



Original

Evaluación del riesgo quirúrgico de 1.000 episodios consecutivos con el sistema POSSUM. Comparación entre cirugía gastrointestinal programada y urgente[☆]

Celia Villodre^{a,*}, Silvia Carbonell^a, Javier Espinosa^a, Jose Antonio Bravo^a, Lorea Zubiaga^a, Silvia Rojas^a, Rafael Abad^a, Luis Campos^a, Cándido Alcázar^a, Mariano Franco^a, Jose Luis Estrada^a, Pedro Zapater^b, Luis Mena^c y Félix Lluís^a

^a Servicio de Cirugía General y del Aparato Digestivo, Hospital General Universitario de Alicante, Alicante, España

^b Servicio de Farmacología Clínica, Hospital General Universitario de Alicante, Alicante, España

^c Unidad de Documentación Clínica y Admisión, Hospital General Universitario de Alicante, Alicante, España

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:

Recibido el 13 de marzo de 2011

Aceptado el 19 de junio de 2011

On-line el 3 de septiembre de 2011

Palabras clave:

Sistema POSSUM

Riesgo quirúrgico

Cirugía programada

Cirugía urgente

RESUMEN

Introducción: El objetivo del presente estudio fue evaluar la capacidad predictiva del sistema POSSUM en nuestro medio y determinar su comportamiento en cirugía gastrointestinal programada y compararla con la urgente.

Pacientes y método: Se analizaron 1.000 episodios quirúrgicos correspondientes a 909 pacientes intervenidos con anestesia general o loco-regional, de forma programada (n = 547 episodios) o urgente (n = 453), que precisaron ingreso hospitalario.

Resultados: La morbilidad total fue 31,9% (32,9% en cirugía programada y 30,7% en cirugía urgente). La capacidad discriminativa de la escala POSSUM evaluada mediante curva ROC fue mayor para la mortalidad Portsmouth (área bajo la curva [ABC] = 0,92) que para la morbilidad (ABC = 0,74). La bondad de ajuste o grado de calibración entre los valores esperados usando la escala POSSUM y los observados fue reducida para la morbilidad (Hosmer-Lemeshow [H-L] = 164,1; p < 0,05). La escala POSSUM predijo un número considerablemente mayor de muertes a las observadas, aunque la variante Portsmouth predijo mejor la mortalidad. El grado de calibración de la morbilidad fue mejor para la cirugía gastrointestinal programada (H-L = 27,7) que para la cirugía gastrointestinal urgente (H-L = 177,3). Mediante análisis de regresión logística se identificó, además del riesgo estimado mediante la propia escala POSSUM, las variables complejidad quirúrgica, tipo de cirugía y edad del paciente como factores significativos predictivos de la aparición de morbilidad y mortalidad.

Conclusiones: En nuestro medio, el sistema POSSUM predice adecuadamente el riesgo de morbilidad de la cirugía gastrointestinal programada, y sobreestima el riesgo de morbilidad de la cirugía gastrointestinal urgente.

© 2011 AEC. Publicado por Elsevier España, S.L. Todos los derechos reservados.

[☆] Presentado en el 28 Congreso Nacional de Cirugía. Madrid 8-11 de noviembre de 2010.

* Autor para correspondencia.

Correo electrónico: celivii@gmail.com (C. Villodre).

0009-739X/\$ - see front matter © 2011 AEC. Publicado por Elsevier España, S.L. Todos los derechos reservados.

doi:10.1016/j.ciresp.2011.06.004

Assessment of the surgical risk of 1,000 consecutive episodes using the POSSUM system. Comparison between elective and emergency gastrointestinal surgery

ABSTRACT

Keywords:

POSSUM system
Surgical risk
Elective surgery
Emergency surgery

Introduction: The aim of the present study was to assess the predictive capacity of the POSSUM system in a Spanish university hospital, and to determine its behaviour in elective gastrointestinal surgery and compare it with emergency gastrointestinal surgery (operation < 24 hours).

Patients and method: A total of 1,000 surgical episodes corresponding to 909 patients who required hospital admission, operated on under general or loco-regional anaesthesia, either in the elective (n= 547 episodes) or the emergency setting (n= 453), were included in the study.

Results: The overall morbidity was 31.9% (32.8% in elective surgery; 30.7% in emergency surgery). The discriminatory capacity of the POSSUM scale, evaluated using receiver operating characteristic (ROC) curves, was higher for the Portsmouth variant of mortality (Area Under the Curve [AUC] = 0,92) than for morbidity (AUC= 0,74). The goodness of fit between the expected values using the POSSUM scale and those observed was reduced for morbidity (Hosmer-Lemeshow [H-L] = 164.1; $p < 0.05$). The POSSUM scale predicted a higher number of deaths than those observed, although the Portsmouth variant was better at predicting mortality. The goodness of fit for morbidity was better for elective gastrointestinal surgery (H-L= 27.7) than emergency gastrointestinal surgery (H-L= 177.3). The logistic regression analysis identified (besides the estimated risk using the POSSUM scale itself), surgical complexity, surgery type (elective, emergency), and age of patient, as significant predictive factors of morbidity and mortality.

Conclusions: In a Spanish university hospital, the POSSUM system adequately predicts morbidity risk in elective gastrointestinal surgery, and over-estimates morbidity risk in emergency gastrointestinal surgery.

© 2011 AEC. Published by Elsevier España, S.L. All rights reserved.

La auditoría de resultados de morbimortalidad debe ser ágil e informativa y permitir la detección de desviaciones y la revisión precoz de procedimientos. Debido a la composición peculiar de cada casuística, las cifras crudas de morbimortalidad son insuficientes para auditar los propios resultados e inadecuadas para compararlos entre cirujanos, servicios u hospitales. Tanto para auditorías como para comparaciones, las cifras de morbimortalidad deben estar corregidas y ajustadas al riesgo inherente de cada casuística. El sistema POSSUM (*Physiological and Operative Severity Score for the enumeration of Mortality and morbidity*) fue desarrollado por Copeland para la estratificación de todo tipo de pacientes en base al riesgo quirúrgico^{1,2}. Su empleo solo requiere una gama limitada de mediciones de laboratorio y datos clínicos habituales en todo pre-operatorio, y los complementa con datos obtenidos en la propia intervención quirúrgica. De este modo, la fórmula para el cálculo del riesgo incorpora una puntuación fisiológica y otra operatoria.

Este sistema permite predecir el riesgo de la cirugía para una determinada población de pacientes y comparar los resultados quirúrgicos obtenidos en distintas poblaciones de pacientes, hospitales, servicios, equipos e, incluso, cirujanos individuales³. También facilita el seguimiento de las variaciones en los resultados que se obtienen a lo largo de periodos de tiempo⁴. Todo ello en base a la comparación de morbilidad y mortalidad esperada (calculada a partir de una fórmula pre-establecida) y la observada en la población a estudio.

Para emplear un sistema de predicción del riesgo quirúrgico fuera del contexto original en el que se desarrolló, es

necesario validar sus prestaciones en la nueva población diana⁵. El sistema POSSUM se diseñó originalmente para Cirugía General y posteriormente se efectuaron varias modificaciones para adaptarlo a distintas subespecialidades. Por ello, tanto la población diana como el contexto en el que se aplica deben ser examinados y, si existen diferencias, la fórmula original debe ser re-evaluada. Así, por ejemplo, la variante Portsmouth (p-POSSUM) se creó porque la fórmula original sobreestimaba el riesgo de mortalidad en Cirugía General⁶.

El objetivo del presente estudio fue evaluar la capacidad predictiva del sistema POSSUM en nuestro medio y determinar su comportamiento en la cirugía gastrointestinal programada y compararla con la cirugía gastrointestinal de urgencias, efectuadas por el mismo equipo de cirujanos, en el seno de un servicio de Cirugía General y del Aparato Digestivo de un hospital terciario español.

Pacientes y métodos

Se estudiaron de forma prospectiva los pacientes intervenidos entre enero y noviembre de 2009 en el seno del Servicio de Cirugía General y del Aparato Digestivo del Hospital General Universitario de Alicante. Se incluyeron pacientes mayores de 14 años, atendidos de forma consecutiva y que recibieron anestesia general o loco-regional. Se excluyeron pacientes intervenidos en quirófanos de tarde o que estuvieron

ingresados menos de 23 horas, así como pacientes intervenidos en la modalidad de cirugía mayor ambulatoria. Se analizaron 1.000 episodios consecutivos correspondientes a dos series de pacientes: a) intervenidos de forma programada, en quirófanos de mañana y que precisaron ingreso hospitalario; o b) intervenidos en quirófanos de Urgencias, durante las 24 horas del día, 7 días a la semana, también con ingreso hospitalario.

Se consideró cirugía urgente o semi-programada cuando se efectuó entre 2 y 24 horas posteriores al ingreso en Urgencias. Excepcionalmente, algún paciente considerado como de cirugía urgente estuvo en el área de Urgencias más de 24 horas antes de ser intervenido. Ello ocurrió en algunos pacientes con oclusión intestinal, colecistitis o apendicitis, y estuvo motivado por la incertidumbre diagnóstica inicial. No obstante, el circuito que siguieron estos pacientes fue el propio de la cirugía de urgencias. Además, también se incluyeron en esta categoría los episodios quirúrgicos correspondientes a reintervenciones de pacientes intervenidos previamente tanto mediante cirugía programada como urgente.

Se efectuó seguimiento durante los 30 días siguientes a la intervención o hasta la muerte del paciente, siempre que ocurriera durante este primer mes. Los episodios se cerraron en sesión semanal de morbimortalidad del Servicio y en consultas externas o Urgencias durante los primeros 30 días del post-operatorio.

La información se recogió de forma prospectiva de acuerdo con la actividad asistencial, y se almacenó en una base de datos en la red interna del hospital, en formato Access 2003® (Microsoft Corporation, Redmont, Washington, EE.UU.), desde donde se generaron tablas de datos, anonimizadas, para su análisis estadístico. Se registraron datos del episodio, diagnóstico e intervención principal, valores de las variables fisiológicas y operatorias descritas por Copeland², e incidentes relacionados con morbimortalidad, incluyendo además cualquier incidencia no descrita por Copeland que influyera sobre la evolución post-operatoria.

La complejidad quirúrgica de los procedimientos se adaptó a los criterios descritos². La puntuación fisiológica osciló entre 12 y 84 puntos, la operatoria entre 6 y 44 puntos. El riesgo individual de morbimortalidad se calculó con las ecuaciones descritas por Copeland, y se añadió el cálculo del riesgo de mortalidad con la variante Portsmouth descrito en estudios posteriores⁶. Morbilidad: $\text{Log}_e R1 / (1-R1) = -5,91 + (0,16 \times \text{Puntuación fisiológica}) + (0,19 \times \text{Puntuación operatoria})$. Mortalidad: $\text{Log}_e R2 / (1-R2) = -7,04 + (0,13 \times \text{Puntuación fisiológica}) + (0,16 \times \text{Puntuación operatoria})$. Mortalidad Portsmouth (p-POSSUM): $\text{Log}_e R2 / (1-R2) = -9,37 + (0,19 \times \text{Puntuación fisiológica}) + (0,15 \times \text{Puntuación operatoria})$. R1: riesgo de morbilidad. R2: riesgo de mortalidad.

Análisis estadístico

Los datos cuantitativos se expresaron como mediana, amplitud intercuartil y rango y los cualitativos como frecuencias o porcentajes. Las diferencias entre grupos en el caso de variables cuantitativas se analizaron usando el test no paramétrico de la u de Mann-Whitney y entre porcentajes o frecuencias mediante el test de la Chi-cuadrado.

Se calculó el número y el porcentaje respecto al total de la muestra de episodios quirúrgicos tras los cuales se desarrollaron complicaciones –morbilidad observada (O)– o se produjo un éxito –mortalidad observada (O)– durante los 30 días del post-operatorio. Cada episodio quirúrgico se clasificó en una de cinco categorías posibles en función de su puntuación en la escala POSSUM y el riesgo de morbilidad/mortalidad asociado: < 20, 20-40, 40-60, 60-80 y > 80%. Para cada categoría se calculó la puntuación POSSUM media y se multiplicó por el número de pacientes para obtener los valores de morbilidad y mortalidad esperadas (E). Se calcularon las razones entre los valores observados y esperados de morbilidad y mortalidad (tasa O/E). Igualmente se calcularon estas tasas tras estratificar la muestra total de episodios en dos grupos: cirugía programada y urgente, y se analizaron las diferencias entre ambos mediante el test estadístico z.

La bondad del ajuste o grado de calibración entre los valores esperados estimados usando la escala POSSUM y los realmente observados se analizó mediante la prueba chi cuadrado de Hosmer-Lemeshow (HL)^{7,8}. La capacidad discriminativa de la escala POSSUM para predecir la morbilidad y mortalidad de los episodios quirúrgicos se evaluó mediante el valor del área bajo la curva (ABC) de curvas Receiver Operating Characteristics (curvas ROC)⁹.

Se realizó un análisis de regresión logística de las variables dependientes morbilidad y mortalidad observadas en el que se incluyeron los factores complejidad quirúrgica, malignidad, tipo de cirugía, presencia de exudado y edad, además de la puntuación POSSUM como variables independientes.

Los cálculos se realizaron usando los programas Excel 2003® y SPSS® para Windows, versión 15.0 (SPSS Inc., Chicago, Illinois, EE.UU.). Un valor de $p < 0,05$ fue considerado estadísticamente significativo.

Resultados

El estudio incluyó 1.000 episodios quirúrgicos (547 pertenecientes a cirugía programada, 453 a cirugía urgente) en 909 pacientes (528 intervenidos con cirugía programada, 381 con cirugía urgente). El 92% de pacientes (n = 835) requirió un único episodio quirúrgico; 60 pacientes, dos episodios quirúrgicos; 11 pacientes, tres episodios quirúrgicos; y 3 pacientes, cuatro episodios quirúrgicos. Un porcentaje significativamente menor de pacientes intervenidos con cirugía programada requirieron más de una intervención quirúrgica comparado con los sometidos a cirugía urgente (4,2 vs 14,6%; $p = 0,001$). No hubo diferencias de sexo entre ambos grupos. Los pacientes intervenidos con cirugía urgente tuvieron una edad media significativamente menor que los intervenidos con cirugía programada (tabla 1). De acuerdo con las estimaciones de la escala POSSUM, los episodios de cirugía urgente se asociaron con un riesgo esperado de morbilidad y mortalidad significativamente mayor que los de cirugía programada (tabla 1). El porcentaje de pacientes que murió en el mes posterior a la cirugía fue significativamente mayor en el grupo de pacientes sometido a cirugía urgente aunque en ambos grupos de pacientes la mortalidad observada fue menor que la esperada por la escala POSSUM y la modificación Portsmouth (tabla 1). Finalmente, el peso medio de los GRD del

Tabla 1 – Características de los episodios quirúrgicos analizados

Características	Total	Cirugía programada	Cirugía urgente	p
Número de episodios	1.000	547	453	
Sexo (hombres/mujeres) (%)	56/44	54/46	59/41	0,188
Edad (mediana [AIQ]) (rango)	58 (30) (14 – 99)	61 (24) (16 – 93)	54 (38) (14 – 99)	0,000
Riesgo esperado de morbilidad según escala POSSUM (mediana [AIQ]) (rango)	0,23 (0,41) (0,05-1,0)	0,19 (0,30) (0,05 – 0,98)	0,33 (0,58) (0,11 – 1,0)	0,000
Riesgo esperado de mortalidad según escala POSSUM (mediana [AIQ]) (rango)	0,04 (0,09) (0,01 – 0,99)	0,03 (0,06) (0,01 – 0,76)	0,06 (0,20) (0,02 – 0,99)	0,000
Riesgo esperado de mortalidad según la variante Portsmouth (mediana [AIQ]) (rango)	0,01 (0,03) (0,0 – 0,99)	0,008 (0,02) (0,0 – 0,57)	0,015 (0,08) (0,0 – 0,99)	0,000
Mortalidad esperada según la variante Portsmouth N (%)	63 (6%)	14 (2,6%)	49 (10,8%)	0,001
Mortalidad observada N (%)	28 (2,8%)	7 (1,3%)	21 (4,6%)	0,003
Peso GRD (mediana [AIQ]) (rango)	1,4 (2,2) (0,4 – 51)	1,5 (1,3) (0,41 – 51,3)	1,3 (3,3) (0,39-51,4)	0,117
Tumores N (%)	219 (21,9%)	179 (32,7%)	40 (8,8%)	0,001
Presencia de líquido seroso, sangre o pus en cavidad abdominal N (%)	339 (33,9%)	49 (9,0%)	290 (64,0%)	0,001

AIQ = amplitud intercuartil (percentil 75% - percentil 25%); Rango = valor máximo – valor mínimo. El peso GRD (Grupos Relacionados con el Diagnóstico) se refiere al episodio asistencial completo en el que se incluye el episodio quirúrgico evaluado.

episodio asistencial en el que se enmarcaba el evento quirúrgico no fue significativamente diferente entre los dos grupos de pacientes estudiados (tabla 1). La proporción de tumores fue significativamente superior en la cirugía programada que en la urgente, mientras que el porcentaje de pacientes que presentaron líquido seroso, sangre o pus en la cavidad abdominal fue mayor en la cirugía urgente (tabla 1).

Las morbilidades observadas en cada grupo se recogen en la tabla 2. Se observó un total de 515 morbilidades correspondientes a un total de 319 episodios, 180 en cirugía programada y 139 en cirugía urgente. La morbilidad total fue 31,9% (32,9% en cirugía programada y 30,7% en cirugía urgente). Se observaron diferencias significativas en la frecuencia de aparición de algunas morbilidades entre el grupo de cirugía programada y urgente. Concretamente, en el grupo de cirugía urgente se apreció un número significativamente mayor de dehiscencias de heridas, infecciones

profundas y pulmonares, cuadros de insuficiencia respiratoria y septicemias. Por el contrario, el número de episodios de hemorragia profunda fue significativamente mayor entre los pacientes sometidos a cirugía programada (tabla 2). El análisis de resultados de morbilidad no cambió de forma relevante al considerar cualquier otra incidencia relacionada con la morbilidad además de las descritas por Copeland.

Al considerar el total de episodios quirúrgicos analizados, se observó que la capacidad discriminativa de la escala POSSUM evaluada mediante el cálculo del área bajo la curva ROC fue mayor para la mortalidad Portsmouth (ABC = 0,92) que para la morbilidad (ABC = 0,74) (fig. 1).

Cuando se estratifica el total de episodios quirúrgicos en quintiles de probabilidad se aprecia que la bondad del ajuste o grado de calibración entre los valores esperados usando la escala POSSUM y los finalmente observados es reducida para la morbilidad (Hosmer-Lemeshow = 164,1; $p < 0,05$) (tabla 3)

Tabla 2 – Distribución de la morbilidad observada en episodios con cirugía programada y cirugía urgente

	Programada (n)	Urgente (n)	p
Dehiscencia anastomótica	36	22	0,245
Dehiscencia de herida	13	22	0,034 ^a
Fallo cardíaco	3	6	0,196
Fiebre de origen desconocido	9	7	0,900
Hemorragia de herida	10	3	0,105
Hemorragia profunda	17	5	0,049 ^a
Hipotensión	9	11	0,379
Infección de herida	52	48	0,567
Infección profunda	17	29	0,015 ^a
Infección pulmonar	7	15	0,032 ^a
Infección urinaria	4	4	0,789
Insuficiencia renal	4	9	0,081
Insuficiencia respiratoria	4	24	0,001 ^a
Septicemia	6	17	0,006 ^a
Trombosis venosa profunda y TEP	1	3	0,232
Otras	57	41	0,468
Total	249	266	
Episodios	180 (32,9%)	139 (30,7%)	

^a Diferencia estadísticamente significativa.

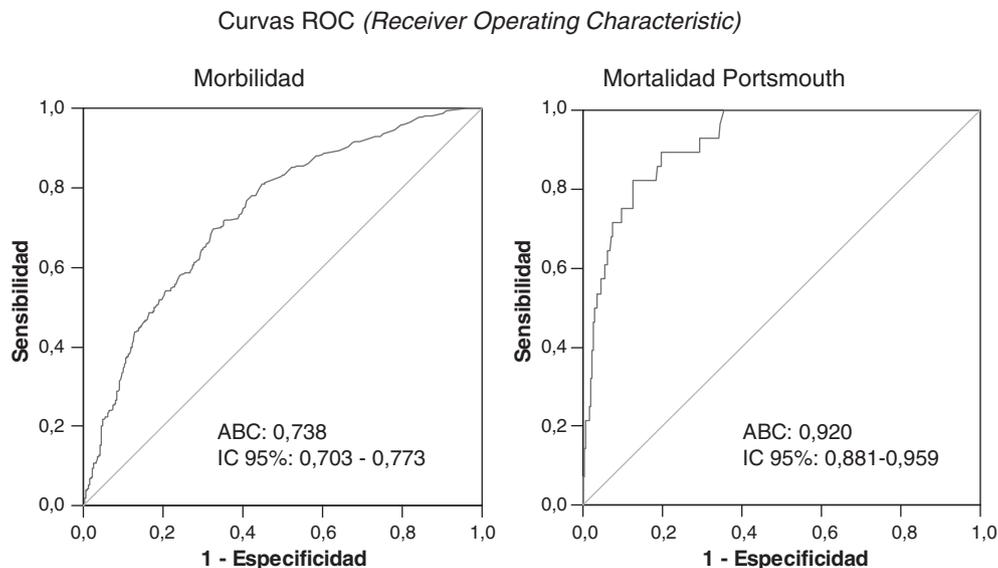


Figura 1 – Curvas ROC correspondientes a la morbilidad y mortalidad (con el empleo de la variante Portsmouth) de los 1.000 episodios quirúrgicos estudiados. ABC: área bajo la curva; IC: intervalo de confianza asintótico al 95%.

indicando la existencia de una diferencia estadísticamente significativa entre los valores esperados y los observados (fig. 2). El número total de episodios que desarrollaron morbilidades fue menor que el previsto por la escala POSSUM aunque es en las categorías de riesgo superior al 40% en las que se apreció una reducción en los valores observados frente a los esperados (tabla 3).

En nuestra muestra se observa que la mayoría de los episodios quedan incluidos en una categoría de riesgo de mortalidad inferior al 20%. La escala POSSUM predijo un número considerablemente mayor de muertes a las observadas (tabla 4). La variante Portsmouth predijo el doble de muertes de las que finalmente ocurrieron, una relación que se mantuvo constante para todas las categorías de riesgo consideradas (tabla 5), con una diferencia estadísticamente significativa entre los valores observados y esperados. A pesar

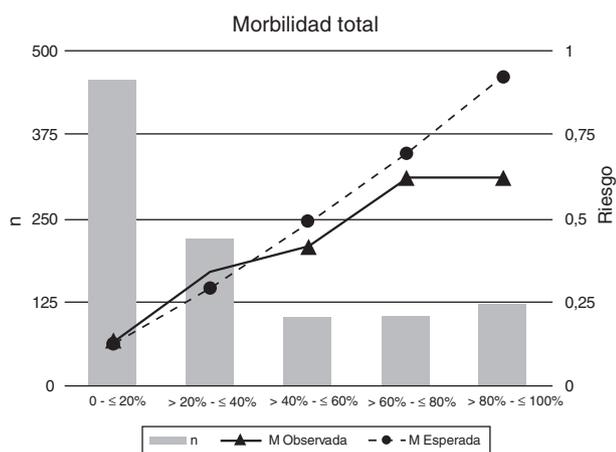


Figura 2 – Morbilidad esperada y observada cuando se estratifican el total de episodios quirúrgicos en cinco quintiles posibles de probabilidad en función de su puntuación en la escala POSSUM.

de ello, se puede considerar a la variante Portsmouth como una herramienta mejor para predecir la mortalidad, como se muestra por un estadístico de Hosmer-Lemeshow de 66,9 (tabla 5) menor al mostrado por la escala POSSUM (Hosmer-Lemeshow = 167,2) (tabla 4).

El riesgo esperado de morbilidad estimado mediante la escala POSSUM se distribuyó de forma diferente en los grupos de cirugía programada y urgente (figs. 3 y 4). Esta diferente distribución fue estadísticamente significativa (Chi cuadrado: 91,1; $p = 0,001$) destacando que tan solo un 13% de los episodios de cirugía programada tenían un riesgo, según la escala POSSUM, superior al 60% de desarrollar morbilidad (fig. 3; tabla 3) frente a un 34% de los episodios de cirugía urgente (fig. 4; tabla 3). En el análisis de los componentes de la escala se observó una puntuación media fisiológica y operatoria superiores para los episodios de cirugía urgente (20,18 y 14,22, respectivamente) que en la cirugía programada (17,20 y 10,27, respectivamente).

En las tablas 3 y 4 y en las figuras 3 y 4, se describen los valores observados y esperados a partir de la escala POSSUM de morbilidad y mortalidad y el análisis del grado de ajuste entre los mismos para el total de la muestra analizada y para los episodios quirúrgicos programados y urgentes. En dicho análisis, realizado con la prueba de Hosmer-Lemeshow (tabla 6), se observa que únicamente en los casos de cirugía programada el número de muertes esperadas por la variante Portsmouth no difiere de forma significativa del número de muertes observadas (Hosmer-Lemeshow = 5,21; $p > 0,05$).

Mediante análisis de regresión logística se identificó, además del riesgo estimado mediante la propia escala POSSUM, las variables complejidad quirúrgica ($p = 0,001$), tipo de cirugía ($p = 0,004$) y edad del paciente ($p = 0,001$) como factores significativos predictivos de la aparición de morbilidad. Para la mortalidad, las mismas variables alcanzaron significación estadística: complejidad quirúrgica ($p = 0,001$), tipo de cirugía ($p = 0,024$) y edad del paciente ($p = 0,011$).

Tabla 3 – Análisis comparativo de la morbilidad observada y esperada estratificadas según el riesgo estimado mediante la escala POSSUM

Riesgo esperado	Episodios (n)			Morbilidad observada (n)			Morbilidad esperada (n)			Tasa O/E		
	Total	Cirugía programada	Cirugía urgente	Total	Cirugía programada	Cirugía urgente	Total	Cirugía programada	Cirugía urgente	Total	Cirugía programada	Cirugía urgente
0 - ≤ 20%	456	293	163	62	45	17	57	32	25	1,08	1,41	0,68
> 20 - ≤ 40%	218	124	94	75	56	19	64	36	27	1,17	1,56	0,70
> 40 - ≤ 60%	101	59	42	42	28	14	50	29	21	0,84	0,97	0,67
> 60 - ≤ 80%	103	51	52	64	36	28	72	35	37	0,89	1,03	0,76
> 80 - ≤ 100%	122	20	102	76	15	61	113	18	94	0,67	0,83	0,65
Total	1.000	547	453	319	180	139	356	150	204	0,90	1,20	0,68

Tasa O/E: relación de episodios con complicaciones observadas frente a episodios con complicaciones esperadas.

Para el total de episodios: Hosmer-Lemeshow: 164,1; $p < 0,05$.

Tabla 4 – Análisis comparativo de la mortalidad observada y esperada estratificadas según el riesgo estimado mediante la escala POSSUM

Riesgo esperado	Episodios (n)			Mortalidad observada (n)			Mortalidad esperada (n)			Tasa O/E		
	Total	Cirugía programada	Cirugía urgente	Total	Cirugía programada	Cirugía urgente	Total	Cirugía programada	Cirugía urgente	Total	Cirugía programada	Cirugía urgente
0 - ≤ 20%	842	511	331	5	5	0	43	24	19	0,12	0,21	0,00
> 20 - ≤ 40%	77	22	55	5	1	4	22	6	16	0,23	0,17	0,25
> 40 - ≤ 60%	32	10	22	4	0	4	16	5	11	0,25	0,00	0,36
> 60 - ≤ 80%	25	4	21	5	1	4	17	3	14	0,29	0,33	0,29
> 80 - ≤ 100%	24	0	24	9	0	9	21	0	21	0,43	0,00	0,43
Total	1.000	547	453	28	7	21	119	38	81	0,24	0,19	0,26

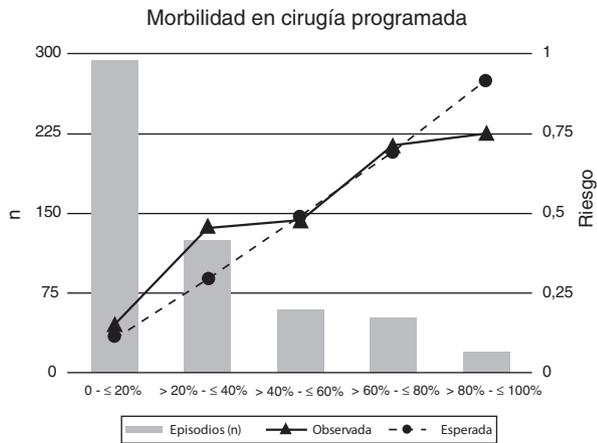
Tasa O/E: relación de muertes observadas frente a muertes esperadas.

Para el total de episodios: Hosmer-Lemeshow: 167,2; $p < 0,05$.

Tabla 5 – Análisis comparativo de la mortalidad observada y esperada estratificadas según el riesgo estimado mediante la variante Portsmouth

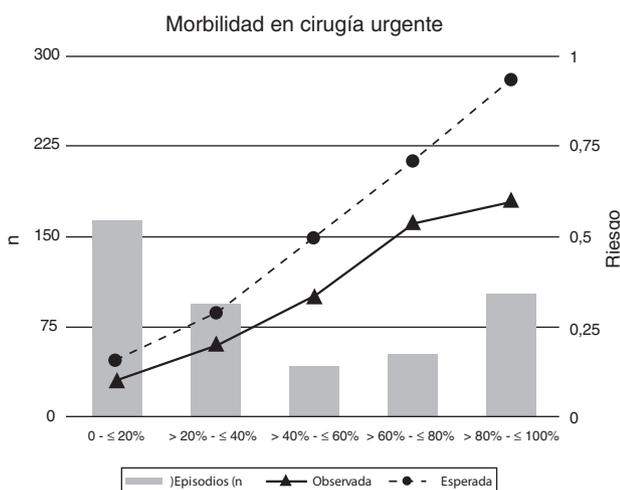
Riesgo esperado	Episodios (n)	Mortalidad observada (n)	Mortalidad esperada (n)	Tasa O/E
0 - ≤ 20%	921	10	22	0,45
> 20 - ≤ 40%	33	4	10	0,40
> 40 - ≤ 60%	19	6	10	0,60
> 60 - ≤ 80%	15	2	11	0,18
> 80 - ≤ 100%	12	6	11	0,55
Total	1.000	28	64	0,44

Tasa O/E: relación de muertes observadas frente a muertes esperadas. Hosmer-Lemeshow: 66,9; p < 0,05.

**Figura 3 – Morbilidad esperada y observada cuando se estratifican los episodios quirúrgicos de cirugía programada en cinco quintiles posibles de probabilidad en función de su puntuación en la escala POSSUM.**

Discusión

En nuestro estudio, el sistema POSSUM estima adecuadamente el riesgo de morbilidad de la cirugía gastrointestinal

**Figura 4 – Morbilidad esperada y observada cuando se estratifican los episodios quirúrgicos de cirugía urgente en cinco quintiles posibles de probabilidad en función de su puntuación en la escala POSSUM.**

programada en pacientes que precisan ingreso hospitalario superior a 24 horas. Sin embargo, sobreestima el riesgo de morbilidad de la cirugía gastrointestinal de urgencias. Los mejores resultados observados en la cirugía gastrointestinal de urgencia se han mantenido constantes a lo largo del periodo de estudio.

La cohorte de pacientes en los que se ha efectuado el presente estudio es, probablemente, una muestra representativa que podría extrapolarse a cualquier otro hospital de nuestro país. La naturaleza del estudio nos ha permitido incluir una amplia gama de diagnósticos y procedimientos habituales en la práctica clínica de Cirugía General y del Aparato Digestivo. La presión de Urgencias de los pacientes del Servicio oscila alrededor de 25%, inferior al 45-50% comunicado en algunos centros. Este valor indica que solo se operan de urgencia los pacientes que lo requieren, y no aquellos que no han podido ser intervenidos de forma programada. Un sesgo del presente estudio podría ser la escasa inclusión de procedimientos programados de bajo riesgo. Posiblemente a ello se deba la mejor coincidencia de la escala POSSUM con nuestros datos de cirugía programada.

Hay varias posibles razones que explicarían nuestros datos. La cirugía programada que requiere ingreso hospitalario superior a 24 horas se efectúa habitualmente en pacientes con elevada co-morbilidad o cuando la enfermedad de base es compleja. Muchos de estos pacientes son portadores de un tumor maligno con distintos grados de progresión (ejemplo: con metástasis ganglionares o a distancia). Asimismo, la complejidad de las intervenciones electivas es superior a la de las intervenciones urgentes, otro factor tenido en cuenta por el sistema POSSUM. La cirugía de urgencia se centra en la resolución de problemas agudos, con un manifiesto componente inflamatorio, y donde la presencia de líquido seroso, contaminación intestinal, sangre o pus en la cavidad abdominal serían circunstancias óptimas para la acción del

Tabla 6 – Bondad del ajuste o grado de calibración^a entre los valores esperados de morbilidad estimados usando la escala POSSUM y los finalmente observados en cirugía programada y urgente

	Cirugía programada	Cirugía urgente
χ^2 mortalidad Portsmouth (p-POSSUM)	5,21 (p > 0,05)	64,43 (p < 0,05)
χ^2 morbilidad	27,7 (p < 0,05)	177,3 (p < 0,05)

^a Análisis efectuado con la prueba de Hosmer-Lemeshow.

cirujano y por ende susceptibles de mejorar el estado del paciente.

Varios estudios previos han detectado que la predicción de mortalidad POSSUM sobreestima el riesgo, en especial en los pacientes estratificados de menor riesgo. Como medida correctora se aconseja el empleo de la variante Portsmouth (p-POSSUM) ya que se ajusta mejor a los resultados reales⁶. Nuestros resultados avalan que la variante Portsmouth estima adecuadamente la mortalidad en cirugía gastrointestinal programada y de urgencia.

La calibración es la capacidad del modelo para asignar las probabilidades correctas de resultado final a pacientes individuales. En el presente trabajo, la comparación entre resultados esperados y observados mediante la prueba de Hosmer-Lemeshaw mostró una correcta capacidad de calibración para la morbilidad en pacientes intervenidos con cirugía programada. Sin embargo, la calibración no fue correcta para los pacientes intervenidos con cirugía urgente/semiprogramada, donde los resultados observados fueron mejores que los esperados. Por otro lado, la potencia discriminativa del modelo –es decir, la capacidad de asignar mayor riesgo a los pacientes que presentan complicaciones o fallecen que a aquellos que no las presentan o sobreviven– se analizó con la ayuda de curvas ROC. En el presente estudio, el área bajo la curva para la mortalidad Portsmouth fue 0,92, lo que significa buena capacidad discriminativa. Sin embargo, para la morbilidad fue 0,74, lo que implica una peor capacidad discriminativa. La suma de ambas consideraciones sugiere que la escala POSSUM es poco fiable para predecir el riesgo de complicaciones para pacientes individuales, aunque puede emplearse para calcular tasas globales de morbilidad esperada en determinadas poblaciones. Y que la escala es mejor herramienta de predicción para la mortalidad que para la morbilidad.

Los índices brutos de morbimortalidad quirúrgica son poco informativos en general ya que no contemplan las características de la población tratada ni la complejidad de los procedimientos empleados. Para paliar esta limitación, se recurre a menudo al denominado *benchmarking*, una técnica de comparación de los resultados obtenidos por un grupo amplio de servicios pertenecientes a distintas instituciones hospitalarias. Sin embargo, el *benchmarking* también adolece de limitaciones, muchas de ellas derivadas de la diversidad en la recogida de datos. Tan es así, que recientemente se aboga por la comparación evolutiva de resultados entre equipos similares. Para esa finalidad, sin duda, las escalas de riesgo ofrecen una herramienta muy valiosa¹⁰. A efectos prácticos, las escalas de riesgo pueden ayudar en la comparación de resultados entre sistemas de salud, o en la evaluación de la calidad de unidades individuales. Hasta ahora no se disponía de datos ajustados por riesgo y referidos exclusivamente a la cirugía gastrointestinal de urgencia, una actividad clínica de gran impacto en nuestro país y, posiblemente también, en el mundo occidental.

También parece necesario adaptar las escalas de predicción de riesgo a cada especialidad quirúrgica o área específica de actividad¹¹⁻¹⁶. Los resultados del presente trabajo avalan la necesidad de desarrollar una fórmula específica para cirugía gastrointestinal urgente que sea validada prospectivamente en el mismo escenario y posteriormente revalidada en otros

escenarios externos. En este sentido, las variables complejidad quirúrgica, tipo de cirugía y edad se han mostrado como factores predictores de la morbilidad y mortalidad independientes de la valoración POSSUM en el análisis de regresión logística.

En conclusión, la escala POSSUM es un instrumento útil para efectuar un seguimiento de la morbimortalidad, permite comparar los resultados propios con los obtenidos en otros centros y facilita el control de calidad de la actividad asistencial. Sin embargo, es necesario contar con datos procedentes del medio específico (tipo de hospital, servicio y especialidad) en el que se va a aplicar. En este sentido, el presente estudio ha detectado que la escala POSSUM sobreestima el riesgo de morbilidad de la cirugía gastrointestinal de urgencia y sugiere que podría ser útil confeccionar una nueva ecuación mejor adaptada a estas necesidades.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

BIBLIOGRAFÍA

- Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. *Br J Surg*. 1991;78:355-60.
- Copeland GP. The POSSUM system of surgical audit. *Arch Surg*. 2002;137:15-9.
- Machado F, Barberousse C, Santiago P, Barrios E, Carriquiry L. Comparación de resultados de la cirugía en dos poblaciones ajustadas por riesgo mediante el sistema POSSUM. *Cir Esp*. 2007;81:31-7.
- Luna A, Rebas P, Navarro S, Montmany S, Coroleu D, Cabrol J, et al. An evaluation of morbidity and mortality in oncologic gastric surgery with the application of POSSUM P-POSSUM, and O-POSSUM. *World J Surg*. 2009;33:1889-94.
- Horzic M, Kopljarić M, Cupurdija K, Bielen DV, Vergles D, Lackovic Z. Comparison of P-POSSUM and Cr-POSSUM scores in patients undergoing colorectal cancer resection. *Arch Surg*. 2007;142:1043-8.
- 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.
- Lemeshow S, Hosmer Jr DW. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol*. 1982;115:92-106.
- Hosmer DW, Lemeshow S. *Applied logistic regression*, 2nd ed. New York: John Wiley & Sons, Inc. 2000.
- Hanley JA, McNeil BJ. The meaning and use of the area under the receiver operating characteristic (ROC) curve. *Radiology*. 1982;143:29-36.
- Vollmer CM, Pratt W, Vanounou T, Maithel SK, Callery MP. Quality assessment in high-acuity surgery: volume and mortality are not enough. *Arch Surg*. 2007;142:371-80.
- Campillo-Soto A, Flores-Pastor B, Soria-Aledo V, Candel-Arenas M, Andrés-García B, Martín-Lorenzo JG, et al. Sistema POSSUM. Un instrumento de medida de la calidad en el paciente quirúrgico. *Cir Esp*. 2006;80:395-9.
- Oomen JLT, Cuesta MA, Engel AF. Comparison of outcome of POSSUM, p-POSSUM, and cr-POSSUM scoring after elective resection of the sigmoid colon for carcinoma or

- complicated diverticular disease. *Scand J Gastroenterol.* 2007;42:841-7.
13. Pratt W, Joseph S, Callery MP, Vollmer Jr CM. POSSUM accurately predicts morbidity for pancreatic resection. *Surgery.* 2008;143:8-19.
 14. De Castro SMM, Houwert JT, Lagarde SM, Reitsma JB, Busch ORC, van Gulik TM, et al. Evaluation of POSSUM for patients undergoing pancreatoduodenectomy. *World J Surg.* 2009;33:1481-7.
 15. Leung E, Ferjani AM, Stellard N, Wong LS. Predicting postoperative mortality in patients undergoing colorectal surgery using P-POSSUM and Cr-POSSUM scores: a prospective study. *Int J Colorectal Dis.* 2009;24:1459-64.
 16. Valenti V, Hernández-Lizoain JL, Baixauli J, Pastor C, Martínez-Regueira F, Beunza JJ, et al. Analysis of POSSUM score and postoperative morbidity in patients with rectal cancer undergoing surgery. *Langenbecks Arch Surg.* 2009;394:55-63.