

Validación prospectiva multicéntrica del EuroSCORE II en Argentina

Multicenter prospective validation of the EuroSCORE II in Argentina

Raúl A. Borracci^{1,2*}, Miguel Rubio¹, Julio Baldi-Jr¹, Julio C. Giorgini² y Claudio C. Higa²

¹Servicio de Cirugía Cardíaca, Hospital de Clínicas, Universidad de Buenos Aires; ²Servicio de Cardiología y Cirugía Cardiovascular, Deutsches Hospital, Buenos Aires, Argentina

Resumen

Objetivo: Validar, en forma prospectiva y en múltiples centros, la precisión y utilidad clínica del European System for Cardiac Operative Risk Evaluation II (EuroSCORE II) para predecir la mortalidad operatoria de la cirugía cardíaca en centros de Argentina. **Método:** Entre enero de 2012 y febrero de 2018 se incluyeron en forma prospectiva 2,000 pacientes consecutivos que fueron sometidos a cirugía cardíaca en diferentes centros de Argentina. El punto final fue mortalidad hospitalaria por cualquier causa. La discriminación, calibración, precisión y utilidad clínica del EuroSCORE II se evaluaron en la cohorte global y en los diferentes tipos de cirugías, basándose en las curvas Receiver Operating Characteristics (ROC), bondad de ajuste de Hosmer-Lemeshow, razón de mortalidad observada/esperada, índice de Shannon y curvas de decisión. **Resultados:** El área ROC del EuroSCORE II estuvo entre 0.73 y 0.80 para todo tipo de cirugía, y el valor más bajo fue para la cirugía coronaria. La mortalidad observada y esperada fue 4.3 y 3.0%, respectivamente ($p = 0.034$). El análisis de la curva de decisión demostró un beneficio neto positivo para los umbrales por debajo de 0.24 para todo tipo de cirugía. **Conclusiones:** El EuroSCORE II tuvo un desempeño adecuado en términos de discriminación y calibración para todos los tipos de cirugía, aunque algo inferior para la cirugía coronaria. Si bien en términos generales subestimó el riesgo en los grupos de riesgo intermedio, el comportamiento global fue aceptable. El EuroSCORE II podría considerarse una opción de modelo genérico y actualizado de estratificación del riesgo operatorio para predecir la mortalidad hospitalaria de la cirugía cardíaca en nuestro contexto.

Palabras clave: European System for Cardiac Operative Risk Evaluation II. Validación. Riesgo. Análisis de curvas de decisión. Procedimientos quirúrgicos cardíacos. Argentina.

Abstract

Objective: To validate prospectively in multiple centers, the accuracy and clinical utility of the European System for Cardiac Operative Risk Evaluation (EuroSCORE II) to predict the operative mortality of cardiac surgery in Argentina. **Methods:** Between January 2012 and February 2018, 2,000 consecutive adult patients who underwent cardiac surgery in different centers in Argentina were prospectively included. The end-point was in-hospital all-cause mortality. Discrimination, calibration, precision and

Correspondencia:

*Raúl A. Borracci

La Pampa, 3030

C.P. 1428, Buenos Aires, Argentina

E-mail: raborracci@gmail.com

Fecha de recepción: 06-02-2019

Fecha de aceptación: 14-03-2019

DOI: 10.24875/ACM.19000059

Disponible en internet: 06-05-2019

Arch Cardiol Mex. 2019;89(4):315-323

www.archivoscardiologia.com

1405-9940/© 2019 Instituto Nacional de Cardiología Ignacio Chávez. Publicado por Permanyer. Este es un artículo open access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

clinical utility of the EuroSCORE II were evaluated in the global cohort and in the different types of surgeries, based on ROC (Receiver Operating Characteristics) curves, Hosmer-Lemeshow goodness-of-fit test, observed/expected mortality ratio, Shannon index and decision curves analysis. **Results:** ROC area of the EuroSCORE II was between 0.73 and 0.80 for all types of surgery, being the lowest value for coronary surgery. The observed and expected mortality was 4.3% and 3.0%, respectively ($p = 0.034$). The decision curve analysis showed a positive net benefit for all thresholds below 0.24, considering all type of surgeries. **Conclusion:** The EuroSCORE II showed an adequate performance in terms of discrimination and calibration for all types of surgery, although somewhat inferior for coronary surgery. Though in general terms this model underestimated the risk in intermediate risk groups, its overall performance was acceptable. The EuroSCORE II could be considered an optional updated generic model of operative risk stratification to predict in-hospital mortality after cardiac surgery in our context.

Key words: European System for Cardiac Operative Risk Evaluation II. Validation. Risk. Decision curve analysis. Cardiac surgical procedures. Argentina.

Introducción

Los modelos de estratificación por riesgo en cirugía cardíaca cumplen dos funciones principales. Por un lado, colaboran en la toma de decisiones terapéuticas de acuerdo al riesgo preoperatorio de mortalidad esperada con la cirugía; y por otro, ayudan en el control de calidad de los resultados quirúrgicos de un servicio¹. Aunque existen múltiples modelos de puntuación por riesgo que pueden desarrollarse con datos locales, el uso de un *score* validado mundialmente permite comparar los resultados propios con un estándar internacional común.

La implementación efectiva de un *score* internacional requiere validar su precisión en términos de discriminación y calibración en la muestra local en la que va a ser aplicada. El European System for Cardiac Operative Risk Evaluation (EuroSCORE II) es un sistema de puntuación que se ha validado varias veces en otros países y que reemplazó y ajustó a las antiguas versiones aditiva y logística del EuroSCORE original². Tres metaanálisis demostraron que, a pesar de la heterogeneidad de los estudios, el EuroSCORE II muestra un buen desempeño global para predecir la mortalidad operatoria; aunque serían necesarias nuevas validaciones prospectivas en otras poblaciones de pacientes³⁻⁵.

También se han señalado algunas inconsistencias de este modelo. Dado que existiría una variación estacional anual de la mortalidad asociada a la cirugía cardíaca, el reclutamiento de datos del EuroSCORE II solo en la primavera-verano del hemisferio norte podría haber introducido un sesgo en la precisión del modelo⁶. Otros autores han criticado la calidad de la calibración del EuroSCORE II, basada solamente en el estadístico de Hosmer-Lemeshow^{7,8}; así como también, algunos aspectos del diseño en cuanto al desarrollo y la validación del modelo⁹. Otros estudios señalaron que este

modelo podría subestimar el riesgo de mortalidad en pacientes de alto riesgo y en aquellos sometidos a cirugía combinada¹⁰. También el EuroSCORE II parece subestimar la mortalidad de la cirugía de la endocarditis¹¹ y, en algunas poblaciones europeas en particular, no estima mejor la mortalidad que las versiones anteriores¹². Sin embargo, muchas de estas críticas fueron realizadas poco tiempo después de presentado el modelo, y suavizadas en los últimos años a medida que aparecieron las validaciones prospectivas externas.

El objetivo de este estudio fue validar en forma prospectiva y en múltiples centros, la precisión y utilidad clínica del EuroSCORE II para predecir la mortalidad operatoria de la cirugía cardíaca en centros de Argentina.

Métodos

Este estudio incluyó una serie prospectiva y consecutiva de 2,000 pacientes adultos que fueron sometidos a cirugía cardíaca en el Hospital de Clínicas de la Universidad de Buenos Aires y sus hospitales y clínicas asociadas, entre enero de 2012 y febrero de 2018. Se incluyeron todos los tipos de cirugía, excepto disección aórtica aguda, trasplante e implante valvular aórtico transcatheter. Los datos clínicos y de laboratorio basales y operatorios se recogieron en forma prospectiva en una base de datos *ad hoc*, que incluía las variables necesarias para estimar el riesgo de mortalidad quirúrgica basándose en el EuroSCORE II. Se adoptaron todas las definiciones propuestas por dicho modelo de estratificación del riesgo, y el puntaje para cada paciente se determinó con una calculadora interactiva en línea (<http://www.euroscore.org/calc.html>). En la base de datos computarizada se agregó información adicional sobre otros factores de riesgo, comorbilidades, complicaciones operatorias mayores y mortalidad hospitalaria. El punto final evaluado fue la mortalidad

hospitalaria por cualquier causa. La discriminación y precisión del EuroSCORE II se evaluaron en la cohorte global y en los diferentes tipos de cirugías realizadas.

Análisis estadístico

Las variables continuas se expresaron como media \pm desviación estándar (DE) y la distribución de frecuencias de los puntajes se complementó con el rango intercuartílico (RIC). Se utilizó la prueba de bondad de ajuste de Kolmogorov-Smirnov para analizar la normalidad de las distribuciones. El cociente de la mortalidad observada sobre la esperada se comparó mediante la prueba χ^2 o la prueba de probabilidad exacta de Fisher para dos colas, y se representó en un gráfico de validación clínica del modelo, separado por grupos de riesgo¹³. La calibración del modelo se evaluó con la prueba de bondad de ajuste de Hosmer-Lemeshow (HL), y el área bajo la curva ROC (*receiver operating characteristics*), con su intervalo de confianza (IC) del 95%, se usó para estimar la capacidad de discriminación del EuroSCORE II en la predicción de la mortalidad hospitalaria. La precisión del EuroSCORE II también se determinó basándose en el modelo de la teoría de la información, teniendo en cuenta la estimación del riesgo y la presencia o no del suceso muerte (índice de Shannon)¹⁴. Finalmente, se calculó el beneficio neto del EuroSCORE II para predecir la mortalidad hospitalaria, basándose en el análisis de la curva de decisión¹⁵. La construcción de la curva de decisión se realizó con una planilla de cálculo en Microsoft Excel®, y el resto del análisis estadístico se realizó con SPSS Statistics para Windows, Versión 17.0 (SPSS, Inc., Chicago, IL, USA).

El estudio se realizó siguiendo las recomendaciones para estudios de investigación en seres humanos y las normas legales vigentes. Se implementaron medidas para proteger la confidencialidad de toda la información de acuerdo con la ley Argentina 25,326 de protección de datos personales. El protocolo del estudio fue aprobado por el Comité Revisor de cada institución participante.

Resultados

En la [tabla 1](#) se resumen las características basales de la población estudiada. La distribución de los valores del EuroSCORE II para toda la población de pacientes se muestra en la [figura 1 A](#), donde se observa una marcada asimetría a la derecha, con una mediana de 1.44% (RIC: 0.85-2.72%). Por su parte, la

Tabla 1. Características basales de la población

Variables	N (%)
<i>Preoperatorias e intraoperatorias</i>	
Edad (media \pm DE) (rango)	66.5 \pm 10.7 (20-92) años
Sexo masculino	1429 (71.5)
Diabetes dependiente de la insulina	52 (2.6)
Diabetes no dependiente de la insulina	386 (19.3)
Insuficiencia cardíaca	95 (4.8)
Accidente cerebrovascular	70 (3.5)
Enfermedad pulmonar obstructiva crónica	112 (5.6)
Arteriopatía periférica	46 (2.3)
Díálisis	16 (0.8)
Angina inestable*†	17 (1.7)
Infarto de miocardio reciente (< 60 días)†	47 (4.7)
Endocarditis activa	57 (2.9)
Cirugía cardíaca previa	29 (1.5)
Disfunción moderada/severa del VI	620 (31.0)
<i>Tipo de cirugía</i>	
Coronario	999 (50.0)
Valvular aórtico	415 (20.7)
Valvular mitral	120 (6.0)
Combinado	206 (10.3)
Miscelánea‡	260 (13.0)
Cirugía de urgencia/emergencia	298 (14.9)
Cirugía coronaria sin bomba†	283 (28.3)
Uso de al menos una arteria mamaria†	954 (95.5)
Tiempo de CEC (media \pm DE)	64.6 \pm 16.0 min
<i>Postoperatorias</i>	
Mortalidad hospitalaria	85 (4.3)
Extubación en quirófano (<i>ultra-fasttrack</i>)	1362 (68.1)
<i>Complicaciones mayores</i>	
Reoperación por sangrado	31 (1.6)
Infarto tipo Q†	18 (1.8)
Bajo gasto cardíaco	51 (2.6)
Ventilación mecánica prolongada (> 48 h)	34 (1.7)
Accidente cerebrovascular	16 (0.8)
Díálisis <i>de novo</i>	21 (1.1)
Infección esternal profunda	24 (1.2)

*Definida como necesidad de nitratos intravenosos al momento de la cirugía.

†Calculado exclusivamente para la cirugía coronaria (n = 999).

‡Miscelánea incluye neoplasias cardíacas, aneurisma ventricular y de aorta ascendente, complicaciones mecánicas del infarto y congénitas del adulto.

CEC: circulación extracorpórea; DE: desviación estándar; VI: ventrículo izquierdo.

curva ROC de la [figura 1 B](#) demuestra un área de 0.80, con una buena discriminación global del modelo para toda la cohorte, así como también, una adecuada calibración (HL $p = 0.178$). El desempeño del EuroSCORE II para cada tipo de cirugía en términos de capacidad predictiva se muestra en la [figura 2](#). La posición de cada burbuja en el gráfico depende del balance entre la calibración del modelo y el poder discriminatorio evaluado con el área ROC. Se observa que el EuroSCORE II presenta una calibración adecuada para cualquier tipo de cirugía con todos los valores de HL $\chi^2 < 15.0$, y una discriminación que varía entre 0.73 y 0.82.

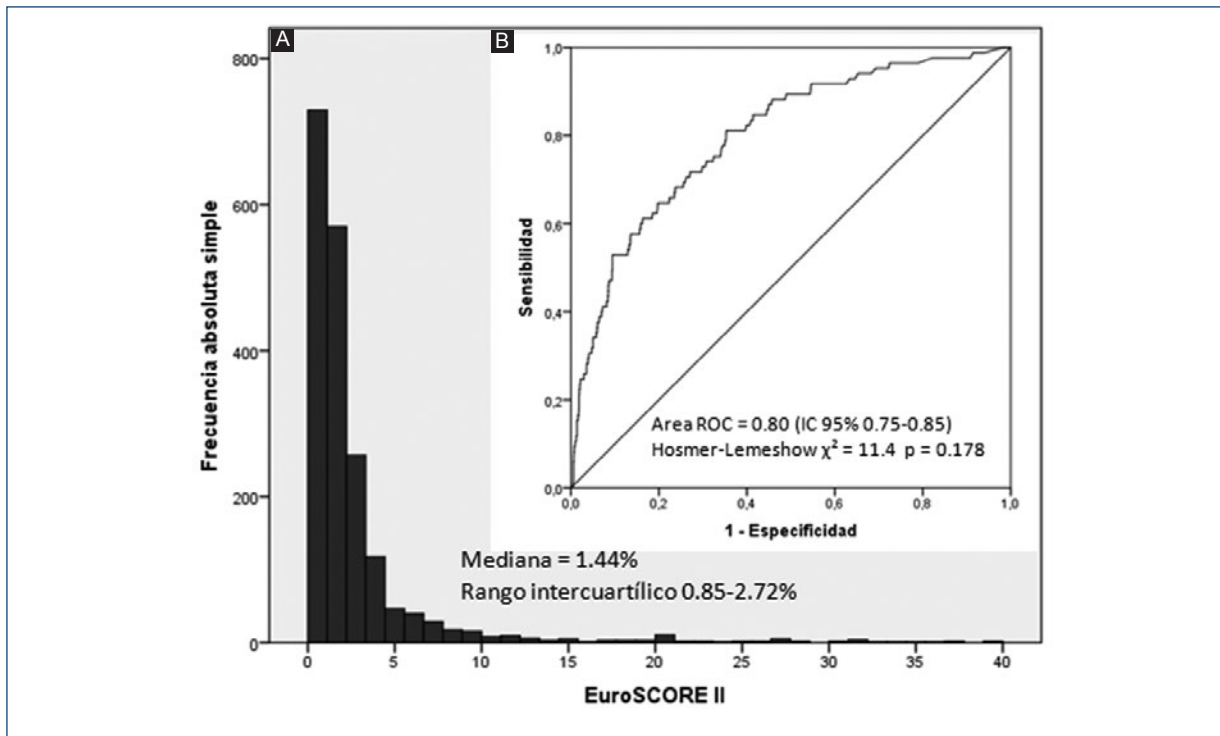


Figura 1. A: Histograma de la distribución de los valores del EuroSCORE II para toda la población de pacientes. Se observa una marcada asimetría a la derecha. **B:** Área bajo la curva ROC (*receiver operating characteristics*) del desempeño del EuroSCORE II para toda la población de pacientes.

En la [figura 3 A](#) se compara la mortalidad hospitalaria observada versus la esperada para todo tipo de cirugía. La razón observada/esperada varió entre 1.0 y 2.1 de acuerdo al grupo de riesgo analizado. En el gráfico de validación clínica del modelo se observa que el EuroSCORE II subestimó el riesgo en los pacientes con riesgos intermedio y alto, pero no los de los extremos de la distribución. De la misma manera, el índice de precisión de Shannon disminuyó desde 0.99 hasta 0.76 a medida que se avanzaba hacia los grupos de mayor riesgo ([Fig. 3 B](#)), lo que indica una reducción de la precisión del modelo. La mortalidad observada y esperada para toda la muestra fue de 4.3 y 3.0%, respectivamente; con una relación observada/esperada de 1.4 ($p = 0.034$). Los datos crudos de la mortalidad observada y esperada separados por grupos de riesgo, así como su comparación estadística, se resumen en la [tabla 2](#). Las complicaciones mayores observadas se presentan al final de la [tabla 1](#).

La utilidad clínica del EuroSCORE II para predecir la mortalidad hospitalaria se analizó con las curvas de decisión de la [figura 4](#), en términos de beneficio neto. Los resultados se presentan con el umbral de

probabilidad en la abscisa (el riesgo de mortalidad operatoria más allá del cual la decisión es no operar al paciente), y los beneficios predichos por el EuroSCORE II en la ordenada, para cada tipo de cirugía.

Discusión

El EuroSCORE original comenzó a usarse en 1999 con sus dos modelos, aditivo y logístico, para estratificar el riesgo operatorio de la cirugía cardíaca, y ha sido ampliamente validado en todo el mundo^{16,17}. A medida que mejoraron los cuidados operatorios y se redujo la mortalidad hospitalaria, el EuroSCORE comenzó a sobreestimar el riesgo quirúrgico¹⁸. Su modelo sucesor es el EuroSCORE II, que apareció en 2012 y en la actualidad sigue un proceso de validación externa en poblaciones distintas a las que fue desarrollado el modelo¹⁹.

En el estudio actual de validación, el EuroSCORE II tuvo un desempeño global adecuado en términos de discriminación y calibración para predecir la mortalidad inmediata de la cirugía cardíaca en esta cohorte. En particular, la capacidad predictiva fue mejor en los pacientes sometidos a cirugía valvular o combinada, que en los operados de revascularización miocárdica

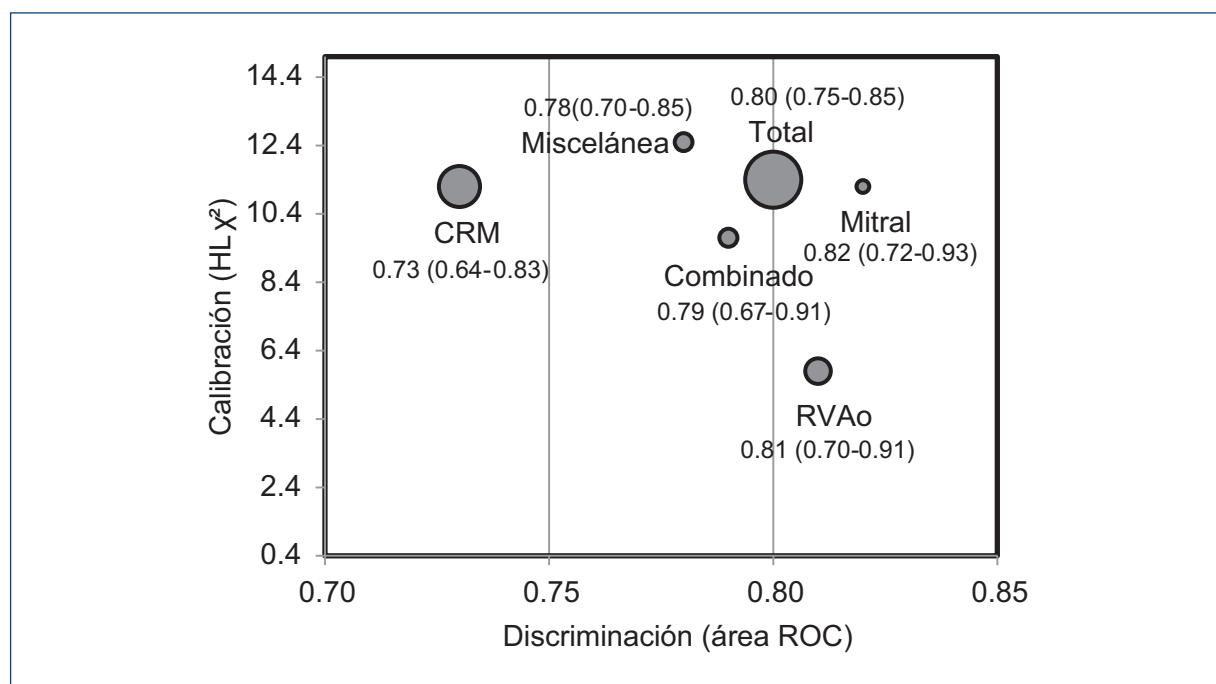


Figura 2. Desempeño del EuroSCORE II para cada tipo de cirugía, en términos de capacidad predictiva. La posición de cada burbuja en el gráfico depende del balance entre la calibración del EuroSCORE II, calculado con el estadístico ch^2 de Hosmer-Lemeshow ($HL\chi^2$), y el poder discriminatorio evaluado con el área bajo la curva ROC (*receiver operating characteristics*). La precisión del EuroSCORE II aumenta de arriba abajo y de izquierda a derecha. El tamaño de los círculos es proporcional al tamaño de la muestra de cada grupo. El umbral de 0.05 para los valores de calibración se halla fuera del gráfico en $HL\chi^2 = 15.0$. Los valores entre paréntesis corresponden a los intervalos de confianza de las áreas ROC. CRM: cirugía de revascularización miocárdica; RVAo: reemplazo valvular aórtico.

Tabla 2. Mortalidad observada y esperada separada por grupos de riesgo

Grupos de riesgo	N	Decesos	Mortalidad observada (O)	Mortalidad esperada (E)	O:E	p
0 a 1	666	5	0.8%	0.7%	1.1	1.000
1 a 2	628	14	2.2%	1.4%	1.6	0.293
2 a 3	292	13	4.5%	2.5%	1.8	0.172
3 a 5	198	11	5.6%	3.8%	1.5	0.481
5 a 10	124	18	14.5%	6.9%	2.1	0.066
10 a 20	44	9	20.5%	13.8%	1.5	0.395
> 20	48	15	31.3%	31.4%	1.0	1.000
Total	2000	85	4.3%	3.0%	1.4	0.034

aislada. La validación clínica del modelo, basada en la relación de la mortalidad observada/esperada y el índice de Shannon, demostró que el sistema se comportó mejor en los grupos de pacientes con riesgo más bajo y más alto, mientras que subestimó el riesgo en los grupos intermedios.

En una publicación previa, se validó el EuroSCORE II en una cohorte de 503 pacientes²⁰, mientras que en esta ocasión se amplió la muestra a 2,000 pacientes y se incluyeron más centros asistenciales. Comparativamente, la serie actual corroboró el mejor desempeño del score en los pacientes sometidos a cirugía cardíaca

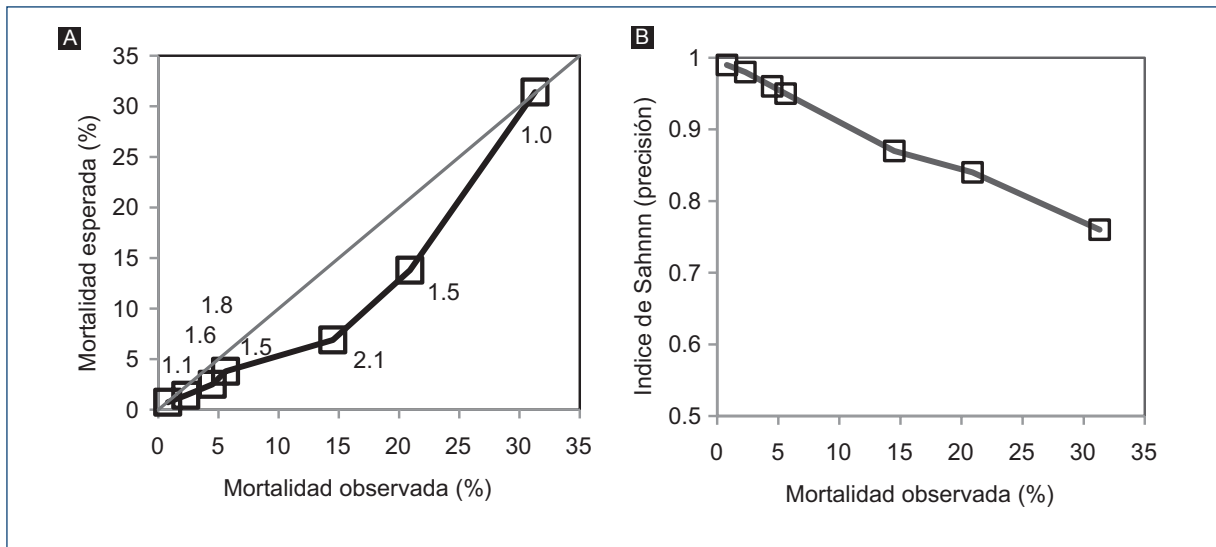


Figura 3. Gráfico de validación clínica del modelo separado por grupos de riesgo. **A:** Mortalidad observada versus esperada de acuerdo al EuroSCORE II para todo tipo de cirugía. Los valores sobre cada punto corresponden a la razón observada/esperada. **B:** Precisión de la predicción basada en el índice de Shannon. Los cuadrados representan pacientes agrupados por el mismo riesgo. Dado que la distribución de riesgos estaba sesgada a la derecha, se prefirió dividir la serie en grupos de diferentes tamaños y no por fractiles.

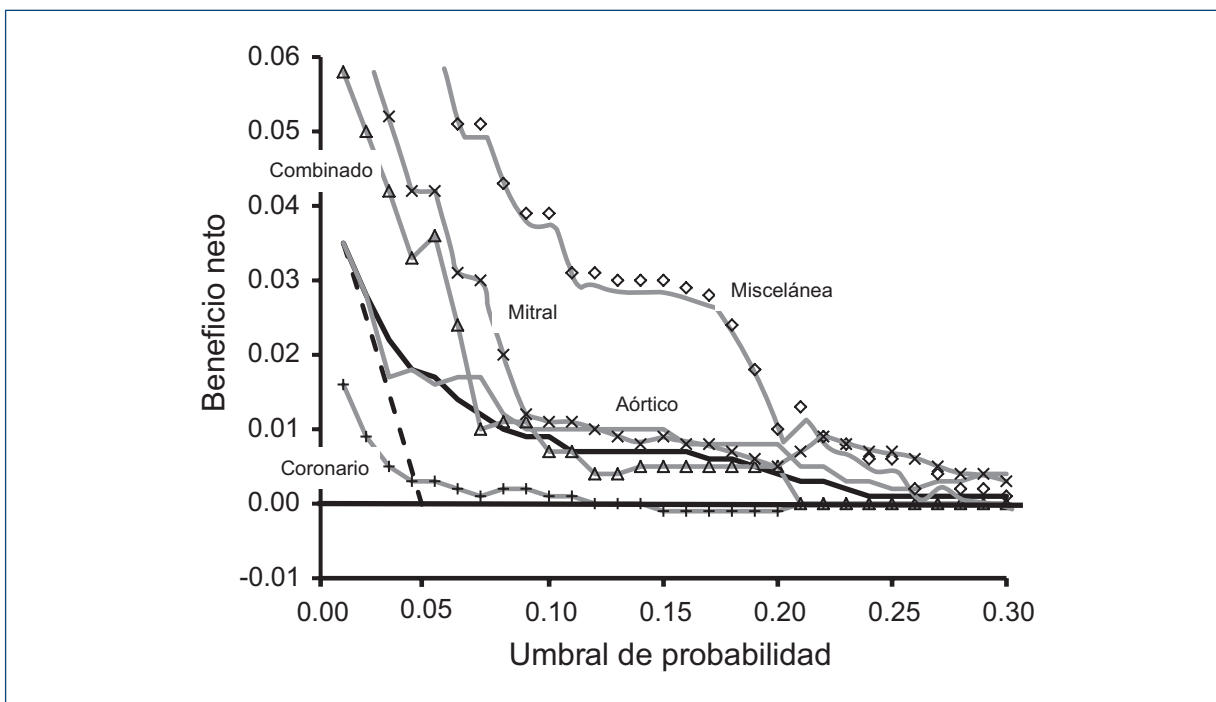


Figura 4. Curvas de decisión para la predicción de la mortalidad hospitalaria del EuroSCORE II aplicada a 2,000 pacientes sometidos a cirugía cardíaca. La línea sólida horizontal que pasa por el cero asume que no se trata ningún paciente (considera que todos morirán después de la cirugía), por lo que el beneficio neto es cero. La línea punteada asume que todos los pacientes son tratados (considera que todos sobrevivirán); mientras que las líneas grises asumen que los pacientes son tratados solo si las predicciones superan cierto umbral, con el riesgo esperado basado en el EuroSCORE II. Cada línea gris con marcas representa un tipo de cirugía distinto. La línea sólida negra resume el beneficio neto de todas las cirugías juntas. El gráfico representa el beneficio neto esperado por paciente para los diferentes umbrales, en comparación a no tratar ningún paciente. Las líneas grises punteadas que asumirían el tratamiento para todos los pacientes en cada tipo de cirugía, no fueron incluidas para dar una mayor claridad al gráfico.

no coronaria. Desde su introducción en la práctica clínica, el EuroSCORE II fue evaluado en más de 50 estudios con diferentes resultados. El metaanálisis de Guida, et al.³, que incluyó 22 estudios, mostró una discriminación global de 0.79, similar a la de nuestro estudio. Las publicaciones más recientes, no incluidas aún en una revisión sistemática, son algunas de las siguientes: Irán²¹, China^{22,23}, Grecia²⁴, España^{25,26}, India²⁷, EE.UU.^{28,29} y Países Bajos³⁰, con áreas ROC que varían entre 0.67 y 0.87.

La comparación de la mortalidad observada en el ámbito local con los mismos resultados internacionales es importante para establecer patrones de calidad y mejoramiento continuo. Por ejemplo, la mortalidad global para todo tipo de cirugía en nuestro estudio fue de 4.3%; mientras que en poblaciones comparables del Reino Unido y de los Países Bajos que emplearon también el EuroSCORE II, la misma tasa de mortalidad fue de 3.0% y 2.7%, respectivamente^{31,32}.

En particular, en nuestro estudio, el EuroSCORE II mostró un buen desempeño para predecir el riesgo del reemplazo valvular aórtico. Aunque sería esperable que este modelo pudiera servir para evaluar también el implante valvular percutáneo, un metaanálisis reciente demostró que el EuroSCORE II alcanzó un poder discriminativo de solo 0.62 con el método endovascular³³. En el metaanálisis de Biancari, et al.⁴, que evaluó exclusivamente el desempeño del EuroSCORE II en el reemplazo valvular aórtico, el área ROC hallada fue de 0.73. Por su parte, Carosella, et al.³⁴ encontraron que un puntaje local, el ArgenSCORE, mostró un mejor poder discriminativo (área ROC = 0.82) que el EuroSCORE II (área ROC = 0.76), cuando compararon los resultados del reemplazo valvular aórtico quirúrgico. En un estudio publicado en 2011, el área ROC del ArgenSCORE fue de 0.80 (0.75-0.85) para todo tipo de cirugía, similar a la hallada en nuestro estudio actual basado en el EuroSCORE II y realizado casi una década después³⁵. De todas formas, por tratarse de un *score* internacional, el EuroSCORE II permitiría la comparación de los resultados locales con los estándares mundiales de calidad, y evitaría así el sesgo de comparación con puntajes desarrollados para uso exclusivamente local.

La variación de la mortalidad quirúrgica entre centros y cirujanos suele ser amplia, y debería tenerse en cuenta al usar cualquier modelo de estratificación del riesgo; por ello, se ha sugerido realizar un ajuste del riesgo predicho por el EuroSCORE II con el llamado Risk-adjusted mortality ratio (RAMR)^{36,37}. El RAMR es un índice

individual para cada centro o cirujano, que corresponde al cociente entre la mortalidad observada sobre la esperada del conjunto de pacientes. En nuestro estudio, este cociente fue 1.4; por lo tanto, para obtener el riesgo esperado real se multiplica este índice por la mortalidad predicha por el EuroSCORE II. Por ejemplo, si el riesgo esperado de mortalidad de un paciente es del 5%, el riesgo real corregido será del 7%.

Cada vez es más frecuente incluir en la evaluación de la utilidad clínica de un modelo un análisis de curvas de decisión. La calibración y la discriminación son aspectos importantes de un modelo de predicción, pero no evalúan la habilidad para tomar mejores decisiones con o sin el uso del modelo. La suma de la sensibilidad y la especificidad basada en el índice de Youden es un indicador promedio de la utilidad de un modelo, pero ignora el peso relativo de los verdaderos y falsos positivos. Recientemente, se ha propuesto el cálculo del beneficio neto basado en diferentes umbrales de decisión para clasificar los pacientes que se beneficiarían o no con el tratamiento. Este análisis compara los riesgos y beneficios de una decisión, calculando la suma ponderada de los casos verdaderos positivos menos los falsos positivos. Según se observa en la **figura 4**, a partir de un umbral de 0.24, el beneficio de tratar un paciente es prácticamente cero, de acuerdo al riesgo esperado calculado con el EuroSCORE II para todo tipo de cirugía.

Al observar la subestimación que hizo el EuroSCORE II en nuestros pacientes de riesgo intermedio se plantean tres interpretaciones posibles: la del comportamiento inadecuado del modelo, la de un pobre desempeño de los cirujanos en ese estrato de riesgo, o ambas a la vez. Cuando se comparan los datos del estudio actual con los resultados hospitalarios de la cirugía valvular aórtica, sola o combinada, en pacientes con riesgo intermedio (EuroSCORE II o STS entre 4 y 7%) realizada en otros dos centros de referencia del país, se observó una menor mortalidad global con respecto a nuestros resultados comunicados en un estudio previo³⁸⁻⁴⁰. Esta observación apoyaría al menos la segunda interpretación sobre la validez del EuroSCORE II en pacientes con riesgo intermedio de nuestra muestra.

Una de las principales fortalezas de nuestro estudio es que los datos fueron recolectados en forma prospectiva desde la implementación del EuroSCORE II, y su utilidad clínica fue analizada con un modelo de curva de decisión. Por su lado, una posible limitación de nuestra cohorte es que casi el 90% de los

pacientes presentaban un riesgo menor o igual a 5%, hecho que podría sesgar el desempeño global del *score*, dado que su comportamiento parece ser mejor en los pacientes con riesgo más bajo. Además, el tamaño de la muestra para algunas categorías o grupos de riesgo podría ser demasiado pequeño para demostrar diferencias estadísticas con respecto a la mortalidad esperada en dicha categoría. Otra limitación es que aunque se trata de una cohorte grande de pacientes de varios centros, la muestra no representa toda la población sometida a cirugía cardíaca en Argentina. Con respecto al buen desempeño del EuroSCORE II en los casos de bajo riesgo, esto podría deberse a que estos pacientes con menos comorbilidades requerirían menor soporte tecnológico de cuidado postoperatorio, lo que suele ser una limitante en los países en desarrollo. En cambio, en los pacientes con riesgo mayor, el desempeño quirúrgico propiamente dicho sería proporcionalmente menos importante que el cuidado postoperatorio. En relación con el buen desempeño del EuroSCORE II en los pacientes con riesgo muy alto, esta observación podría estar sesgada por el pequeño tamaño de la muestra en este estrato ($n = 15$).

Conclusiones

El EuroSCORE II tuvo un desempeño adecuado en términos de discriminación y calibración para todos los tipos de cirugía, aunque algo inferior para la cirugía coronaria. Si bien en términos generales subestimó el riesgo en los grupos de riesgo intermedio, el comportamiento global fue aceptable. El análisis de la curva de decisión para todo tipo de cirugía demostró un beneficio neto positivo para todos los umbrales por debajo de 0.24. El EuroSCORE II podría considerarse una opción de modelo genérico y actualizado de estratificación del riesgo operatorio para predecir la mortalidad hospitalaria de la cirugía cardíaca en nuestro contexto.

Financiamiento

La presente investigación no ha recibido ayudas específicas provenientes de agencias del sector público, sector comercial o entidades sin ánimo de lucro.

Conflicto de intereses

Los autores declaran no tener conflicto de intereses con respecto a este artículo.

Responsabilidades éticas

Protección de personas y animales. Los autores declaran que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

Confidencialidad de los datos. Los autores declaran que han seguido los protocolos de su centro de trabajo sobre la publicación de datos de pacientes.

Derecho a la privacidad y consentimiento informado. Los autores declaran que en este artículo no aparecen datos de pacientes.

Bibliografía

1. Borraacci RA, Arribalzaga EB. Fuzzy logic-based model to stratify cardiac surgery risk. *Rev Argent Cardiol*. 2015;83:304-11.
2. Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. EuroSCORE II. *Eur J Cardiothorac Surg*. 2012; 41:734-44.
3. Guida P, Mastro F, Scarscia G, Whitlock R, Paparella D. Performance of the European System for Cardiac Operative Risk Evaluation II: a meta-analysis of 22 studies involving 145,592 cardiac surgery procedures. *J Thorac Cardiovasc Surg*. 2014;148:3049-57.
4. Biancari F, Juvonen T, Onorati F, Faggian G, Heikkinen J, Airaksinen J, et al. Meta-analysis on the performance of the EuroSCORE II and the Society of Thoracic Surgeons Scores in patients undergoing aortic valve replacement. *J Cardiothorac Vasc Anesth*. 2014;28:1533-9.
5. Sullivan PG, Wallach JD, Ioannidis JP. Meta-Analysis Comparing Established Risk Prediction Models (EuroSCORE II, STS Score, and ACEF Score) for Perioperative Mortality During Cardiac Surgery. *Am J Cardiol*. 2016;118:1574-82.
6. Poullis M, Fabri B, Pullan M, Chalmers J. Sampling time error in EuroSCORE II. *Interact Cardiovasc Thorac Surg*. 2012;14:640-1.
7. Collins GS, Altman DG. Calibration of EuroSCORE II. *Eur J Cardiothorac Surg*. 2013;43:654.
8. Hickey GL, Bridgewater B. How well calibrated is EuroSCORE II? *Eur J Cardiothorac Surg*. 2013;43:208.
9. Collins GS, Altman DG. Design flaws in EuroSCORE II. *Eur J Cardiothorac Surg*. 2013; 43:871.
10. Spiliopoulos K, Bagiatis V, Deutsch O, Kemkes BM, Antonopoulos N, Karangelis D, et al. Performance of EuroSCORE II compared to EuroSCORE I in predicting operative and mid-term mortality of patients from a single center after combined coronary artery bypass grafting and aortic valve replacement. *Gen Thorac Cardiovasc Surg*. 2014;62:103-11.
11. Patrat-Delon S, Rouxel A, Gacouin A, Revest M, Flécher E, Fouquet O, et al. EuroSCORE II underestimates mortality after cardiac surgery for infective endocarditis. *Eur J Cardiothorac Surg*. 2016;49:944-51.
12. Janikowski K, Morawiec R, Jegier B, Jaszewski R, Lelonek M. EuroSCORE II does not show better accuracy nor predictive power in comparison to original EuroSCORE: a single-centre study. *Kardiol Pol*. 2016;74:469-75.
13. Steyerberg EW, Vergouw Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35:1925-31.
14. Borraacci RA, Tajer CD. Aplicación de la teoría de la información a la investigación clínica. *Rev Argent Cardiol*. 2006;74:483-6.
15. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Mak*. 2006;26:565-74.
16. Nashef SA, Roques F, Michel P. European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Cardiothorac Surg*. 1999;16:9-13.
17. Borraacci RA, Rubio M, Baldi J (Jr), Arribalzaga EB, Poveda Camargo RL. Cardiac surgery stratified by EuroSCORE. Long-term survival. *Medicina (Buenos Aires)*. 2013;73:438-42.
18. Parolari A, Pesce LL, Trezzi M, Loardi C, Kassem S, Brambillasca C, et al. Performance of EuroSCORE in CABG and off-pump coronary artery bypass grafting: single institution experience and meta-analysis. *Eur Heart J*. 2009;30:297-304.
19. Collins GS, Altman DG. Design flaws in EuroSCORE II. *Eur J Cardiothorac Surg*. 2013;43:871.
20. Borraacci RA, Rubio M, Celano L, Ingino CA, Allende NG, Ahuad Guerrero RA. Prospective validation of EuroSCORE II in patients undergoing cardiac surgery in Argentinean centres. *Interact Cardiovasc Thorac Surg*. 2014;18:539-43.
21. Atashi A, Amini S, Tashnizi MA, Moeinipour AA, Aazami MH, Tohidnezhad F, et al. External Validation of European System for Cardiac

- Operative Risk Evaluation II (EuroSCORE II) for Risk Prioritization in an Iranian Population. *Braz J Cardiovasc Surg.* 2018;33:40-6.
22. Shen L, Chen X, Gu J, Xue S. Validation of EuroSCORE II in Chinese Patients Undergoing Coronary Artery Bypass Surgery. *Heart Surg Forum.* 2018;21:E036-9.
23. Ma X, Wang Y, Shan L, Cang Z, Gu C, Qu N, et al. Validation of SinoSCORE for isolated CABG operation in East China. *Sci Rep.* 2017;7:16806.
24. Stavridis G, Panaretos D, Kadda O, Panagiotakos DB. Validation of the EuroSCORE II in a Greek Cardiac Surgical Population: A Prospective Study. *Open Cardiovasc Med J.* 2017;11: 94-101.
25. Mateos-Pañero B, Sánchez-Casado M, Castaño-Moreira B, Paredes-Astillero I, López-Almodóvar LF, Bustos-Molina F. Assessment of Euroscore and SAPS III as hospital mortality predicted in cardiac surgery. *Rev Esp Anestesiol Reanim.* 2017;64:273-81.
26. García-Valentín A, Mestres CA, Bernabeu E, Bahamonde JA, Martín I, Rueda C, et al. Validation and quality measurements for EuroSCORE and EuroSCORE II in the Spanish cardiac surgical population: a prospective, multicentre study. *Eur J Cardiothorac Surg.* 2016;49:399-405.
27. Pillai BS, Baloria KA, Selot N. Validation of the European System for Cardiac Operative Risk Evaluation-II model in an urban Indian population and comparison with three other risk scoring systems. *Ann Card Anaesth.* 2015;18:335-42.
28. Osnabrugge RL, Speir AM, Head SJ, Fonner CE, Fonner E, Kappetein AP, et al. Performance of EuroSCORE II in a large US database: implications for transcatheter aortic valve implantation. *Eur J Cardiothorac Surg.* 2014;46:400-8; discussion 408.
29. Hernández-Vaquero D, Díaz R, Meana B, Morís C. External validation of the EuroSCORE II risk stratification model in the USA. *Eur J Cardiothorac Surg.* 2015;48:177.
30. Hogervorst EK, Rosseel PMJ, van de Watering LMG, Brand A, Bentala M, van der Meer BJM, et al. Prospective validation of the EuroSCORE II risk model in a single Dutch cardiac surgery centre. *Neth Heart J.* 2018;26(11):540-51.
31. Bridgewater B, Hickey GL, Cooper G, Deanfield J, Roxburgh J; Society for Cardiothoracic Surgery in Great Britain and Ireland; National Institute for Clinical Outcomes Research, UCL. Publishing cardiac surgery mortality rates: lessons for other specialties. *BMJ.* 2013;346:f1139.
32. Siregar S, Groenwold RH, Versteegh MI, Takkenberg JJ, Bots ML, van der Graaf Y, et al. Data Resource Profile: adult cardiac surgery database of the Netherlands Association for Cardio-Thoracic Surgery. *Int J Epidemiol.* 2013;42:142-9.
33. Wang TKM, Wang MTM, Gamble GD, Webster M, Ruygrok PN. Performance of contemporary surgical risk scores for transcatheter aortic valve implantation: A meta-analysis. *Int J Cardiol.* 2017;236:350-5.
34. Carosella VC, Mastantuono C, Golovonevsky V, Cohen V, Grancelli H, Rodríguez W, et al. Prospective and Multicentric Validation of the ArgenSCORE in Aortic Valve Replacement Surgery. Comparison with the EuroSCORE I and the EuroSCORE II. *Rev Argent Cardiol.* 2014;82:6-12.
35. Carosella VC, Grancelli H, Rodríguez W, Sellanes M, Cáceres M, Cohen Arazi H, et al. External and Temporal Validation 10 Years after the Development of the First Latin-American Risk Stratification Score for Cardiac Surgery (ArgenSCORE). *Rev Argent Cardiol.* 2011;79:500-7.
36. García-Villarreal OA. EuroSCORE II. Cómo se usa en la práctica diaria. *Rev Mex Cardiol.* 2014;25:50-51.
37. García-Villarreal OA: Risk-adjusted mortality ratio for EuroSCORE II in everyday clinical practice. *Interact Cardiovasc Thorac Surg.* 2014;18:450.
38. Navia D, Piccinini F, Vranvic M, Camporotondo M, Espinoza J, Simonetto B, et al. Early and Long-term Outcomes of Aortic Valve Replacement Surgery in Low- and Intermediate-risk Patients. *Rev Argent Cardiol.* 2018;86:190-5.
39. Fortunato GA, Marenquino RG, Cirio S, Rossi E, Domenech A, Kotowicz V. Aortic Valve Replacement in Intermediate Risk Patients: Surgical Outcomes. *Rev Argent Cardiol.* 2018;86:114-7.
40. Borracci RA, Rubio M, Baldi J Jr, Ahuad Guerrero RA, Mauro V, Ingino CA. In-hospital Outcomes of Surgical Aortic Valve Replacement: The Benchmark for Transcatheter Valve Implant. *Rev Argent Cardiol.* 2018;86:196-9.