, ICU patient, prognosis, cost.

puesto por algunos investigadores, que pueden servir como consideraciones accesorias para las decisiones de ingreso de un paciente a la UTI, el momento de su egreso o del retiro de apoyo vital.

Además, el empleo de estos sistemas facilita la categorización de pacientes para la realización de estudios prospectivos multicéntricos y la comparación de resultados de manejo entre diferentes unidades de terapia intensiva.

Inicialmente, Cullen et al en 1974 publicaron el Sistema de Calificación de Intervención Terapéutica (TISS) con el que proponían clasificar a los pacientes críticamente enfermos de acuerdo a la gravedad de su padecimiento, establecer la relación necesaria de enfermera(s): paciente, evaluar el

* Unidad de Terapia Intensiva, Clínica Londres.

empleo de las camas de la UTI y establecer necesidades de las mismas en el futuro.²

En 1981, un grupo de médicos dirigidos por Knauth asignó puntajes de 1 a 4 a la alteración de 34 variables fisiológicas, la edad y la presencia de enfermedades previas, marcando un hito en la historia de la construcción de escalas de gravedad con la creación de la Evaluación de la Fisiología Aguda y del Estado Crónico de Salud (APACHE).³

En 1984 Le Gall et al en Francia, usando sólo 14 variables fisiológicas tomadas del APACHE, crean la Calificación Simplificada de la Fisiología Aguda (SAPS) observando una clasificación correcta de los pacientes en grupos de probabilidad de muerte.⁴ Los anteriores índices fueron el producto del consenso de varios expertos y revisiones bibliográficas.

Lemeshow un año después, empleando un método refinado de asociaciones de variables independientes categóricas, creó un Modelo de Probabilidad de Mortalidad (MPM) en la Universidad de Massachusetts, EUA. con el que reporta una confiabilidad aceptable para la predicción de la gravedad de la enfermedad.⁵

Desde su creación hasta la fecha se han realizado actualizaciones con mejoras de cada sistema, empleando técnicas estadísticas avanzadas para la selección y la asignación del peso de cada variable, estimando el riesgo de mortalidad a través del uso de un modelo de regresión logística múltiple.⁶⁻¹⁰

El propósito de este estudio es encontrar al SCGE más económico y discriminativo a través del cálculo del área bajo la curva característica del receptor operativo (ROC) \pm EE, entendiendo la discriminación como la capacidad de un modelo de separar aquellos pacientes que se ha predicho que sobrevivirán de aquéllos que se ha predicho que morirán.^{11,12}

PACIENTES Y MÉTODOS

Diseño del estudio. Es un estudio de cohorte de pacientes ingresados a la unidad de terapia intensiva con estancia de 24 horas o mayor. Se excluyeron del estudio a pacientes quemados o menores de 16 años de edad. Se les siguió hasta su egreso hospitalario. La información se colectó diariamente en un formato diseñado *exprofeso*, se almacenó y procesó en la hoja de cálculo Excel 97.

Sitio donde se realizó el estudio. La investigación se realizó en la unidad de terapia intensiva multidisciplinaria de la Clínica Londres, hospital general privado, situado en la ciudad de México.

Variables estudiadas y definiciones. Se consignaron las siguientes variables: Demográficas, calificación de la gravedad de la enfermedad de acuerdo a los lineamientos establecidos para cada SCGE (TISS, SAPS, SAPS II, APACHE II, APACHE III y MPM II₀ y II₂₄), el costo derivado del cálculo de cada sistema y el estado vital al momento del alta hospitalaria. La calificación de cada sistema se realizó sólo con el puntaje obtenido de sus tablas respectivas sin considerar las ecuaciones sugeridas para el cálculo de la mortalidad en algunos de ellos (v.g. APACHE II, APACHE III). Por ello no hubo necesidad de adquirir el programa de cómputo APACHE III. El costo derivado del cálculo de cada sistema se cuantificó por el cargo de los paraclínicos requeridos para medir cada puntaje. Los paraclínicos requeridos para la determinación de cada sistema están incluidos en la rutina de laboratorio solicitada al ingreso a la unidad de terapia intensiva de cada paciente. No se midió el tiempo necesario para el cálculo de cada sistema. Para evitar errores de concordancia interobservador en la calidad de los datos, los procesos de calificación, recolección y captura de variables fue realizado por el autor.

Aspectos éticos y de bioseguridad. El proyecto fue autorizado por el Comité de Ética del hospital, quien no consideró necesario el consentimiento informado.

Análisis estadístico. Se emplearon razones para el sexo de los pacientes, medias \pm desviación estándar para la edad y el puntaje de cada SCGE, y proporciones para el estado vital al egreso hospitalario. Se calculó el costo de la elaboración del puntaje de cada SCGE, consignando el costo de cada paraclínico requerido en cada sistema.

Se realizaron los cálculos de intervalos de clase y sus límites de acuerdo a fórmulas convencionales para cada sistema, manteniendo el mismo número de intervalos de clase para todos. Se establecieron la sensibilidad y la especificidad a diferentes puntos de corte para cada SCGE y con ellos se elaboraron las curvas características de receptor operativo (ROC). Dichas curvas se usan actualmente para juzgar la capacidad discriminativa de diversos métodos estadísticos que combinan varios datos para propósitos predictivos. En la mayoría de los casos se grafican e interpretan en forma cualitativa, sin embargo, en el presente estudio se realizará el cálculo del área bajo la curva ROC \pm error estándar (EE) y se expresará en porcentaje.

Cuadro I. Características demográficas de los grupos estudiados.

Parámetro evaluado	Grupo sobreviviente	Grupo fallecido
Número de pacientes	223 (82.0%)	49 (18.0%)
Género femenino	114 (48.9%)	18 (36.8%)
Género masculino	109 (51.1%)	31 (63.2%)
Edad (años)*	58.2 ± 19.6	65.4 ± 19.4
Estancia hospitalaria (días)*	10.9 ± 9.2	7.5 ± 7.4
Estancia en la UTI (días)*	4.1 ± 3.3	4.1 ± 4.0

* Media ± SE.

RESULTADOS

Descripción general de los pacientes. Durante el periodo de estudio ingresaron 283 pacientes al estudio. Se excluyeron del análisis estadístico a 11 pacientes (estancia en la UTI menor de 24 horas). Sesenta y tres pacientes (22.2%) ingresaron para vigilancia postoperatoria. El grupo de estudio tuvo una edad promedio de 60.0 ± 19.6 años (16-96 años), con una relación hombre: mujer de 1:1. La estancia hospitalaria total fue de 10.2 ± 9.2 días, y de estancia en la UTI de 4.1 ± 3.5 días. La mortalidad al egreso hospitalario del grupo total fue del 18.0% ($n = 49$). Las características demográficas de los dos grupos, sobreviviente y fallecido se muestran en el *cuadro I*. No se observó diferencia de acuerdo al género. Los pacientes que fallecieron tuvieron mayor edad, 65.4 años contra 58.2 años en el grupo sobreviviente. La estancia hospitalaria fue mayor en el grupo sobreviviente, 10.9 días contra 7.5 días en el grupo de fallecidos, sin embargo, la estancia en la UTI fue similar en ambos grupos.

Intervalos de clase. Se calcularon los intervalos de clase necesarios de acuerdo a la fórmula de Sturges, $k = 1 + 3.322 (\log_{10}n)$, donde k es el número de intervalos de clase y n el número de valores en el conjunto de datos en observación. Se obtuvieron 8 intervalos de clase para cada sistema de calificación y se establecieron los límites de intervalo de acuerdo a la fórmula $w = R/k$, donde w es la dimensión de los intervalos de clase, R es la amplitud o diferencia entre la observación más pequeña y la más grande dentro de cada conjunto de datos y k es el número de intervalos de clase. Los resultados se ilustran en el *cuadro II*. El número de pacientes sobrevivientes y fallecidos en cada intervalo de clase para cada sistema de calificación de gravedad de la enfermedad se muestra en el *cuadro III*.

Se calculó sensibilidad y especificidad para cada intervalo de clase. Con estos datos se construyeron las curvas ROC y se calcularon las áreas bajo la curva y se expresaron en porcentaje junto con sus errores estándar *cuadro IV*. Se aprecia la mayor área para el APACHE III con $85.89\% \pm 2.62\%$, seguida por las áreas del SAPS II $85.58\% \pm 2.68\%$, y el APACHE II con $83.42\% \pm 2.80\%$, encontrando que entre ellas no hay gran diferencia porcentual. Para los demás sistemas los resultados fueron MPM II₂₄ $81.52 \pm 3.67\%$, SAPS $80.36 \pm 3.25\%$, MPM II₀ $76.43 \pm 3.70\%$ y TISS $75.51 \pm 3.81\%$.

Los costos derivados del cálculo de cada sistema se aprecian en el mismo *cuadro IV*. Los sistemas más económicos fueron el TISS y el MPM II₀, cuyo costo fue nulo, mientras que los más costosos fueron el APACHE III con un total de \$ 646.00, seguido por el SAPS II \$ 598.00, el SAPS \$ 550.00, APACHE II \$ 543.00, MPM II₂₄ \$ 316.00.

Cuadro II. Intervalos de clase y sus límites para cada sistema de calificación de severidad de la enfermedad.

Intervalos de clase	TISS	SAPS	SAPS II	APACHE II	APACHE III	MPM II ₀	MPM II ₂₄
1	< 14	< 6	< 24	< 6	< 27	< 5.0	< 5.0
2	15-21	7-9	25-36	7-12	28-54	-4.9 a -4.0	-4.9 a -4.0
3	22-28	10-12	37-48	13-18	55-81	-3.9 a -3.0	-3.9 a -3.0
4	29-35	13-15	49-60	19-24	82-108	-2.9 a -2.0	-2.9 a -2.0
5	36-42	16-18	61-72	25-30	109-135	-1.9 a -1.0	-1.9 a -1.0
6	43-49	19-21	73-84	31-36	136-162	-0.9 a 0	-0.9 a 0
7	50-56	22-24	85-96	37-42	163-189	0.1 a 1	0.1 a 1
8	> 57	> 25	> 97	> 43	> 190	> 1	> 1

Cuadro III. Número de pacientes sobrevivientes (S) y fallecidos (F) en cada intervalo de clase por sistema de calificación de severidad de la enfermedad.

Intervalos de clase	TISS		SAPS		SAPS II		APACHE II		APACHE III		MPM II ₀		MPM II ₂₄	
	S	F	S	F	S	F	S	F	S	F	S	F	S	F
1	60	3	51	1	92	1	34	0	56	0	10	0	0	0
2	68	6	49	1	60	4	64	1	91	4	87	5	17	0
3	46	10	45	9	36	8	63	8	50	14	46	6	35	0
4	24	12	44	9	20	18	37	16	17	16	33	11	40	3
5	17	8	17	10	9	7	16	9	6	9	24	8	49	4
6	6	4	8	6	6	5	6	8	3	5	20	9	34	8
7	1	3	4	6	0	4	2	4	0	0	3	9	8	8
8	1	3	5	7	0	2	1	3	0	1	0	1	13	10

Cuadro IV. Áreas bajo la curva ± error estándar de cada sistema y su costo.

Sistema	Área bajo la curva ± error estándar (%)	Costo (\$)
TISS	75.51 ± 3.81	0.00
SAPS	80.36 ± 3.25	550.00
SAPS II	85.58 ± 2.68	598.00
APACHE II	83.42 ± 2.80	543.00
APACHE III	85.89 ± 2.62	646.00
MPM II ₀	76.43 ± 3.70	0.00
MPM II ₂₄	81.52 ± 3.67	316.00

DISCUSIÓN

El diseño y aplicación de las escalas de gravedad para predecir morbilidad ha sido una de las preocupaciones en la UTI, pues permiten elaborar planes diagnósticos y de manejo. Así, en el caso de la pancreatitis, la escala de Ranson predice mortalidad y justifica manejos más agresivos. La Escala de Coma de Glasgow orienta a la gravedad y evolución del estado neurológico, permitiendo una vigilancia clínica muy sensible. En el área crítica el desarrollo de los sistemas de calificación de la gravedad de la enfermedad se ha orientado a responder las siguientes preguntas:

1. ¿Son útiles las Unidades de Terapia Intensiva?
2. ¿Puede confiablemente predecirse la mortalidad esperada de un paciente?, y si es así, ¿Pueden tomarse decisiones tales como negar el ingreso a la UTI o retirar el apoyo vital apoyándose en estas predicciones?
3. ¿Puede mejorarse la relación costo/beneficio en las UTI's con el desarrollo de algoritmos basados en índices predictivos de mortalidad?.
4. ¿Son éticas tales consideraciones?.

En el momento actual no es posible responder concluyentemente a ninguna de ellas, sin embargo, la evolución de la medicina crítica ha permitido mejorar la sobrevida de padecimientos que antaño eran mortales en todos los casos.

El incremento en la sensibilidad y la especificidad de los índices pronósticos de mortalidad durante los últimos años ha permitido la comparación en los resultados obtenidos por diferentes unidades de terapia intensiva y ha facilitado la estratificación de pacientes para la elaboración de proyectos de investigación a nivel multicéntrico.

Sin embargo, la sofisticación de algunos de ellos, como el caso del APACHE III que requiere la adquisición de un programa de cómputo y paraclinicos adicionales obliga a incrementar los costos, por lo que el desarrollo de sistemas más económicos sería idóneo, sobre todo para la mayoría de las unidades de terapia intensiva nacionales que requieren de un sistema que reúna ambas características, alta precisión y economía.

En el presente estudio se ha documentado la superioridad del APACHE III al ser el más discriminativo, seguido estrechamente por el SAPS II y APACHE II; sólo se empleó la tabla para el cálculo del APACHE III publicada por los autores. Se considera que al emplear el programa de cómputo del APACHE III el área bajo la curva será mayor dado que incluye variables adicionales como el área de procedencia del enfermo y su categoría diagnóstica, haciendo la diferencia entre el APACHE III y los demás índices más notorios. Estos resultados son congruentes con los reportados por el Grupo Europeo/Norteamericano de Estudio de la Gravedad que reportan las áreas de las curvas ROC para APACHE III, SAPS II y APACHE II

de 86.6%, 85.5% y 84.8%, respectivamente.¹³ En este estudio también se encuentra que la diferencia porcentual en la capacidad discriminativa entre el APACHE III y el APACHE II es de tan sólo 1.8%, mientras que en nuestra investigación fue de 2.4%, ambas cifras sin repercusión clínica ni estadística. Si bien este último estudio no consideró al TISS dentro de su investigación, es necesario considerar que éste no es un sistema diseñado específicamente para evaluar la gravedad de la enfermedad, pero sí indicador indirecto de gravedad de la enfermedad al calificar el apoyo diagnóstico y terapéutico.

En el presente estudio sólo se calculó el costo de los paracéntricos, puesto que el costo derivado de los salarios del personal y aparatos necesarios para el procesamiento de las muestras, es similar para todos los sistemas. No se midió el tiempo de realización de cada sistema, sin embargo, oscila entre 3 y 5 minutos para cada uno.

Desafortunadamente, se encuentra que la mayor capacidad predictiva se asocia también con costos más elevados de elaboración. Así, el APACHE III requiere para su cálculo el empleo de una biometría hemática, química sanguínea, electrólitos séricos, albúmina y bilirrubina total, mientras que el MPM II, por ejemplo, sólo requiere la evaluación de algunas variables por parte del médico para su cálculo.

En los estudios realizados acerca de la predicción de mortalidad, el costo necesario para el cálculo de cada sistema es obviado, por considerarse que los paracéntricos usados forman parte del perfil solicitado al ingreso a la unidad del paciente. Sin embargo, ante la diversidad de sistemas de calificación de gravedad de la enfermedad, las grandes bases de datos, por ejemplo la Base de Datos Multicéntrica en México y The Project Impact de Norteamérica, recolectan la información de más de dos sistemas, incrementando sus tiempos de captura y gastos. El objetivo del estudio fue determinar cuál sistema era el más eficiente y económico con la finalidad de emplear sólo uno en las diversas unidades de nuestro país y con ello economizar en tiempo, dinero y esfuerzo.

En este estudio para evitar la variabilidad interobservador un solo investigador calculó el puntaje de cada sistema en todos los pacientes, si bien esto puede originar un error sistemático, que no se descarta por no haber medido la variabilidad intraobservador, no se considera que sea significativa dado que hay validez externa al comparar los resultados con los del Grupo Europeo/Norteamericano de Estudio de la Gravedad.

El índice predictivo de mayor capacidad discriminativa a un menor costo en este estudio fue el APACHE II, con un costo de 103.00 pesos menor que el APACHE III, el sistema con mayor capacidad discriminativa, del cual difiere en sólo 2.5%. Se recomienda seguir empleando esta escala hasta no encontrar otra opción con mayor capacidad discriminativa y de menor costo.

BIBLIOGRAFÍA

1. Relman AS. Intensive care units: Who needs them? *N Engl J Med* 1980; 302: 965.
2. Cullen DJ, Civetta JM, Briggs BA, Ferrara LC. Therapeutic intervention scoring system: A method for quantitative comparison of patient care. *Crit Care Med* 1974; 2: 57.
3. 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: 591-7.
4. 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: 975-7.
5. 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.
6. Keene AR, Cullen DJ. Therapeutic intervention scoring system: Update 1983. *Crit Care Med* 1983; 11: 1.
7. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II. A severity of disease classification system. *Crit Care Med* 1985; 13: 818-29.
8. Knaus WA, Wagner DP, Draper EA, Zimmerman DP, Bergner M, Bastos PG, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 1991; 100: 1619-36.
9. 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: 2957-63.
10. 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: 2478-86.
11. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143: 29-36.
12. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983; 148: 839-43.
13. Castella X, Artigas A, Bion J, Kari A. The European/North American Severity Study Group. A comparison of severity of illness scoring systems for intensive care unit patients: Results of a multicenter, multinational study. *Crit Care Med* 1995; 23: 1327-35.

Correspondencia:
Dr. Luis David Sánchez-Velázquez.
Durango 50, 3er. Piso.
Col. Roma.
México, D.F. 06700,
E-mail: dbryan@df1.telmex.net.mx