

Original

Validación de EuroSCORE II en España[☆]



Antonio García-Valentín^{a,b,*}, Eduardo Bernabeu^b, Daniel Pereda^c, Miguel Josa^c, José María Cortina^d, Carlos A. Mestres^c, Jacobo Silva^e, Patricio Llamas^b, Eduardo Otero^a, Ángel L. Fernández^f y el Grupo de Trabajo para el Proyecto de Validación de EuroSCORE II en España

^a Servicio de Cirugía Cardiovascular, Hospital Clínico Universitario de Valencia, Valencia, España

^b Servicio de Cirugía Cardiovascular, Hospital General Universitario de Alicante, Alicante, España

^c Servicio de Cirugía Cardiovascular, Hospital Clínic de Barcelona, Barcelona, España

^d Servicio de Cirugía Cardiovascular, Hospital Universitario 12 de Octubre, Madrid, España

^e Servicio de Cirugía Cardiovascular, Hospital Universitario Ramón y Cajal, Madrid, España

^f Servicio de Cirugía Cardiovascular, Hospital Clínico Universitario de Santiago, Santiago de Compostela, España

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:

Recibido el 15 de junio de 2014

Aceptado el 17 de julio de 2014

On-line el 26 de septiembre de 2014

Palabras clave:

Estadística

Cirugía cardíaca

Estudios de validación

Mortalidad

Estudios prospectivos

R E S U M E N

Introducción: EuroSCORE ha sido el modelo de mortalidad quirúrgica de referencia desde finales del siglo xx, aunque diversos estudios han demostrado una pérdida progresiva de calibración. Esto condujo al desarrollo de su versión actualizada, EuroSCORE II. El objetivo de este trabajo fue el diseño y la realización de un proyecto de validación a nivel nacional en España.

Métodos: Se diseñó un estudio prospectivo multicéntrico de participación voluntaria. Se obtuvieron los datos de EuroSCORE, EuroSCORE II y la mortalidad observada. Los datos fueron depurados hasta obtener una tasa de datos perdidos o incoherentes del 0%. La calibración de los modelos se evaluó mediante la prueba de bondad de ajuste de Hosmer-Lemeshow y la discriminación mediante el área bajo la curva ROC.

Resultados: Veinte centros españoles participaron en el estudio, reclutándose 4.034 pacientes entre el 1 de octubre del 2012 y el 31 de marzo del 2013. Se analizó la prevalencia de factores de riesgo. La mortalidad observada fue del 6,5%. Los valores medios de EuroSCORE aditivo, logístico y EuroSCORE II fueron del 6,5, el 9,8 y el 5,7%, respectivamente. Las áreas bajo la curva ROC fueron EuroSCORE: 0,77 (IC del 95%, 0,74–0,80), EuroSCORE II: 0,79 (IC del 95%, 0,76–0,82). Los resultados de la prueba de bondad de ajuste fueron EuroSCORE: 33,02 ($p < 0,001$), EuroSCORE II: 38,98 ($p < 0,001$).

Conclusiones: Los pacientes españoles presentan un perfil de riesgo alto. La mortalidad cruda es aceptable, más cercana al valor de EuroSCORE II que de EuroSCORE. Ambas escalas muestran mala calibración, EuroSCORE por sobreestimación y EuroSCORE II por infraestimación. El área bajo la curva ROC muestra buena discriminación para ambos modelos.

© 2014 Sociedad Española de Cirugía Torácica-Cardiovascular. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Validation of EuroSCORE II in Spain

A B S T R A C T

Keywords:

Statistics

Cardiac surgery

Validation studies

Mortality

Prospective studies

Introduction: EuroSCORE has been the predictive model of choice for surgical mortality since the late 1990's, although several studies have shown a progressive loss of calibration. This led to the development of EuroSCORE II, a new version of the model. The aim of the study was to design and conduct a national validation study.

Methods: A prospective and multicenter study was designed based on volunteer participation. Data regarding EuroSCORE and EuroSCORE II, and observed mortality were acquired. Data were refined to obtain an actual rate of loss, or incoherent, data of 0%. The calibration of the models was evaluated with the Hosmer-Lemeshow goodness-of-fit test, and discrimination with the area under the ROC curve.

Results: A total of 20 Spanish centres participated in this study, with 4034 patients recruited between 1 October 2012 and 31 March 2013. Risk factor prevalence was analysed. Observed mortality was 6.5%. The mean values for additive and logistic EuroSCORE and EuroSCORE II were 6.5, 9.8% and 5.7%, respectively. Areas under the ROC curves were EuroSCORE: 0.77 (95% CI; 0.74–0.80), EuroSCORE II: 0.79 (95% CI; 0.76–0.82). Results for the goodness-of-fit test were EuroSCORE: 33.02 ($P < .001$), EuroSCORE II: 38.98 ($P < .001$).

Conclusions: The risk profile in Spanish patients is high. Crude mortality is acceptable, closer to the value predicted by EuroSCORE II than by EuroSCORE. Both models show failure in calibration; EuroSCORE

[☆] Este artículo ha recibido el premio a mejor comunicación del XXII Congreso SECTCV 2014 – Santiago de Compostela.

* Autor para correspondencia.

Correo electrónico: antonio@garciavalentin.es (A. García-Valentín).

by over-prediction and EuroSCORE II by under-prediction of surgical risk. Areas under the ROC curve show good discrimination for both models.

© 2014 Sociedad Española de Cirugía Torácica-Cardiovascular. Published by Elsevier España, S.L.U. All rights reserved.

Introducción

La práctica de la cirugía cardiovascular incluye de forma habitual la evaluación del riesgo quirúrgico de los pacientes¹, de forma que no se puede separar el estudio de un caso clínico del cálculo de su correspondiente probabilidad de mortalidad estimada en el momento de sentar una indicación quirúrgica.

Las herramientas matemáticas en las que se ha basado esta tarea han sido los modelos predictivos, que son métodos estadísticos de integración de las magnitudes de asociación estadística de las distintas variables predictoras con el resultado que deseamos pronosticar, en especial la mortalidad peroperatoria. En la breve historia de nuestra especialidad han existido numerosos y diversos sistemas de predicción de mortalidad²⁻⁴, la mayoría ya en desuso. En la época actual, debemos destacar como modelos de referencia a nivel internacional al desarrollado por la Society of Thoracic Surgeons (conocido como STS Score, y de uso principalmente en Norteamérica) y EuroSCORE, el modelo predictivo más utilizado en Europa.

El STS Score es un modelo construido a partir de la recopilación de datos de los pacientes sometidos a una intervención cardiaca mayor (incluyendo adultos e infantil) en más de 250 hospitales norteamericanos. Esta recogida de datos comenzó en 1989 y continúa en la actualidad, habiéndose registrado más de 4 millones y medio de pacientes. El modelo de predicción desarrollado presenta actualizaciones periódicas (la última es del año 2009), que se efectúan sobre la base de la totalidad de los datos, lo que supone un excelente y constante método de revisión y ajuste⁵.

El EuroSCORE, sin embargo, fue un modelo desarrollado a partir de datos recogidos durante en el año 1995⁶. Participaron 132 centros de 8 países europeos, incluyendo el nuestro, con cuyos datos se construyó un modelo que fue finalmente publicado en 1999⁷. La popularidad de esta escala alcanzó un nivel muy notable en los años siguientes a su publicación. Fueron llevados a cabo multitud de estudios de validación externa que alabaron la buena capacidad predictiva del modelo en poblaciones geográfica, social y culturalmente muy dispares⁸⁻¹¹, y que incluso demostraron excelentes parámetros de rendimiento en subgrupos específicos de pacientes¹² o incluso para técnicas que no habían sido incluidas inicialmente en el análisis de desarrollo, como la cirugía coronaria sin circulación extracorpórea (CEC)¹³⁻¹⁵. EuroSCORE nunca fue validado a nivel multicéntrico en nuestro país, aunque los análisis por subgrupos nacionales que se realizaron sobre la base de datos original mostraron datos de interés, como una mayor mortalidad cruda en España, derivada de una mayor complejidad del perfil de riesgo de los pacientes y de un mayor porcentaje de cirugía no coronaria¹⁶.

Sin embargo, el éxito inicial de este modelo se desvaneció a finales de la década de los 2000, momento en el que comenzó a hacerse frecuente la publicación de estudios de validación externa que acusaban al modelo de fallo de calibración por sobreestimación del riesgo de mortalidad y redundaban en la superioridad predictiva de otros modelos, especialmente STS Score¹⁷⁻²⁰. Asimismo resultó de especial interés la pérdida de calibración en grupos, como pacientes de alto riesgo o aquellos sometidos a cirugía valvular aórtica²¹⁻²³, en la época crucial de expansión de la técnica de implante de prótesis aórticas transcáteter, y que incluso requirió la adaptación de los criterios de alto riesgo quirúrgico en los principales estudios realizados al respecto²⁴.

En resumen, y como exponía Cortina en un editorial previo¹, EuroSCORE quedó obsoleto por su falta de actualización, suponiendo en la actualidad un eco de la actividad y resultados que se desarrollaban y obtenían hace casi 20 años. La respuesta del equipo de EuroSCORE fue la creación de un nuevo modelo basado en el esquema del ya existente, aunque con datos actualizados y reestructuración de las variables. El proyecto se ejecutó en 2010 y EuroSCORE II fue presentado a finales del 2011 y publicado en febrero del 2012, contando con la participación de 154 centros de 43 países de los 5 continentes y siendo España el que más centros aportó al estudio (aunque no el que más pacientes)²⁵. La publicación de este trabajo vino envuelta en polémica por distintas causas; la primera fue el cambio de horizonte temporal de predicción del nuevo modelo, que pasó de pronosticar la mortalidad intrahospitalaria y a 30 días a limitarse al periodo únicamente intrahospitalario. La única razón para este ajuste fue la escasa aportación de datos de seguimiento a 90 días por parte de los centros participantes en el desarrollo, hecho que, por otro lado, indica una escasa calidad de los datos recogidos en Europa. La segunda causa fue el súbito descenso en la mortalidad estimada para pacientes complejos, que si bien corregía el efecto de sobreestimación del modelo previo, aparentaba caer en un fallo por infraestimación, con las implicaciones que ello podría llevar a efectos de control de la calidad asistencial. Es por todo esto que este modelo se acompañó en su publicación de un durísimo editorial que criticaba su calidad científica, acusándola de graves errores metodológicos, y desaconsejaba su utilización para fines asistenciales y de control de calidad²⁶. Los estudios de validación publicados hasta el momento muestran resultados dispares, incluyendo algún trabajo realizado en un centro de nuestro país, que mostró un fallo de calibración de la nueva escala²⁷.

En este escenario, y con el soporte de la Sociedad Española de Cirugía Torácica-Cardiovascular (SECTCV), se decidió llevar a cabo un proyecto de validación nacional de EuroSCORE II en España, en lo que, hasta la fecha, supondría el primer estudio multicéntrico y prospectivo de validación externa del modelo.

Métodos

Diseño del estudio

El proyecto fue concebido y diseñado por miembros de la SECTCV, sin financiación externa ni remuneración para sus autores. Fueron invitados a participar todos los centros del país miembros de la SECTCV que realizasen cirugía cardiaca del adulto. La llamada a la colaboración se realizó por distintas vías: directamente por entrevista personal, a través de publicación del mismo durante el XXI Congreso de la SECTCV, celebrado en Sevilla en junio del 2012, así como a través del envío global de correos electrónicos a los socios desde la secretaría de la organización. La decisión final de participar en el proyecto fue voluntaria.

Definiciones

EuroSCORE: European System for Cardiac Operative Risk Evaluation. Versión 1999.

EuroSCORE II: European System for Cardiac Operative Risk Evaluation. Versión 2012.

Cirugía cardiaca mayor: intervención quirúrgica para tratamiento de patología cardiaca o del pericardio que se realiza habitualmente bajo anestesia general, requiere apertura del tórax en forma de esternotomía o toracotomía, así como del pericardio para su realización, y trata generalmente patología estructural del corazón y grandes vasos, en la mayoría de las ocasiones con soporte de CEC.

Cirugía cardiaca menor: intervención quirúrgica para tratamiento de patología cardiaca o del pericardio que se realiza habitualmente bajo anestesia local, requiere pequeñas incisiones de acceso, no suele tratar patología estructural y no se realiza bajo CEC.

SICCS: Sistema Informático de Cirugía Cardiovascular de Sorin. Base de datos para Cirugía Cardiovascular comercializada por Biomeconco SL.

Prueba de bondad de ajuste: prueba estadística que compara los datos esperados con los observados, divididos en deciles de riesgo. Emite un valor de estadístico y de p como resultado, y se aplica a la calibración de modelos predictivos.

Curva ROC: del inglés Receiver Operating Characteristic. Representación gráfica de la sensibilidad frente a la tasa de falsos positivos de una prueba diagnóstica o modelo de predicción. El área bajo la curva representa la discriminación del modelo.

SPSS: aplicación informática de análisis estadístico comercializada por IBM.

Microsoft Excel: aplicación informática de hoja de cálculo comercializada por Microsoft Corporation.

Objetivos del estudio

El objetivo primario del estudio fue evaluar los parámetros de rendimiento (calibración y discriminación) del modelo EuroSCORE II en la población de pacientes sometidos a cirugía cardiaca en España. Secundariamente, también se evaluó el rendimiento de la escala EuroSCORE en los mismos pacientes, así como un análisis descriptivo de los factores de riesgo.

Tamaño muestral

En aras de conseguir un tamaño muestral adecuado, se calculó un periodo de inclusión de 6 meses. Estimando una participación del 25% de los centros del país y una mortalidad cruda del 5% (extrapolada de registros previos), el periodo seleccionado aparentaba ser óptimo para obtener los datos de aproximadamente 2.000 pacientes, con un número de eventos (mortalidad) estimado de 100 pacientes, lo cual otorgaba una potencia estadística adecuada para este estudio^{28,29}.

Recogida de datos

La recogida de datos se realizó finalmente entre el 1 de octubre del 2012 y el 31 de marzo del 2013, y fueron incluidos todos aquellos procedimientos de cirugía cardiaca mayor con o sin CEC realizados en pacientes adultos. Se excluyó a los pacientes pediátricos, los sometidos a cirugía menor, cirugía de marcapasos que no precisa CEC, cirugía de la pared torácica, implante de válvulas transcáteter o trasplantes, ya que ninguno de estos procedimientos se incluyó en la base de datos de desarrollo de EuroSCORE II. También fueron excluidas para su análisis las reintervenciones por cualquier causa que se realizaran en el mismo ingreso de la cirugía por la que ingresó el paciente originalmente.

Fueron recogidos todos los datos relativos al cálculo de EuroSCORE y EuroSCORE II, así como la mortalidad observada. Dado que el objetivo principal del estudio fue la validación de EuroSCORE II, que contempla el resultado de mortalidad como aquella ocurrida en el periodo postoperatorio, se registró la mortalidad de los pacientes

intervenidos hasta la fecha del alta. Para su obtención, se emplearon diversos sistemas informáticos: los centros que disponían de la base de datos SICCS fueron invitados a usarla, de forma que se diseñó una herramienta de extracción de los datos necesarios para el estudio, implementándose también un módulo de cálculo de EuroSCORE II. Los centros que disponían de otra base de datos capaz de proporcionar los datos necesarios para el cálculo de EuroSCORE II fueron autorizados para utilizarla. Finalmente, a los centros que no disponían de ninguno de estos programas, les fue facilitada una herramienta informática basada en Microsoft Excel® (Microsoft Corporation, Redmond, EE. UU.), diseñada para el cálculo de los datos y el almacenamiento de los pacientes intervenidos, diseñada por los autores.

Finalizado el periodo de reclutamiento, cada centro envió sus datos a una única persona, que realizó el depurado de los mismos en busca de datos perdidos e incoherentes. Para ello se crearon herramientas lógicas que permitían detectar estos defectos. La totalidad de los datos obtenidos fue sometida a este proceso y los centros participantes fueron invitados a corregir todos los errores detectados y reenviar nuevamente los datos, con el objetivo de obtener un porcentaje de datos perdidos o incoherentes cercano al 0%. Tras ello, todos los archivos fueron enmascarados y el dato de mortalidad fue añadida al conjunto de datos, que se volcó en su totalidad a un único archivo de SPSS® (IBM, Armonk, EE. UU.) para su análisis estadístico.

Análisis estadístico

No se realizaron análisis intermedios ni por subgrupos o centros de procedencia. La evaluación del rendimiento del modelo se realizó mediante análisis de la capacidad de discriminación y la calibración del mismo. Para el cálculo de la discriminación se utilizó el área bajo la curva ROC (valor calculado e IC del 95%). La calibración se analizó mediante la prueba de bondad de ajuste de Hosmer-Lemeshow (valor del estadístico y p-valor). El análisis descriptivo de los datos se expresó mediante media $\pm$ desviación estándar para las variables continuas (edad, valores medios de EuroSCORE) y frecuencias (porcentajes) para las discontinuas y la mortalidad cruda.

Resultados

Veinte centros de toda España participaron finalmente en el proyecto, aportando por completo los datos solicitados y la mortalidad peroperatoria (tabla 1). El conjunto de datos definitivo incluyó a

Tabla 1

Centros participantes en el Proyecto de Validación Nacional de EuroSCORE II

Complejo Asistencial Universitario de León (León)
Complejo Hospitalario de Navarra (Pamplona)
Complejo Hospitalario Universitario A Coruña (A Coruña)
Clínica Capio Albacete (Albacete)
Hospital Clínic de Barcelona (Barcelona)
Hospital Clínico San Carlos (Madrid)
Hospital Clínico Universitario de Salamanca (Salamanca)
Hospital Clínico Universitario de Santiago (Santiago de Compostela)
Hospital Clínico Universitario de Valladolid (Valladolid)
Hospital Universitario de Cruces (Bilbao)
Hospital de la Santa Creu i Sant Pau (Barcelona)
Hospital General Universitario de Alicante (Alicante)
Hospital Universitari Germans Trias i Pujol (Badalona)
Hospital Universitari Vall d'Hebron (Barcelona)
Hospital Universitario Central de Asturias (Oviedo)
Hospital Universitario 12 de Octubre (Madrid)
Hospital Universitario Fundación Jiménez Díaz (Madrid)
Hospital Universitario La Paz (Madrid)
Hospital Universitario Virgen de las Nieves (Granada)
Hospital Universitario Virgen del Rocío (Sevilla)

Tabla 2
Análisis demográfico y de comorbilidades

Edad, media $\pm$ DE	66,6 $\pm$ 12,3
Sexo femenino, n (%)	1.459 (36,2)
Enfermedad pulmonar, n (%)	334 (8,3)
Arteriopatía extracardiaca, n (%)	479 (11,9)
Movilidad limitada, n (%)	150 (3,7)
Disfunción renal	
CrCl > 85 ml/min, n (%)	1.364 (33,6)
CrCl 51-85 ml/min, n (%)	1702 (42,2)
CrCl < 51 ml/min n (%)	934 (23,2)
Díalisis crónica n (%)	34 (0,8)
Diabetes mellitus ID, n (%)	272 (6,7)
Reintervención, n (%)	410 (10,2)
Endocarditis activa, n (%)	126 (3,1)
Estado crítico preoperatorio, n (%)	276 (6,8)
Angina CCS 4, n (%)	302 (7,5)
Clase funcional	
NYHA I, n (%)	819 (20,3)
NYHA II, n (%)	1.426 (35,3)
NYHA III, n (%)	1.439 (35,7)
NYHA IV, n (%)	350 (8,7)
Disfunción ventricular izquierda	
FE > 50%, n (%)	3213 (79,6)
FE 31-50%, n (%)	395 (9,8)
FE 21-30%, n (%)	132 (3,3)
FE < 21%, n (%)	294 (7,3)
Infarto reciente	491 (12,2)
Hipertensión pulmonar	
PAPs 31-55 mmHg, n (%)	871 (21,6)
PAPs > 55 mmHg, n (%)	497 (12,3)
Prioridad de la intervención	
Electiva, n (%)	2260 (56,1)
Paciente ingresado, n (%)	1582 (39,2)
Urgencia, n (%)	182 (4,5)
Salvamento, n (%)	10 (0,2)
Complejidad	
Un procedimiento coronario, n (%)	1.024 (25,4)
Un procedimiento no coronario, n (%)	1.610 (39,9)
2 procedimientos, n (%)	1.131 (28,0)
3 o más procedimientos, n (%)	269 (6,7)
Cirugía sobre aorta torácica, n (%)	373 (9,2)

CCS: Canadian Cardiovascular Society; CrCl: aclaramiento de creatinina; DE: desviación estándar; FE: fracción de eyección; ID: insulinodependiente; n: número de pacientes; NYHA: New York Heart Association; PAPs: presión arterial pulmonar sistólica.

4.034 pacientes. La media de pacientes por centro fue de 201,7 para dicho periodo, y que debe suponer más del 50% de la actividad media anual, dado que no incluyó el periodo estival. El porcentaje de datos perdidos o incoherentes fue del 0%, lo cual refleja una muy alta calidad de los datos recogidos en España^{25,30}. Los datos relativos a los factores de riesgo de la población del estudio están recogidos en la [tabla 2](#), de la que podemos destacar la baja representación de la cirugía coronaria dentro de la actividad global, el avanzado estado clínico (más del 44% en clase funcional III o IV de la New York Heart Association) y el gran porcentaje de pacientes intervenidos de forma no electiva (casi la mitad). Estos datos serán convenientemente analizados y discutidos en trabajos posteriores.

La mortalidad cruda fue de 262 pacientes (6,5%). La caracterización del riesgo según la versión aditiva de EuroSCORE fue de 6,5 puntos (riesgo alto). El valor promedio $\pm$ desviación estándar de la mortalidad estimada fue de 9,8% $\pm$ 12,3% para EuroSCORE y 5,7% $\pm$ 8,7% para EuroSCORE II.

En cuanto a la evaluación de la calibración, EuroSCORE mostró fallo por sobreestimación del riesgo quirúrgico, con un resultado fallo para la prueba de bondad de ajuste de 33,02 ($p < 0,001$). EuroSCORE II mostró también fallo de calibración con un valor para la prueba de Homer-Lemeshow de 38,98 ($p < 0,001$), aunque en su caso por

Tabla 3
Prueba de bondad de ajuste para EuroSCORE

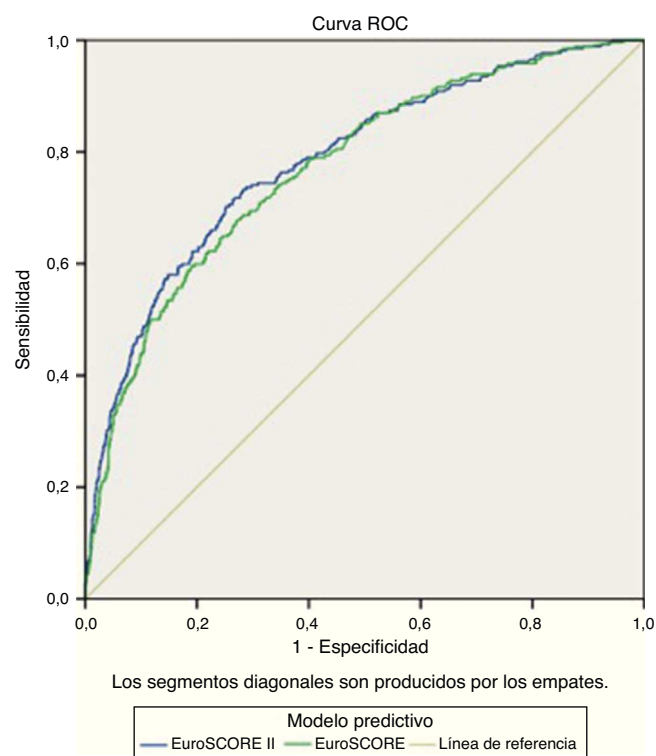
Grupo de riesgo	Pacientes fallecidos		Pacientes vivos		Total
	Observado	Esperado	Observado	Esperado	
1	3	13.911	394	383.089	397
2	8	14.877	401	394.123	409
3	6	15.339	397	387.661	403
4	12	16.239	393	388.761	405
5	15	16.887	387	385.113	402
6	20	18.076	384	385.924	404
7	20	19.644	383	383.356	403
8	30	22.762	375	382.238	405
9	47	30.094	356	372.906	403
10	101	94.171	302	308.829	403

Tabla 4
Prueba de bondad de ajuste para EuroSCORE II

Grupo de riesgo	Pacientes fallecidos		Pacientes vivos		Total
	Observado	Esperado	Observado	Esperado	
1	3	14.629	405	393.371	408
2	7	14.623	390	382.377	397
3	9	15.327	396	389.673	405
4	11	15.654	390	385.346	401
5	15	16.424	389	387.576	404
6	14	17.236	390	386.764	404
7	15	18.569	389	385.431	404
8	32	21.315	373	383.685	405
9	47	28.561	356	374.439	403
10	109	99.660	294	303.340	403

infraestimación del riesgo. Las tablas de mortalidad esperada y observada por deciles de riesgo muestran similar distribución del error para ambos modelos, con infraestimación en deciles de alto riesgo y sobreestimación en deciles de bajo riesgo ([tablas 3 y 4](#)).

Las curvas ROC mostraron muy buena capacidad de discriminación en ambos modelos, con áreas bajo la curva ROC virtualmente

**Figura 1.** Curvas ROC para EuroSCORE y EuroSCORE II.

idénticas: 0,77 (IC del 95%, 0,74-0,80) para EuroSCORE y 0,78 (IC del 95%, 0,76-0,82) para EuroSCORE II (fig. 1).

Discusión

Los resultados de este estudio reflejan una realidad interesante, ya apreciada en los estudios previos que se realizaron con EuroSCORE. El nivel de complejidad en la población española sometida a cirugía cardíaca, reflejado en la estimación de riesgo de mortalidad peroperatoria, es especialmente alto (6,5 en la escala aditiva, 9,8% según EuroSCORE logístico y 5,7% según EuroSCORE II). El estudio sobre EuroSCORE que incluyó a población española fue el de análisis por países del conjunto de datos original de EuroSCORE y reflejó un riesgo medio según la escala aditiva, con un valor de 4,7¹⁶. En nuestro estudio, este valor fue de 6,5, lo que indica un empeoramiento del perfil de comorbilidad de nuestros pacientes desde esa fecha. Los estudios de validación de EuroSCORE realizados en otros tiempos y poblaciones estimaron una mortalidad esperada entre el 3 y el 4%^{11,31,32}, lo que confirma el empeoramiento del perfil de riesgo de los pacientes españoles (mortalidad estimada 9,8%). Los estudios de validación para EuroSCORE II publicados hasta el momento concuerdan con estos datos, estimando una mortalidad del 3-4%^{33,34}, mientras que en nuestro país este valor se eleva hasta el 5,7%. Estos datos son muy similares a los publicados por Silva et al. para su centro en este mismo medio²⁷, en el que comunicaron una mortalidad esperada según EuroSCORE y EuroSCORE II del 9,1 y el 3,6%, respectivamente. Posiblemente, el valor del 5,7% para EuroSCORE II supera al del cálculo que realizaron estos autores en su entorno, pero es más acorde y proporcional a la elevada mortalidad pronosticada por EuroSCORE.

A pesar del elevado riesgo quirúrgico que presentan nuestros pacientes, la mortalidad cruda observada que revela nuestro estudio es del 6,5%, que es un valor aceptable en el contexto de alto riesgo descrito. En el único precedente publicado con pacientes españoles por los autores de EuroSCORE, ya comentado anteriormente, la mortalidad cruda de los pacientes españoles fue del 8,3%¹⁶, por lo cual podemos concluir que, a pesar del empeoramiento del perfil de riesgo de los pacientes, los resultados han mejorado desde entonces.

Los datos demográficos y el análisis de comorbilidad muestran datos muy interesantes, como la baja proporción de cirugía coronaria en nuestro país, o el apreciable número de pacientes que se intervienen durante el ingreso hospitalario y que serán objeto de análisis en estudios posteriores.

La discriminación de ambos modelos ha sido muy buena, con valores del área bajo la curva ROC cercanos a 0,8, que son concordantes con los estudios realizados desde la aparición de EuroSCORE. De hecho, el problema principal de este modelo ha sido la progresiva pérdida de calibración, como ya se comentó, dado que los datos de discriminación han sido adecuados y estables hasta la fecha. La supuesta superioridad de EuroSCORE II sobre el modelo anterior en términos de discriminación²⁵ no ha sido constatada en nuestro estudio, en el que la discriminación de ambos modelos es prácticamente idéntica, aunque este estudio no fue diseñado para comparar ambas curvas ROC.

El análisis de los datos de calibración coloca el nivel de mortalidad peroperatoria en nuestro país en un punto intermedio entre las predicciones de ambos modelos. La prueba de bondad de ajuste muestra fallo de calibración en las 2 escalas, en EuroSCORE por sobreestimación del riesgo quirúrgico (problema ampliamente debatido en la literatura, como ya se comentó anteriormente) y EuroSCORE II por ligera infraestimación. En análisis intuitivo de los mismos datos revela que nuestra mortalidad cruda es relativamente cercana, aunque mayor a la esperada por EuroSCORE II y mucho más lejana a la pronosticada por EuroSCORE, por lo que

podríamos asumir que el mejor modelo en nuestro medio sería EuroSCORE II, si bien es muy probable, por los datos analizados y los resultados obtenidos, que el modelo ideal se sitúe entre EuroSCORE y EuroSCORE II. Este dato también será fruto de un análisis más profundo en otro estudio. La tendencia a la infraestimación de EuroSCORE II ha sido admitida por los mismos autores en la publicación original (que la consideraba «aceptable»)²⁵, aunque la calibración que se realizó dentro del mismo conjunto de datos de desarrollo sido objetivo de controversia, ya que el valor de la prueba de bondad de ajuste rozó el fallo en la validación interna original. En cartas al director y respuestas posteriores^{35,36}, fue comunicado un p-valor de 0,09 para la prueba de bondad de ajuste, dato que, a pesar de ser superior al límite de significación estadística habitual, no resulta convincente para justificar una buena calibración, ya que teóricamente el modelo estará mejor calibrado cuanto más cercano a 1 sea dicho valor de p³⁷. Un fallo de calibración puede ser justificado por múltiples y diversas causas, entre las que se pueden mencionar el mal diseño del modelo predictivo o del estudio de validación, las diferencias en el perfil de riesgo de los pacientes o la deficiencia real en la calidad asistencial, que puede ser justificada por motivos socioeconómicos, de protocolización, técnica quirúrgica, cuidados postoperatorios, etc. Es complicado definir cuál ha sido la causa del fallo de calibración en este estudio, aunque ya se ha mencionado el perfil de alto riesgo de los pacientes españoles como uno de los posibles motivos. Las ya conocidas debilidades de EuroSCORE II (tendencia innata a la infraestimación, falta de seguimiento de pacientes) también podrían justificarlo. Por último, el diseño del estudio de validación ha sido adecuado a nuestro entender, con un tamaño suficiente de la muestra, buena representación de los diferentes centros de toda la geografía española, así como un diseño prospectivo y una proporción de datos perdidos de cero, lo cual no parece explicar a priori los resultados descritos.

Es evidente que EuroSCORE II ha sido diseñado con algunos problemas, principalmente por la deficiencia en la recogida de datos de mortalidad a 30 y 90 días aunque, a la luz de los resultados de este trabajo, parece funcionar de forma aceptable en nuestro país (y mejor que su versión anterior), teniendo en cuenta la tendencia a la infraestimación del riesgo descrita, que debería ser corregida.

Hay diversas enseñanzas que se pueden extraer de este proyecto. Una, lejos de los datos matemáticos, es el enorme potencial de producción científica del que disponemos en nuestro país. Este proyecto ha sido generado con 3 ingredientes básicos, ilusión, esfuerzo y organización. Una llamada a la participación logró interesar a más de 30 centros que se brindaron a colaborar. Las circunstancias particulares y la disponibilidad de datos en los diferentes servicios del país produjeron que algunos de estos centros no pudiesen aportar finalmente sus datos (algunos clásicos participantes en estudios sobre calidad asistencial)³⁰, aunque la respuesta fue abrumadora y el tamaño final de la muestra, óptimo para este estudio. El esfuerzo en la recogida permanente y veraz de los datos por parte de los centros participantes, en la revisión de los datos incorrectos y en la coordinación y análisis del conjunto final han dado su fruto en un trabajo, a nuestro entender, metodológicamente correcto y adecuadamente conducido.

En segundo lugar, confirmar que la calidad de los datos recogidos en España es óptima, tal y como se deduce de una tasa de falta de datos o de datos incoherentes del 0%. El primer estudio de calidad de la SECTCV de Josa et al.³⁰ ha demostrado que los datos generados en España tienen una calidad por encima de la mayoría de países del entorno europeo, lo que es un valor añadido del estudio.

Por último, el esfuerzo del grupo coordinador del estudio ha logrado generar el primer estudio multicéntrico y prospectivo de validación a gran escala hasta la fecha, al integrar la información aportada por los grupos participantes, que es la base del estudio.

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 en este artículo no aparecen datos de pacientes.

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

Conflicto de intereses

Ninguno declarado.

Agradecimientos

A los responsables directos en cada uno de los centros, que han integrado el Grupo de Trabajo para el Proyecto de Validación de EuroSCORE II en España (por orden alfabético): Alejandro Adsuar (Sevilla), Antonino Ginel (Barcelona), Carlos Velasco (A Coruña), Elena Arnáiz (Salamanca), Félix Gómez (Pamplona), Gonzalo Aldámiz (Albacete, Madrid), José Antonio Blázquez (Madrid), José Ignacio Aramendi (Bilbao), José López (Oviedo), José Manuel Martínez Cereijo (Santiago), José María González Santos (Salamanca), Manuel Calleja (Granada), Mario Castaño (León), Marisa Cámara (Badalona), Miguel Ángel Castro (Barcelona) y Yolanda Carrascal (Valladolid). También un sincero agradecimiento a Albert Carbonell y Biomeno SL, por su labor en la adecuación de SICCS para la realización de este estudio.

Bibliografía

- Cortina J. ¿Es EuroSCORE II el nuevo patrón como modelo de riesgo en cirugía cardíaca? Uso, aplicación clínica, evaluación y consecuencias. *Cir Cardio*. 2013;20:55–8.
- Higgins TL, Estafanos FG, Loop FD, Beck GJ, Blum JM, Parandani L. Stratification of morbidity and mortality outcome by preoperative risk factors in coronary artery bypass patients. A clinical severity score. *JAMA*. 1992;267:2344–8.
- Parsonnet V, Dean D, Bernstein AD. A method of uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease. *Circulation*. 1989;79:13–12.
- Pons JM, Granados A, Espinas JA, Borrás JM, Martín I, Moreno V. Assessing open heart surgery mortality in Catalonia (Spain) through a predictive risk model. *Eur J Cardiothorac Surg*. 1997;11:415–23.
- Surgeons TSoT. Online STS Risk Calculator. 2012.
- Roques F, Nashef SA, Michel P, Gauducheau E, de Vincentiis C, Baudet E, et al. Risk factors and outcome in European cardiac surgery: Analysis of the EuroSCORE multinational database of 19030 patients. *Eur J Cardiothorac Surg*. 1999;15:816–22 [discussion 22–3].
- Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European System for Cardiac Operative Risk Evaluation (EuroSCORE). *Eur J Cardiothorac Surg*. 1999;16:9–13.
- Pitkanen O, Niskanen M, Rehnberg S, Hippelainen M, Hynynen M. Intra-institutional prediction of outcome after cardiac surgery: Comparison between a locally derived model and the EuroSCORE. *Eur J Cardiothorac Surg*. 2000;18:703–10.
- Huijskes RV, Rosseel PM, Tijssen JG. Outcome prediction in coronary artery bypass grafting and valve surgery in the Netherlands: Development of the Amphiascore and its comparison with the Euroscore. *Eur J Cardiothorac Surg*. 2003;24:741–9.
- Andrade IN, Moraes Neto FR, Oliveira JP, Silva IT, Andrade TG, Moraes CR. Assessment of the EuroSCORE as a predictor for mortality in valve cardiac surgery at the Heart Institute of Pernambuco. *Rev Bras Cir Cardiovasc*. 2010;25:11–8.
- Lafuente S, Trilla A, Bruni L, Gonzalez R, Bertran MJ, Pomar JL, et al. Validation of the EuroSCORE probabilistic model in patients undergoing coronary bypass grafting. *Rev Esp Cardiol*. 2008;61:589–94.
- Mestres CA, Castro MA, Bernabeu E, Josa M, Cartana R, Pomar JL, et al. Preoperative risk stratification in infective endocarditis. Does the EuroSCORE model work? Preliminary results. *Eur J Cardiothorac Surg*. 2007;32:281–5.
- Youn YN, Kwak YL, Yoo KJ. Can the EuroSCORE predict the early and mid-term mortality after off-pump coronary artery bypass grafting? *Ann Thorac Surg*. 2007;83:2111–7.
- Riha M, Danzmayr M, Nagele G, Mueller L, Hoefer D, Ott H, et al. Off pump coronary artery bypass grafting in EuroSCORE high and low risk patients. *Eur J Cardiothorac Surg*. 2002;21:193–8.
- Al-Ruzzeh S, Asimakopoulos G, Ambler G, Omar R, Hasan R, Fabri B, et al. Validation of four different risk stratification systems in patients undergoing off-pump coronary artery bypass surgery: A UK multicentre analysis of 2223 patients. *Heart*. 2003;89:432–5.
- Roques F, Nashef SA, Michel P, Pinna Pintor P, David M, Baudet E. Does EuroSCORE work in individual European countries? *Eur J Cardiothorac Surg*. 2000;18:27–30.
- Wang C, Yao F, Han L, Zhu J, Xu ZY. Validation of the European system for cardiac operative risk evaluation (EuroSCORE) in Chinese heart valve surgery patients. *J Heart Valve Dis*. 2010;19:21–7.
- Yap CH, Reid C, Yui M, Rowland MA, Mohajeri M, Skillington PD, et al. Validation of the EuroSCORE model in Australia. *Eur J Cardiothorac Surg*. 2006;29:441–6 [discussion 6].
- Qadir I, Salick MM, Perveen S, Sharif H. Mortality from isolated coronary bypass surgery: A comparison of the Society of Thoracic Surgeons and the EuroSCORE risk prediction algorithms. *Interact Cardiovasc Thorac Surg*. 2012;14:258–62.
- Zheng Z, Li Y, Zhang S, Hu S, Chinese CRS. The Chinese coronary artery bypass grafting registry study: How well does the EuroSCORE predict operative risk for Chinese population? *Eur J Cardiothorac Surg*. 2009;35:54–8.
- Kalavrouziotis D, Li D, Butth KJ, Legare JF. The European System for Cardiac Operative Risk Evaluation (EuroSCORE) is not appropriate for withholding surgery in high-risk patients with aortic stenosis: A retrospective cohort study. *J Cardiothorac Surg*. 2009;4:32.
- Parolari A, Pesce LL, Trezzi M, Cavallotti L, Kassem S, Loardi C, et al. EuroSCORE performance in valve surgery: A meta-analysis. *Ann Thorac Surg*. 2010;89:787–93, 93 e1–2.
- Ranucci M, Castelvécchio S, Menicanti LA, Scolletta S, Biagioli B, Giomarelli P. An adjusted EuroSCORE model for high-risk cardiac patients. *Eur J Cardiothorac Surg*. 2009;36:791–7.
- Smith CR, Leon MB, Mack MJ, Miller DC, Moses JW, Svensson LG, et al. Transcatheter versus surgical aortic-valve replacement in high-risk patients. *N Engl J Med*. 2011;364:2187–98.
- Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. EuroSCORE II. *Eur J Cardiothorac Surg*. 2012;41:734–44 [discussion 44–5].
- Sergeant P, Meuris B, Pettinari M. EuroSCORE II, illum qui est gravitates magni observe. *Eur J Cardiothorac Surg*. 2012;41:729–31.
- Silva F CM, Reguillo F, Cobiella J, Villagrán E, Montes L, Garcés Z, et al. Validación del EuroSCORE II: ¿funciona en nuestro medio? *Cir Cardio*. 2013;20:59–64.
- Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143:29–36.
- Cortina Romero JM. Complexity and severity scores in cardiac surgery. Uses and limitations. *Rev Esp Cardiol*. 2005;58:473–6.
- Josa M, Cortina J, Mestres C, Pereda D, Walton P, Kinsman R. Primer informe del proyecto español de calidad de cirugía cardiovascular del adulto. Barcelona: Publicaciones Permanyer; 2013.
- Nashef SA, Roques F, Hammill BG, Peterson ED, Michel P, Grover FL, et al. Validation of European System for Cardiac Operative Risk Evaluation (EuroSCORE) in North American cardiac surgery. *Eur J Cardiothorac Surg*. 2002;22:101–5.
- 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.
- Borracci 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.
- Paparella D, Guida P, di Eusanio G, Caparrotti S, Gregorini R, Cassese M, et al. Risk stratification for in-hospital mortality after cardiac surgery: External validation of EuroSCORE II in a prospective regional registry. *Eur J Cardiothorac Surg*. 2014. Jan 30 [Epub ahead of print]. Disponible en: <http://ejcts.oxfordjournals.org/content/early/2014/01/30/ejcts.ezt657.full.pdf+html>.
- Hickey GL, Bridgewater B. How well calibrated is EuroSCORE II? *Eur J Cardiothorac Surg*. 2013;43:208.
- Sharples LD, Nashef SA, Euro SPG. Reply to Hickey and Bridgewater. *Eur J Cardiothorac Surg*. 2013;43:208–9.
- Steyerberg E. En: Gail M, Krickeberg K, Samet J, Tsiatis A, Wong W, editores. *Clinical prediction models*. 1st ed New York: Springer Science + Business Media, LLC; 2010.



BIOMED



unidix

Especialistas en cirugía cardiovascular

desde 1977 al cuidado de tu salud



91 803 28 02



info@biomed.es