

Comparación de resultados de la cirugía en dos poblaciones ajustadas por riesgo mediante el sistema POSSUM

Fernando Machado^a, Carlos Barberousse^a, Pablo Santiago^a, Enrique Barrios^b y Luis Carriquiry^a

^aClínica Quirúrgica. Hospital Maciel. Unidad de Bioestadística. Departamento de Biofísica. Facultad de Medicina de Montevideo. Montevideo. Uruguay.

^bUnidad de Bioestadística. Departamento de Biofísica. Facultad de Medicina de Montevideo. Uruguay.

Resumen

Objetivo. Validar la aplicación en Uruguay del sistema POSSUM como puntuación predictiva de los resultados de la cirugía y como método para comparar los resultados de la asistencia entre instituciones con características diferentes en Montevideo.

Pacientes y método. Se estudió prospectivamente a 815 pacientes, de 2 instituciones de asistencia diferentes, una pública y otra privada, con asignación de puntuaciones POSSUM, cálculo de morbilidad y mortalidad esperadas de acuerdo a esas puntuaciones (y los de la variante p-POSSUM) y registro de las complicaciones y muertes postoperatorias.

Resultados. Las cifras de morbilidad y mortalidad observadas estaban dentro de los rangos predichos por POSSUM (y p-POSSUM) y el uso de estas puntuaciones mostró que no había una diferencia significativa en los resultados de la cirugía entre ambas instituciones, teniendo en cuenta las diferentes categorías de pacientes y operaciones realizadas.

Conclusiones. Los autores concluyen que POSSUM es una puntuación de riesgo quirúrgico que puede usarse en diferentes contextos quirúrgicos y que será muy útil en el futuro para comparar diferentes resultados de la cirugía entre instituciones, regiones y países de todo el mundo.

Palabras clave: Índices de gravedad. Riesgo quirúrgico. Sistema POSSUM.

COMPARISON OF SURGICAL OUTCOMES IN TWO POPULATIONS WITH RISK ADJUSTMENT USING THE POSSUM SYSTEM

Objective. To validate the application of the Physiological and Operative Severity Score for the enumeration of Mortality and morbidity (POSSUM) system in Uruguay in order to validate its broader utility not only for predicting surgical results but also for comparing surgical outcomes in two different healthcare institutions in Montevideo.

Patients and method. A total of 815 patients in 2 institutions (one public and the other private) were prospectively studied. A POSSUM score was assigned, expected morbidity and mortality according to POSSUM [(and its variant, Portsmouth PSSUM (p-POSSUM))] were calculated, and postoperative complications and deaths were registered.

Results. Observed morbidity and mortality were within the range predicted by POSSUM (and p-POSSUM) application. The use of these scores showed no significant differences in surgical outcomes between the two institutions, considering the different mix of patients and interventions.

Conclusion. The authors conclude that the POSSUM scoring system can be used in different surgical settings. In the future, this system will be highly useful to compare different surgical outcomes among institutions, regions and countries all over the world.

Key words: Severity illness index. Surgical risk. POSSUM system.

La Comisión Sectorial de Investigación Científica (CSIC) de la Universidad de la República ha financiado en parte este trabajo.

Correspondencia: Dr. C. Barberousse Carrara.
Simón Bolívar, 1359, apdo. 801. CP11300 Montevideo. Uruguay.
Correo electrónico: carbarbe@hotmail.com

Manuscrito recibido el 10-11-2005 y aceptado el 4-7-2006.

Introducción

Los recientes cambios sociales están teniendo un profundo impacto en la práctica médica, en general, y quirúrgica, en particular. El cirujano ha perdido parte de su protagonismo y los resultados de la cirugía se toman en consideración y debaten no sólo en publicaciones espe-

cializadas, sino también diariamente en la prensa ordinaria. Esta situación ha llevado en las últimas décadas al desarrollo y a la aplicación de sistemas objetivos de auditoría en los que se puede establecer el margen de riesgo para un paciente y una intervención determinados y permitir de esta manera una evaluación más equilibrada del resultado final de la práctica quirúrgica.

Estos sistemas, que utilizan puntuaciones objetivas, pueden predecir el riesgo de la cirugía para una determinada población de pacientes. A su vez, permiten comparar los resultados quirúrgicos entre servicios o cirujanos dentro de una misma institución o entre diferentes instituciones a partir de la medida de morbilidad y mortalidad para situaciones parecidas.

Desde la introducción del sistema ASA de la Sociedad Americana de Anestesiología, se han desarrollado múltiples sistemas de evaluación quirúrgica y del riesgo general de los pacientes, y probablemente el sistema APACHE sea el más utilizado en la actualidad.

El sistema POSSUM (Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity) lo desarrolló Copeland et al¹ en 1991 como un método de estratificación de todos los tipos de pacientes quirúrgicos de acuerdo con su grado de riesgo, y en el que únicamente se utilizan hallazgos clínicos, exámenes de laboratorio y los datos del protocolo operatorio, de modo que puede utilizarse en una gran variedad de situaciones diferentes.

Se desarrolló a partir de un análisis aleatorizado, multivariado, de 62 variables diferentes, durante 6 meses en una etapa inicial y luego con 35 variables durante un período similar, antes de llegar a su actual formulación. Este índice permite evaluar los rangos esperados de morbilidad y mortalidad ajustadas por riesgo para una determinada situación, a partir de la consideración de 2 puntuaciones: una fisiológica, con 12 factores, y la otra de gravedad operatoria, que utiliza 6 factores.

Distintos estudios²⁻⁴ han validado este sistema, especialmente en el Reino Unido, y, en menor medida, en otros países europeos, que tanto consideran su versión original como la modificación metodológica realizada por Whiteley et al⁵ en Portsmouth (p-POSSUM).

Pacientes y método

Se estudiaron 2 series consecutivas de pacientes pertenecientes a 2 unidades quirúrgicas diferentes, entre febrero de 1998 y junio de 2000. La unidad 1 correspondió a un servicio de cirugía general de un Hospital público en el que se realiza docencia universitaria (Clínica Quirúrgica 2, Hospital Maciel). La unidad 2 correspondió a un servicio de cirugía general en un centro de asistencia de salud privado.

Se incluyó en el estudio a todos los pacientes mayores de 15 años, tanto los sometidos a procedimientos quirúrgicos de elección, como de urgencia. Su seguimiento se extendió hasta los 30 días del postoperatorio o hasta su muerte, si ésta ocurrió dentro de este período. Únicamente se excluyó a los pacientes a los que se operó de modo ambulatorio y a los que requirieron procedimientos muy complejos, con participación de equipos multidisciplinarios en su tratamiento.

La información se recolectó de modo prospectivo en una hoja de recogida de datos similar a la publicada por Copeland et al¹. Se registraron datos demográficos, diagnóstico clínico e intraoperatorio, procedimientos quirúrgicos realizados y evolución final de los pacientes.

La muerte intrahospitalaria o dentro los primeros 30 días del postoperatorio se aceptó como definición de mortalidad. Se aceptaron las definiciones de complicación postoperatoria establecidas por Copeland en su publicación original.

A cada paciente se le asignó una puntuación fisiológica antes de la intervención quirúrgica y una puntuación de gravedad operatoria después de ella. Los datos de la puntuación fisiológica se recogieron al momento del ingreso del paciente, una vez que se contó con el resultado de los exámenes de laboratorio. La puntuación de gravedad operatoria se pudo completar una vez definidos todos los aspectos inherentes a la cirugía requeridos por ésta: entidad del sangrado, presencia o no de contaminación peritoneal, etc. No hubo diferencias entre las unidades respecto a la forma de preparación preoperatoria de los pacientes.

La puntuación fisiológica tomó en cuenta 12 factores (edad, estado cardíaco, estado respiratorio, presión arterial, pulso, escala de Glasgow, hemoglobina, leucocitosis, nitrógeno ureico en sangre, ionograma y electrocardiograma) y se establecieron 4 grados con asignación de puntuaciones de modo exponencial (1, 2, 4, 8) con el consiguiente rango entre 12 y 84 puntos (tabla 1). La puntuación de gravedad operatoria incluyó 6 factores (tipo y número de procedimientos, pérdida total estimada de sangre, contaminación peritoneal, presencia de malignidad y oportunidad de la cirugía).

En cuanto a la apreciación de la entidad del procedimiento quirúrgico efectuado (cirugía menor, moderada, mayor o mayor +), se utilizaron los criterios originales; de esta forma, la cirugía de moderada gravedad incluyó: apendicectomía, colecistectomía, mastectomía; ejemplos de cirugía mayor fueron: resección intestinal, colecistectomía con coledocotomía, cirugía vascular periférica o amputación mayor. Se consideró como "cirugía mayor+" la cirugía aórtica, la cirugía de resección rectal, la resección pancreática o hepática y la esofagectomía.

Se asignó una puntuación en escala exponencial similar a la puntuación fisiológica, resultando un rango de 6 a 44 puntos (tabla 2). La puntuación POSSUM total tuvo un rango de 18 a 128 puntos.

Toda la información fue recolectada manualmente en formularios impresos y posteriormente transferidos a una planilla de Microsoft Excel (Microsoft Corporation, Washington, District of Columbia, USA) para su análisis.

El riesgo individual de morbilidad se calculó a partir de la ecuación: $\text{Log}_e (R/ [1-R]) = -5,91 + (0,16 \times \text{puntuación fisiológica}) + (0,19 \times \text{puntuación de gravedad operatoria})$ propuesto por Copeland et al y validado en publicaciones previas.

El riesgo individual de mortalidad se calculó a partir de la ecuación: $\text{Log}_e (R/ [1-R]) = -9,37 + (0,19 \times \text{puntuación fisiológica}) + (0,15 \times \text{puntuación de gravedad operatoria})$, propuesto en p-POSSUM, que se ha considerado más seguro en los estudios ulteriores⁶.

Se compararon las tasas de morbilidad y mortalidad entre las unidades y se efectuó, además, mediante un análisis lineal, una estratificación en grupos de riesgo en la que para cada estrato se comparó la morbilidad y la mortalidad observadas/esperada.

Las comparaciones se ajustaron mediante el test χ^2 para las variables cualitativas y mediante el test de la t de Student para las variables cuantitativas. Las diferencias de distribución de las puntuaciones entre ambas unidades quirúrgicas se ajustaron mediante el test de Kolmogorov-Smirnov. El nivel de significación utilizado fue del 5%.

Para evaluar la validez externa de los modelos ajustados hay que considerar 2 aspectos: la discriminación y la calibración. La discriminación es la capacidad del modelo para distinguir entre 2 individuos con resultados opuestos y puede cuantificarse mediante el área bajo la curva ROC⁷.

Puede demostrarse que cuanto mayor sea esta área, tanto mejor será la capacidad discriminadora del modelo.

La calibración establece el grado de semejanza entre el resultado observado y el predicho por el modelo, tomando en cuenta la evaluación de las probabilidades de ocurrencia del fallo en las distintas categorías (deciles). Esto se cuantifica mediante la prueba de Hosmer y Lemeshow (HL)⁸, que tiene como resultado el valor de un coeficiente cuya significación puede ensayarse mediante una prueba de χ^2 . Cuanto mayor sea el valor de p en este test, tanto mejor será la calibración del modelo.

El área bajo la curva ROC se calculó a partir de la población total (ambas unidades) para analizar la capacidad de discriminación de los modelos de ajuste⁷. La calibración del modelo se evaluó mediante el test estadístico χ^2 de HL⁸.

Resultados

Se estudió a un total de 815 pacientes, 471 pertenecientes a la unidad 1 y 344 a la unidad 2. La distribución por edad, sexo y oportunidad de la cirugía fue similar en

TABLA 1. POSSUM. Puntuación fisiológica

	Puntuación			
	1	2	4	8
Edad (años)	≤ 60	61-70	≥ 71	
Signos cardíacos	Sin falla	Diuréticos, digoxina, tratamiento anti-hipertensivo o por cardiopatía isquémica	Edema periférico; tratamiento con warfarina	Hipertensión venosa yugular
Radiografía de tórax			Cardiomegalia límite	Cardiomegalia
Historia respiratoria	Sin disnea	Disnea de esfuerzo	Disnea limitante	Disnea de reposo
Radiografía de tórax		EPOC leve	EPOC moderado	Fibrosis o consolidación
Tensión sistólica (mmHg)	110/130	131/170	≥ 171	-
Pulso (lat/min)	50-80	81-100 40-49	101-120	≥ 121
Escala de coma de Glasgow	15	12-14	9-11	≤ 8
Hemoglobina (g/100 ml)	13-16	11,5-12,9 16,1-17,0	10,0-11,4 17,1-18,0	≥ 18,1
Leucocitosis ($\times 10^{12}/l$)	4-10	10,1-20,0 o 3,1-4,0	≥ 20,1 o ≤ 3,0	
Urea (mmol/l)	≤ 7,5	7,6-10,0	10,1-15,0	≥ 15,1
Sodio (mmol /l)	≥ 136	131-135	126-130	≤ 125
Potasio (mmol/l)	3,5-5,0	3,2-3,4 o 5,1-5,3	2,9-3,1 o 5,4-5,9	≤ 2,8 o ≥ 6,0
Electrocardiograma	Normal		Fibrilación auricular	Cualquier arritmia o ≥ 5 extrasístoles/min Ondas Q o cambios ST/T

EPOC: enfermedad pulmonar obstructiva crónica.

TABLA 2. POSSUM. Puntuación de gravedad operatoria

	Puntuación			
	1	2	4	8
Gravedad operatoria	Menor	Moderada	Mayor	Mayor+
Múltiples procedimientos	1		2	> 2
Pérdida total de sangre (ml)	≤ 100	101-500	501-999	≥ 1.000
Contaminación peritoneal	Ninguna	Menor (fluido seroso)	Pus localizado	Contenido intestinal libre, sangre o pus
Presencia de malignidad	Ninguna	Sólo tumor primario	Metástasis ganglionares	Metástasis a distancia
Modo de cirugía	Electiva		Reanimación urgente > 2 h intervención quirúrgica < 24 h después de la admisión	Emergencia (cirugía inmediata) < 2 h desde la admisión

ambas unidades (tabla 3). El tipo de cirugía realizada también era comparable en ambas, a pesar de un predominio de la cirugía colorrectal y mamaria en la unidad 2 y de la cirugía de trauma en la unidad 1 (tabla 4).

La mortalidad fue de 3,6% en la unidad 1 y de 2,9% en la unidad 2, diferencia no significativa ($\chi^2 = 0,31$, $p > 0,05$).

La morbilidad fue de 27,3% para la unidad 1 y de 30,5% para la unidad 2, diferencia que tampoco fue significativa ($\chi^2 = 0,95$, $p > 0,05$).

Como se ha expresado antes, la capacidad discriminante del modelo se evaluó mediante el cálculo del área bajo la curva ROC (tabla 5).

La distribución de las puntuaciones fisiológica y de gravedad operatoria presentó curvas similares en ambas poblaciones (tabla 6).

Si se considera la población total de pacientes estratificados de acuerdo con los deciles de probabilidad, el modelo en general predijo un número mayor de muertes esperadas en relación con las finalmente observadas, excepto en el segundo decil (tabla 7). De todas formas, el modelo fue un predictor aceptable de mortalidad (HL = 9,21 $p = 0,10$) pero un predictor menos seguro de morbilidad (HL = 82,5 $p < 0,001$) (tabla 8).

Un grupo importante de pacientes en ambas poblaciones se encontró dentro de puntuaciones bajas; es decir,

TABLA 3. Distribución demográfica de los pacientes

	Unidad 1	Unidad 2	p
Edad (media)	50,7	58,5	NS
Sexo V : M	248 : 223	154 : 190	NS
Cirugía urgente (%)	175 (37)	140 (41)	
Cirugía electiva (%)	296 (63)	204 (59)	NS

M: mujer; NS: no significativo; V: varón.

TABLA 4. Operaciones realizadas

	Unidad 1	Unidad 2	p
Pared abdominal	64 (13,6)	43 (12,5)	NS
Mama	3 (0,6)	25 (7,3)	< 0,05
Tiroides	15 (3,2)	13 (3,8)	NS
Vesícula biliar	100 (21,2)	59 (17,2)	NS
Gastroduodenal	25 (5,3)	11 (3,3)	NS
Intestino delgado	20 (4,3)	20 (5,9)	NS
Hepato-bilio-pancreática	53 (11,3)	34 (10,0)	NS
Esófago	4 (0,8)	5 (1,5)	NS
Colorrectal	32 (6,8)	47 (13,7)	0,002
Apéndice	46 (9,8)	31 (9,0)	NS
Amputaciones	21 (4,4)	10 (3,0)	NS
Vascular	31 (6,6)	20 (5,8)	NS
Trauma	17 (3,6)	0 (0)	-
Miscelánea	40 (8,5)	22 (6,4)	NS
Total	471 (100)	344 (100)	

NS: no significativo.
Los valores entre paréntesis son porcentajes.

TABLA 5. Área bajo la curva ROC

ABC	LIIC	LSIC
0,71	0,66	0,75
0,81	0,72	0,89

ABC: área bajo la curva ROC; IC: intervalo de confianza; LIIC: límite inferior del IC del 95%; LSIC: límite superior del IC del 95%.

TABLA 6. Distribución de las puntuaciones POSSUM

Puntuación POSSUM*	Unidad 1 (%)	Unidad 2 (%)
18-25	187 (40)	138 (40)
26-35	199 (42)	134 (39)
36-45	62 (13)	47 (14)
46-55	17 (4)	21 (6)
56-65	6 (1)	4 (1)
Total	471	344

ks = 0,0486, p = 0,6975

ks: test de Kolmogorov-Smirnov (véase el texto).

*Suma de la puntuación fisiológica más la puntuación de gravedad operatoria.

pacientes de bajo riesgo, sometidos a operaciones de complejidad baja. La media de la puntuación fisiológica fue de 17,68 para la unidad 1 y de 17,89 para la unidad 2 (no significativa). La puntuación de gravedad operatoria fue igualmente similar: 11,55 y 11,47, respectivamente.

Sin embargo, si se tiene en cuenta la evolución final de los pacientes, se encontró una marcada diferencia en la media de la puntuación POSSUM para los pacientes que tuvieron una buena evolución postoperatoria y fueron da-

TABLA 7. Estadístico de Hosmer-Lemeshow (HL) para mortalidad

	Número	Observados	Esperados	O/E
0,0 < R ≤ 0,1	729	9	11,239481	0,8007487
0,1 < R ≤ 0,2	38	6	5,339956	1,1236049
0,2 < R ≤ 0,3	20	5	5,102255	0,9799588
0,3 < R ≤ 0,4	13	3	4,343168	0,6907401
0,4 < R ≤ 0,5	4	0	1,742062	0,0000000
0,5 < R ≤ 0,6	5	1	2,809651	0,3559160
0,6 < R ≤ 1,0	6	3	4,576594	0,6555093

HL: 9,21; O/E: relación de muertes observadas frente a esperadas.
p = 0,10.

Se presentan las muertes observadas y las esperadas. Se estratificó a los pacientes en deciles de riesgo (véase p. 142-3 en Hosmer y Lemeshow⁸), donde R representa la probabilidad de muerte en cada grupo. Los pacientes con probabilidad mayor de 0,6 se consideraron como un grupo.

TABLA 8. Estadístico Hosmer-Lemeshow (HL) para morbilidad

	Número	Observados	Esperados	O/E
0,0 < R ≤ 0,1	143	17	11,18372	1,5200660
0,1 < R ≤ 0,2	209	30	29,42145	1,0196643
0,2 < R ≤ 0,3	122	45	29,84714	1,5076822
0,3 < R ≤ 0,4	84	22	29,30025	0,7508468
0,4 < R ≤ 0,5	72	27	32,26561	0,8368042
0,5 < R ≤ 0,6	46	22	24,90864	0,8832275
0,6 < R ≤ 1,0	139	71	109,53273	0,6482081

HL: 82,51288; O/E: relación de muertes observadas frente a esperadas.
p < 0,001.

Se presentan complicaciones observadas y esperadas. Los pacientes se han estratificado en deciles de riesgo (véase tabla 7), donde R representa la probabilidad de morbilidad en cada grupo. Los pacientes con probabilidad mayor de 0,6 se consideraron como un grupo.

dos de alta (unidad 1: 28,70; unidad 2: 28,90) comparados con los que murieron después de la intervención quirúrgica (unidad 1: 43,88; unidad 2: 44,40), sin diferencias significativas entre ambas unidades (p = 0,90).

El análisis lineal (tablas 9-12) de la morbilidad y la mortalidad observada/esperada mostró en ambas unidades resultados globales similares, y se confirmaron las predicciones de la ecuación POSSUM de morbilidad y p-POSSUM de mortalidad.

En la tabla 13 se muestra la relación de mortalidad estandarizada, que expresa la relación observado/esperado de la mortalidad total para cada unidad.

Discusión

Las tasas brutas de morbilidad y mortalidad son medidores poco fieles de los resultados de la intervención quirúrgica, ya que no consideran los caracteres particulares de la población, así como tampoco la complejidad del procedimiento a realizar. En los últimos años, han surgido como respuesta a esta deficiencia sistemas más adecuados para evaluar los resultados de la cirugía. En algunas revisiones se han analizado los diferentes sistemas de evaluación de riesgo y pronósticos, propuestos en las últimas décadas⁹. Pese a ser uno de los sistemas de aparición más reciente —a principios de la década de los años noventa—, el POSSUM ha sido validado como un instrumento muy útil y seguro de un muy amplio espectro

TABLA 9. Muertes observadas/esperadas en la unidad 1. Análisis lineal

Mortalidad (%)	Número de pacientes	Mortalidad media (%)	Muertes observadas	Muertes esperadas, p-POSSUM	Muertes observadas/esperadas
< 10	427	1,58	6	6,75	0,89
10-19	21	14,15	5	2,97	1,68
20-29	9	25,68	1	2,31	0,43
30-39	8	32,39	2	2,59	0,77
40-49	1	40,85	0	0,41	0,00
50-59	2	55,72	0	1,11	0,00
60-69	1	65,70	1	0,66	1,52
70-79	0	-	-	-	-
80-89	2	86,95	2	1,74	1,15
≥ 90	0	-	-	-	-
Total	471	3,94	17	18,56	0,92

TABLA 10. Morbilidad observada/esperada en la unidad 1. Análisis lineal

Morbilidad (%)	Número de pacientes	Morbilidad media (%)	Morbilidad observada	Morbilidad esperada, p-POSSUM	Morbilidad observada/esperada
< 10	74	8,02	10	5,94	1,68
10-19	126	13,83	23	17,42	1,32
20-29	68	24,65	25	16,76	1,49
30-39	55	35,12	14	19,32	0,72
40-49	45	44,52	13	20,04	0,65
50-59	31	53,92	12	16,71	0,72
60-69	21	64,00	7	13,44	0,52
70-79	16	73,90	6	11,82	0,51
80-89	19	85,46	8	16,24	0,49
≥ 90	16	94,79	11	15,17	0,73
Total	471	32,45	129	152,53	0,85

TABLA 11. Muertes observadas/esperadas en la unidad 2. Análisis lineal

Mortalidad (%)	Número de pacientes	Mortalidad media (%)	Muertes observadas	Muertes esperadas, p-POSSUM	Muertes observadas/esperadas
< 10	302	1,46	3	4,40	0,68
10-19	17	13,82	1	2,35	0,43
20-29	11	25,45	4	2,80	1,43
30-39	5	35,20	1	1,76	0,57
40-49	3	44,67	0	1,34	0,00
50-59	2	54,50	1	1,09	0,92
60-69	3	63,33	0	1,90	0,00
70-79	0	-	-	-	-
80-89	1	88,00	0	0,88	0,00
≥ 90	0	-	-	-	-
Total	344	4,80	10	16,51	0,61

TABLA 12. Morbilidad observada/esperada en la unidad 2. Análisis lineal

Morbilidad (%)	Número de pacientes	Morbilidad media (%)	Morbilidad observada	Morbilidad esperada, p-POSSUM	Morbilidad observada/esperada
< 10	62	7,35	7	4,56	1,54
10-19	89	14,09	6	12,54	0,48
20-29	55	24,04	21	13,22	1,59
30-39	28	34,18	7	9,57	0,73
40-49	25	44,68	13	11,17	1,16
50-59	17	53,41	11	9,08	1,21
60-69	22	64,32	11	14,15	0,78
70-79	15	74,67	7	11,20	0,62
80-89	12	84,58	10	10,15	0,99
≥ 90	19	94,47	12	17,95	0,67
Total	344	33,02	105	113,58	0,92

TABLA 13. Relación de mortalidad estandarizada (SMR, *standardized mortality ratio*, razón de observados/esperados para datos totales de mortalidad) para cada unidad. Resultados de la calibración de los datos de mortalidad por unidad

	Unidad 1	Unidad 2
Estadístico HL	8,13	5,61
Valor p	0,42	0,69
SMR	0,92	0,60

HL: Hosmer y Lemeshow.

de la actividad quirúrgica. Éste ofrece 3 ventajas en relación con el APACHE II, probablemente el sistema más utilizado en los últimos años: la consideración tanto de la mortalidad como de la morbilidad quirúrgica, su capacidad para usarlo en todo tipo de pacientes, y no sólo en los internados en unidades de cuidados intensivos o similares, y su fácil accesibilidad, dado que requiere únicamente la recolección de datos clínicos y de laboratorio simples. Se ha considerado superior al APACHE II incluso en pacientes internados en unidades de alta dependencia¹⁰. Pero, al igual que el POSSUM, la mayoría de los estudios de validación se ha diseñado en el Reino Unido y muy pocos en otros países europeos.

Una revisión en PubMed reveló únicamente un estudio publicado acerca de la aplicación del POSSUM en sistemas de control sanitario en países subdesarrollados, igualmente con muy buenos resultados¹¹. Pero, como sus autores reconocen, la mayoría de los cirujanos en Malasia ha estudiado y obtenido sus títulos de posgrado en el Reino Unido y se rigen por el sistema británico de atención.

De este modo, hemos decidido evaluar la aplicabilidad y la validez del POSSUM en un país de Sudamérica, con una tradición diferente en el cuidado sanitario y en el entrenamiento quirúrgico, para así reafirmar la universalidad del sistema.

Algunos autores británicos¹² han cuestionado la aplicabilidad del POSSUM en poblaciones quirúrgicas corrientes y han encontrado en su experiencia que sólo el 30% de los pacientes operados se han asignado adecuadamente a una puntuación POSSUM determinada. Nuestra experiencia ha sido totalmente distinta: los datos clínicos y operatorios, así como los estudios de laboratorio, forman parte de la información corrientemente requerida en la evaluación preoperatoria de todo paciente sometido a cirugía. En lo que respecta al cuestionamiento acerca de los estudios de laboratorio requeridos, quizás el ionograma pudo haberse excluido en la evaluación de pacientes de bajo riesgo, pero, en la actualidad, dado que la mayor parte de los estudios se realizan con máquinas automáticas, éste no supone un esfuerzo adicional ni aumenta significativamente los costes.

Desde el inicio seleccionamos la ecuación POSSUM original para el estudio de la morbilidad y la ecuación p-POSSUM para el estudio de la mortalidad. La aplicación de la ecuación POSSUM para analizar la mortalidad sistemáticamente sobrestima la mortalidad global, especialmente en el grupo de menor riesgo. A pesar de que esta sobrestimación puede prevenirse mediante el uso de un

análisis exponencial, ello exigiría un trabajo matemático adicional que no estaría justificado.

El análisis de la población global reveló un ajuste mejor con la ecuación p-POSSUM para los datos de mortalidad (HL = 9,21), en comparación con la aplicación del POSSUM para la morbilidad (HL = 82,5).

También intentamos utilizar el POSSUM para establecer una comparación válida entre la evolución postoperatoria en 2 grupos diferentes de pacientes, en diferentes situaciones de atención sanitaria. En Uruguay, se ha expresado, y en general se acepta, aunque sin cifras que lo demuestren, que los resultados de la actividad quirúrgica desarrollada en los servicios universitarios son peores que los de la actividad privada. Esto obedecería a factores múltiples y bien conocidos: desnutrición y peor medio socioeconómico en los pacientes asistidos en medios universitarios, dificultades logísticas de la asistencia pública de Uruguay, gran número de intervenciones quirúrgicas llevadas a cabo por cirujanos en formación, etc. Todas estas variables contribuirían a generar esta supuesta diferencia. Hemos utilizado el POSSUM para determinar esta verdad asumida empíricamente de un modo científico y nuestros resultados no han confirmado esta hipótesis. La mortalidad esperada y la observada fueron similares entre ambas unidades, sin diferencias significativas a pesar de las ligeras diferencias entre ambos grupos, con un predominio de la cirugía mamaria y colorrectal en la unidad 2 y de la cirugía del trauma en la unidad 1 (tabla 4). Un aspecto interesante encontrado en el análisis de la mortalidad, y no mencionado en publicaciones previas, fue el hallazgo de una marcada diferencia en la puntuación POSSUM entre los pacientes que sobrevivieron y los que fallecieron después de la intervención quirúrgica. La puntuación fue inferior a 30 en los pacientes supervivientes y mayor de 40 en los fallecidos. A pesar de que POSSUM y p-POSSUM se elaboraron para auditar poblaciones y no para predecir evoluciones individuales¹³, esta diferencia podría ofrecer una orientación sobre la alta probabilidad de una evolución desfavorable en un paciente determinado.

Conclusiones

No hay antecedentes en Uruguay de la aplicación de sistemas objetivos de auditoría para evaluar y comparar los resultados de la práctica quirúrgica en diferentes circunstancias.

La utilización del sistema POSSUM con esta intención ha demostrado que es posible y razonablemente seguro dentro de un sistema asistencial muy diferente de aquél donde se creó. Por otra parte, su aplicación ha permitido a los autores mostrar que no hay diferencias significativas en los resultados entre 2 unidades diferentes de este sistema, contrariamente a lo ampliamente aceptado, pero que nunca se ha demostrado científicamente.

Agradecimientos

Los autores agradecen a Rafael Alonso, profesor adjunto de Bioestadística en la Unidad de Métodos Cuanti-

tativos de la Facultad de Medicina de Montevideo, por su invalorable colaboración en este proyecto.

Bibliografía

1. Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. *Br J Surg.* 1991;78:355-60.
2. Sagar PM, Hartley MN, Mancey-Jones B, Sedman PC, May J, MacFie J. Comparative audit of colorectal resection with the POSSUM scoring system. *Br J Surg.* 1994;81:1492-4.
3. Sagar PM, Hartley MN, MacFie J, Taylor BA, Copeland GP. Comparison of individual surgeon's performance: risk-adjusted analysis with POSSUM scoring system. *Dis Colon Rectum.* 1996;39:654-8.
4. Brunelli A, Fianchini A, Gesuita R, Carle F. POSSUM scoring system as an instrument of audit in lung resection surgery. *Ann Thorac Surg.* 1999;67:329-31.
5. Whiteley MS, Prytherch DR, Higgins B, Weaver PC, Prout WG. An evaluation of the POSSUM surgical scoring system. *Br J Surg.* 1996;83:812-5.
6. Prytherch DR, Whiteley MS, Higgins B, Weaver PC, Prout WG, Powell SJ. POSSUM and Portsmouth-POSSUM for predicting mortality. *Br J Surg.* 1998;85:1217-20.
7. Hanley JA, McNeil BJ. The meaning and use of the area under the receiver operating characteristic (ROC) curve. *Radiology.* 1982;143:29-36.
8. Hosmer DW, Lemeshow S. *Applied logistic regression.* 2nd ed. New York: John Wiley; 2000.
9. Jones HJ, De Cossart L. Risk scoring in surgical patients. *Br J Surg.* 1999;86:149-57.
10. Jones DR, Copeland GP, De Cossart L. Comparison of POSSUM with APACHE II for prediction of outcome from a surgical high-dependency unit. *Br J Surg.* 1992;79:1293-6.
11. Yip MK, Ng KJ. Risk-adjusted surgical audit with the POSSUM scoring system in a developing country. *Br J Surg.* 2002;89:110-3.
12. Bann SD, Sarin S. Comparative audit: the trouble with POSSUM. *J R Soc Med.* 2001;94:632-4.
13. Copeland GP. The POSSUM system of surgical audit. *Arch Surg.* 2002;137:15-9.