

Editorial

¿Es EuroSCORE II el nuevo patrón como modelo de riesgo en cirugía cardíaca? Uso, aplicación clínica, evaluación y consecuencias

Is EuroSCORE II the new standard as risk model in cardiac surgery? Use, clinical application, evaluation and consequences

José M.^a Cortina*

Servicio de Cirugía Cardíaca, Hospital 12 de Octubre, Madrid, España

La cirugía cardíaca es una especialidad pionera en el uso sistemático de sistemas de predicción de riesgo

Puede decirse, sin temor a equivocarse, que los cirujanos cardíacos hemos introducido la dialéctica de la evaluación del riesgo quirúrgico de nuestros pacientes como una práctica habitual, cotidiana y sistemática. La introducción de esta metodología ha instalado en la actualidad, una cultura de la evaluación objetiva del riesgo en la mayor parte de los grupos quirúrgicos. Esta tendencia entronca con otra ya existente en la especialidad que no es otra que la de analizar y publicar los resultados de los procedimientos que realizamos. Como consecuencia de todo esto, podemos decir que, posiblemente, la práctica de la cirugía cardíaca sea uno de los campos más analizados y escrutados, no solo por los cirujanos cardíacos, sino por cardiólogos, especialistas en salud pública, administradores, etc. Curiosamente, procedimientos alternativos a la cirugía como el intervencionismo percutáneo también disponen de herramientas equivalentes en forma de modelos predictivos, pero no se han instalado en la práctica clínica diaria.

En este sentido, es interesante reseñar las referencias que se hacen en relación con el uso de sistemas predictivos en las guías de práctica clínica actuales. Las guías norteamericanas de 2011 sobre intervencionismo percutáneo¹ y sobre cirugía de revascularización coronaria² establecen como recomendación IIa, nivel de evidencia B que: “El cálculo de *score* de la Sociedad de Cirujanos Torácicos (STS) y el *score* SYNTAX son razonables en pacientes con enfermedad de tronco izquierdo no protegido y en enfermedad coronaria compleja”. El *score* o estimación de riesgo de la STS es una predicción del riesgo quirúrgico. No se considera recomendar el uso de sistemas predictivos del riesgo de procedimientos percutáneos aunque sí que están disponibles en la literatura.

Las guías europeas introducen unas recomendaciones más ponderadas aconsejando el uso de *scores* quirúrgicos para cirugía coronaria como STS y EuroSCORE y también específicos para intervencionismo percutáneo como Mayo Clinic Score³ y el *score* del National Cardiovascular Data Registry⁴. Estas recomendaciones, que son razonables, no han conseguido la generalización de modelos aplicables a procedimientos percutáneos que también implican riesgo inmediato y tardío.

En mi opinión, cualquier procedimiento terapéutico, cardiológico o de cualquier índole debería estar sometido a una metodología de estimación de su riesgo adecuadamente validada, que permita, fun-

damentalmente, comparar los resultados obtenidos con los esperables en un entorno de análisis de la calidad real.

En la actualidad, EuroSCORE es el modelo predictivo más usado en nuestro entorno y posiblemente en el ámbito mundial

El comienzo de la popularización del uso de modelos de riesgo en cirugía cardíaca hay que atribuirlo a Victor Parsonnet⁵. Este sistema fue ampliamente utilizado durante toda la década de los noventa. El proyecto del EuroSCORE se originó en 1993 como respuesta europea al Parsonnet. La recogida de datos ocurrió en otoño de 1995 publicándose el modelo en 1999⁶. Fue un proyecto multiinstitucional que implicó la participación de 128 grupos quirúrgicos de 8 Estados europeos⁷. España participó con 25 grupos y 2.444 pacientes. De él se derivó un modelo que ha sido usado desde principios de la década anterior hasta la actualidad. Su éxito ha sido incuestionable. Una búsqueda en PubMed con la palabra EuroSCORE devuelve un resultado de 1.344 referencias sin filtrar. Con el mismo criterio, Parsonnet *score* resulta con 161 y STS *score* con 267.

El desarrollo del modelo de riesgo de la STS ha sido paralelo al crecimiento de su base de datos, y ha seguido una dinámica distinta. Esta ha sido en el sentido de la actualización progresiva de los modelos predictivos con el tiempo y del desarrollo de modelos específicos para cirugía coronaria y valvular independientes. Su uso es discretamente más complejo que el EuroSCORE.

El éxito del EuroSCORE se basa fundamentalmente en su simplicidad de uso, tanto del modelo aditivo como logístico, pero también en que se percibe como un modelo de todos debido a la importante participación multiinstitucional. El EuroSCORE I, después de muchos años de uso, perdió capacidad de calibración, sobrestimando en general los riesgos⁸, debido a que en realidad, era representativo de una “foto” de la práctica de la cirugía cardíaca en 1995. No obstante, su capacidad de discriminación seguía aceptablemente preservada. En función de esta obsolescencia, los autores del proyecto inicial gestaron una actualización del modelo persiguiendo la generación de un EuroSCORE II. Esta vez, el proyecto se diseñó mediante la participación de los centros vía web, implicando la participación de 154 grupos de 43 países en el ámbito mundial agrupando información útil para el análisis de 22.381 pacientes. Es llamativamente reseñable que España sea el país que más grupos aporta, 19, seguida de Francia

*Autor para correspondencia. Servicio de Cirugía Cardíaca, Hospital 12 de Octubre, Avda. Córdoba s/n - 28041 Madrid, España. Tel. +34639958078
Correo electrónico: josecortina@sctcv.es

con 16, Italia con 15 y Reino Unido con 12. La recogida de datos se realizó entre mayo y julio de 2010, siendo presentados los datos preliminares en Lisboa en el congreso de la European Association for Cardio-Thoracic Surgery en 2011 y publicándose en febrero de 2012⁹. Aunque su publicación es reciente y ya ha generado comentarios editoriales y evaluaciones en todos los sentidos, todo hace pensar que su aplicación se generalizará por el mismo tipo de razones que el modelo inicial.

El modelo predictivo EuroSCORE permite distintos tipos de uso. En función de cada uno de ellos presenta virtudes y limitaciones determinadas

En una extraordinariamente crítica editorial de P. Sergeant¹⁰ publicada simultáneamente con el EuroSCORE II, en la que, además de otras consideraciones, llega a cuestionar la ética profesional de los grupos participantes, el autor considera que: "Este modelo no debería ser usado, como tal, para la monitorización o comparación de la calidad, terapias alternativas o consentimiento informado y mucho menos para la información pública de las actuaciones médicas". Es evidente que este tipo de usos que este autor desaconseja absolutamente, sobre la base de defectos metodológicos de su desarrollo, son suficientemente serios e importantes como para requerir una herramienta que sea 100% fiable para no sufrir consecuencias desagradables. Sin embargo, también es cierto que es imposible impedir que todo esto ocurra, máxime cuando el desarrollo del modelo EuroSCORE se ha basado en una colaboración tan extendida.

En relación con el uso para la monitorización y comparación de la calidad, hemos de decir que esta aplicación es insoslayable. Todos los interesados en la cuestión de la calidad, que es toda la sociedad, persiguen herramientas que sean canon de comparación. Y el EuroSCORE II es una nueva "foto" de la práctica de la cirugía cardíaca en un entorno muy amplio como para ser desechado. Las objeciones metodológicas, aun siendo serias, no impedirán este tipo de uso.

La aplicación del modelo para estimaciones de riesgo individuales, para el proceso de decisión terapéutica, tiene más limitaciones. Tiene solo un valor orientativo dado que el número de variables que utiliza es limitado y es fácil que un paciente determinado presente un factor de riesgo importante no incluido en el modelo. Se trata de una estimación probabilística que sería próxima a la verdad si se interviesen cientos de pacientes con esas mismas características de ese enfermo dado. Por esta misma razón, el uso para evaluar el riesgo utilizando otro tipo de tecnología como implantación de válvulas transcáteter también sería inapropiado. Para estimaciones individuales, es más próximo a la verdad el uso de sistemas que incluyan más factores de riesgo o el uso de sistemas específicos para un procedimiento dado.

Lo mismo que en el ámbito de la calidad, estas razones no impedirán que se use extensamente a título individual.

Un uso más aséptico y que tiene menos limitaciones es como medidor objetivo del *case-mix* o complejidad de las poblaciones que atendemos. Un sistema de puntuación simple permite cuantificar cómo es de grave nuestra población en términos objetivos. Esta aplicación es mucho más sencilla e intuitiva con *scores* tipo "estáticos" como EuroSCORE. Con modelos de riesgo tipo STS que se recalibran con más frecuencia es más difícil la evaluación en el tiempo de la gravedad de las poblaciones.

El uso de un modelo de riesgo como EuroSCORE bajo unas circunstancias determinadas no implica necesariamente un ensayo de validación correcto

El trabajo de Silva et al.¹¹ "Validación del EuroSCORE II: ¿funciona en nuestro medio?" suscita, una vez más, interrogantes sobre la

cuestión de los modelos predictivos, que nos pueden enfrentar problemas serios en la manera de realizar y evaluar nuestra práctica diaria. El citado artículo concluye acerca del modelo y en su experiencia en una aceptable capacidad discriminatoria, con un área bajo la curva ROC de 0,78 y una pobre calibración resumida en una mortalidad observada de 6,3 frente a una esperada de EuroSCORE II de 3,46 intervalo, de confianza 95% (3,3-3,6) resultando en un test de Hosmer-Lemeshow insatisfactorio con $\chi^2 = 58,19$; $p < 0,0001$. Dado que se acepta, que la validación de un modelo predictivo requiere una buena discriminación y una buena calibración, al ser esta última insatisfactoria se ha de inferir una respuesta negativa a la pregunta que plantea en el título del ensayo y por ende una invalidación del mismo.

Sin embargo, antes de dar por buena esa aseveración, deberíamos plantearnos, por un lado, cuestiones conceptuales y metodológicas a la hora de la investigación del modelo y, por otro, las inferencias que se podrían derivar de nuestras conclusiones.

Desde el punto de vista conceptual, deberíamos preguntarnos: ¿qué significa validación de un sistema predictivo?

Desde un análisis estadístico, la validación externa de un modelo pronóstico significa documentar que funciona correctamente sobre una población distinta de la que se derivó. Posiblemente, el concepto de validación de estas herramientas se esté usando de forma incorrecta a la hora de publicar el uso de este tipo de sistemas. Es equívoco pretender validar un modelo simplemente por el hecho de usarlo en unas circunstancias determinadas como por ejemplo utilizarlo en un grupo de patología determinado. La adecuada validación de un modelo probabilístico presenta una serie de condiciones que no siempre se dan cuando se aplican estos modelos¹².

La primera condición para pretender la validación de un modelo es la del tamaño muestral. Harrel¹³ considera que para estudios en que la variable de salida es binaria, vivo o muerto, se requieren al menos cien casos con este *outcome* para la validación del modelo. Condición esta que se cumple en el estudio que comentamos. Además del tamaño, se requiere la aplicación estricta del *score* a todos y cada uno de los pacientes correspondientes, sin asunciones sobre factores perdidos, y lo que es más importante, sin pérdidas en el evento que sea la variable dependiente.

Idealmente, la correcta validación de un modelo de este tipo debería realizarse de forma prospectiva y multiinstitucional. La aplicación de modelos predictivos en una sola institución con intención de validarlos plantea problemas cuando se concluye en la no validación.

Metodológicamente, la pregunta es obvia: ¿hemos aplicado correctamente el modelo?

Una condición evidente para sustentar las conclusiones que se publican es conseguir la certeza de una correcta aplicación. Un nuevo modelo como EuroSCORE II, que ha recodificado variables y ha introducido alguna nueva, requeriría una aplicación prospectiva para evitar demasiadas inferencias en la recodificación y una recogida de información ajustada a las definiciones. El trabajo que comentamos aplica el nuevo *score* a una serie retrospectiva respecto a la publicación del modelo. Esto tiene 2 problemas posibles. El primero en forma de posibles errores en el cálculo del *score* y el segundo que se ocupa de una serie pasada respecto a un modelo actual. Llama la atención en el trabajo de Silva la escasa proporción de cirugía no electiva (11,8%) frente a 23,3% de la serie de EuroSCORE. Este dato solo puede significar 2 cosas, o una óptima gestión de la programación quirúrgica o un error en la adscripción de las prioridades que se derivaría en un infracálculo del *score*. Por otro lado, su trabajo obtiene un EuroSCORE I de 9,1, uno de los más altos de la literatura, frente a

un EuroSCORE II de 3,46. Sin embargo, en el estudio de Di Dedda¹⁴ estas cifras son 7,33 y 3,1, respectivamente. Aunque no quepa esperar una correlación lineal entre los 2 *scores*, llama la atención la discordancia.

Por todo ello, antes de emitir conclusiones y plantear hipótesis sobre la validez de un modelo dado es necesaria la certeza sobre la correcta aplicación.

Si el modelo está correctamente aplicado, y nuestra conclusión es que no se valida, se plantean varias explicaciones que no son excluyentes entre ellas. Estas serían: una metodología estadística de validación inadecuada, una real mala calibración del nuevo modelo y, por último, una diferencia real y significativa frente a los resultados esperables según el modelo.

Existe consenso en la literatura sobre que significa la validación externa de un modelo pronóstico¹³

Actualmente, la gran mayoría de los modelos pronósticos en cirugía cardíaca utilizan la regresión logística como herramienta de desarrollo, ya que se ajusta bien al modelo de variable de salida binaria.

Para estos modelos, y con las condiciones previas que hemos comentado de correcta aplicación del mismo, la validación externa requiere una buena discriminación y una buena calibración o fiabilidad.

Una correcta calibración o fiabilidad debe dar idea de 2 parámetros: por un lado, el promedio de la estimación debe ser estadísticamente correcto y, por otro, debe predecir bien en todos los rangos, desde bajo a alto riesgo. Clásicamente, se acepta el test de Hosmer-Lemeshow para este análisis.

Una buena discriminación debe dar idea de la capacidad del modelo de separar correctamente las respuestas del modelo que se trate. Da idea de la correlación entre la probabilidad predicha del evento con la ocurrencia real del mismo. Para este parámetro, se utiliza el área bajo la curva ROC.

Los trabajos que se publican intentando validar el EuroSCORE II están usando esta metodología pensamos que correctamente.

Con la información actualmente disponible en estudios multiinstitucionales se puede decir que el nuevo modelo ha mejorado discretamente su capacidad de discriminación y substancialmente su calibración

En general, parece que existe consenso en que la capacidad de discriminación del nuevo modelo mejora discretamente la del modelo previo. Sirvan de ejemplo las siguientes: en el trabajo que estamos comentando, el área bajo la curva pasa de 0,75 a 0,78; en la publicación de la serie del EuroSCORE II pasa de 0,79 a 0,81⁹; en la serie multiinstitucional de Gran Bretaña e Irlanda, 0,81¹⁵ y en la serie del Policlínico San Donato pasa de 0,79 a 0,81¹⁴.

Sobre la calibración existen discrepancias en la literatura que ahora comentaremos. La disquisición sobre la correcta calibración es la más sensible. Por regla general, los hallazgos son de 3 tipos: riesgo esperado bajo frente al observado, riesgo esperado alto frente al observado y buena concordancia entre ambos. En los 2 primeros casos casi siempre se concluye en mala calibración, pero en el primero nunca se invoca a otro tipo de posibles razones. La mayoría de los trabajos disponibles entran en estas 3 categorías y la decisión se plantea en a cuál dar más crédito. Por otro lado, por definición, una buena calibración implica que el modelo estime bien el resultado en todos los niveles de riesgo. Normalmente, se dividen los estratos en deciles para analizar el funcionamiento por segmentos de riesgo. El trabajo de Silva no estratifica por grupos de riesgo y sí por grupos de patología, encontrando que el nuevo modelo infraestima el riesgo en todos ellos y esta es su hipótesis para una mala calibración. Este

hallazgo es similar en la serie de Zhang¹⁶ donde encuentran infraestimación del riesgo en pacientes valvulares múltiples. En el mismo sentido es el estudio de Kunt¹⁷ con una diferencia muy elevada entre mortalidad observada y riesgo esperado (7,9 vs. 1,7).

Sin embargo, en la serie multiinstitucional ya citada¹⁵ observan que la calibración es buena en promedio (3,1 observada frente a 3,4 esperada por EuroSCORE II) pero el modelo falla en el test de Hosmer-Lemeshow, sobre todo debido a una sobrestimación en los pacientes de más bajo y más alto riesgo. Así, en cirugía coronaria aislada, observan una mortalidad de 1,5 frente a una esperada de 2,1. La observación de mala estimación en pacientes de alto riesgo también es compartida por Howell¹⁸ y por Barili; en este caso, sobrestimación en riesgos a partir del 30%¹⁹.

Por otro lado, Di Dedda¹⁴ encuentra buena calibración tanto en promedio como en los quintiles de riesgo que analiza.

En general, casi todos los autores están de acuerdo en que la calibración del nuevo modelo es claramente mejor que el previo. Las observaciones sobre mala calibración son absolutamente discordantes, y se refieren a infraestimaciones del riesgo, por un lado, o a sobrestimaciones en los segmentos de riesgo bajo y alto.

Tratar de explicar todo esto a la vez es complejo y sería necesario entrar en la dialéctica del peso asignado a las variables incluidas en el modelo y de la exclusión de variables que pudieran ser importantes. Creo que esta última hipótesis se puede obviar, pues el modelo presenta suficientes variables de gravedad de la enfermedad y de patología asociada como para funcionar adecuadamente en las poblaciones globales que se intervienen. No así, y ya se ha comentado, en la estimación de grupos quirúrgicos específicos.

Solo queda considerar si el peso asignado a las variables en el modelo es el adecuado a nuestra situación concreta. Para considerar la hipótesis de que la calibración es inadecuada en nuestro medio, habría que aceptar que determinadas variables están infraponderadas. No tenemos medios con la información disponible para responder a esta cuestión. Solo se puede concluir que si fuera inadecuado el peso asignado a las variables, para nuestro medio, la participación española no pudo impactar suficientemente con su aportación al modelo final. Hay que recordar que participaron 19 grupos de un total de 154 (12,3%) con una contribución de 1.806 casos sobre 24.385 totales (7,4%) (comunicación personal).

Puede ocurrir que una supuesta mala calibración, con mortalidades esperadas inferiores a las observadas en realidad pueda estar evidenciando diferencias reales y significativas de resultados frente a los esperables, según el modelo predictivo

Esta posibilidad es la más desfavorable y normalmente no se contempla cuando se realizan ensayos de aplicación de los modelos de riesgo. Sin embargo, es obligatorio plantearse al menos desde el punto de vista teórico. El uso de un modelo de riesgo u otro implica conceptualmente un canon frente al que nos evaluamos. Hablando del EuroSCORE II, este representaría lo que cabe esperar en cuanto a resultados, de la práctica actual de la cirugía cardíaca en un ámbito internacional. Los autores construyeron el modelo con la información aportada voluntariamente por un número elevado de centros vía web. No hubo auditoría ni control de la información remitida salvo la que en el ámbito local establecieran los participantes. Por ello, solo se puede confiar en todos y cada uno de ellos en la calidad de la información suministrada respecto a todas las variables, siendo la más crítica la de salida, es decir el evento muerte. Si este hubiera sido infrarreportado, el modelo estaría mal calibrado por infraestimador.

Contra esta posibilidad, tenemos las experiencias de mala calibración por sobrestimación que hemos comentado antes.

No es descartable la posibilidad de que las mortalidades observadas superiores a las esperadas estén traduciendo una diferencia de calidad prestada. Sin embargo, la conclusión de que uno u otro mo-

delo no funciona en nuestro medio porque infraestima los resultados también puede ser pertinente.

Es conveniente recordar aquí que la teoría al uso sobre la evaluación de la calidad infiere que si hacemos que las poblaciones tratadas sean comparables mediante ajustes con modelos de riesgo, las diferencias significativas en los resultados se deberían a diferencias de calidad asistencial. La evaluación de la calidad puede ser hecha desde múltiples perspectivas pero la medida de los resultados en términos de mortalidad ajustada en cirugía cardíaca es un estándar actualmente.

La consecución de unos determinados resultados tiene que ver con el tipo de poblaciones que atendemos, y sabemos que el perfil de los pacientes operados de cirugía cardíaca en España es más grave que en la mayoría de los países europeos²⁰. Pero también, y no menos importante, los resultados dependen de cómo están diseñados y realizados los procesos y de la estructura en términos amplios donde se ejecutan²¹. En general, se tiende a pensar que el resultado quirúrgico depende exclusivamente de lo que ocurra en quirófano, pero hay un sinnúmero de dimensiones que no se ponderan, no se evalúan y sobre las que los cirujanos a veces tienen algún control y otras ninguno. Por supuesto, todas estas variaciones y diferencias estructurales no están ajustadas de la misma manera que se ajustan las variables clínicas de los pacientes y, por consiguiente, solo se pueden hacer hipótesis sobre su impacto en los resultados.

Con independencia de los problemas expuestos, EuroSCORE II se va a convertir en la referencia mundial como modelo de riesgo en cirugía cardíaca como ya lo es su antecedente. Todavía existen dudas sobre la óptima calibración del modelo para el estado del arte actual de la cirugía cardíaca, pero ha mejorado su capacidad predictiva respecto a la versión I tanto en discriminación como en calibración.

Los ensayos para su validación externa requieren una aplicación estricta del modelo, ateniéndose rigurosamente a la definición de las variables. Idealmente, han de realizarse de forma prospectiva, con participación multiinstitucional y aplicándolo al universo de pacientes sobre el que está derivado. La aplicación del modelo sobre grupos específicos de patología no debe considerarse conceptualmente una validación del modelo.

Hay que recordar que la primera utilidad de este tipo de herramientas es la del control de calidad, y que, desviaciones significativas de los parámetros esperados pueden estar evidenciando problemas distintos a los puramente referidos a la naturaleza del modelo.

Bibliografía

1. Levine GN, Bates ER, Blankenship JC, Bailey SR, Bittl JA, Cercek B, et al. 2011 ACCF/AHA/SCAI Guideline for Percutaneous Coronary Intervention. A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines and the Society for Cardiovascular Angiography and Interventions. *J Am Coll Cardio*. 2011;58:e44-122.
2. Hillis LD, Smith PK, Anderson JL, Bittl JA, Bridges CR, Byrne JG, et al. 2011 ACCF/AHA Guideline for Coronary Artery Bypass Graft Surgery. A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. Developed in collaboration with the American Association for Thoracic. *J Am Coll Cardio*. 2011;58:e123-210.
3. Singh M, Rihal CS, Lennon RJ, Spertus J, Rumsfeld JS, Holmes DR Jr. Bedside estimation of risk from percutaneous coronary intervention: the new Mayo Clinic risk scores. *Mayo Clinic proceedings Mayo Clinic*. 2007;82:701-8.
4. Peterson ED, Dai D, DeLong ER, Brennan JM, Singh M, Rao SV, et al. Contemporary mortality risk prediction for percutaneous coronary intervention: results from 588,398 procedures in the National Cardiovascular Data Registry. *J Am Coll Cardio*. 2010;55:1923-32.
5. 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.
6. Nashef S, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). *European journal of cardio-thoracic surgery: official journal of the European Association for Cardio-thoracic Surgery*. 1999;16:9-13.
7. 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. *European journal of cardio-thoracic surgery: official journal of the European Association for Cardio-thoracic Surgery*. 1999;15:816-22; discussion 822-3.
8. Siregar S, Groenewold RHH, de Heer F, Bots ML, van der Graaf Y, van Herwerden LA. Performance of the original EuroSCORE. *European journal of cardio-thoracic surgery: official journal of the European Association for Cardio-thoracic Surgery*. 2012;41:746-54.
9. Nashef S, Roques F, Sharples L, Nilsson J, Smith C, Goldstone AR, et al. EuroSCORE II. *European journal of cardio-thoracic surgery*. 2012;41:1-12.
10. Sergeant P, Meuris B, Pettinari M. EuroSCORE II, illum qui est gravitates magni observe*. *European Journal of Cardio-Thoracic Surgery*. 2012;41:1-3.
11. Silva J, Carnero M, Reguillo F, Cobiella J, Villagrán E, Montes L, et al. Validación del EuroSCORE: ¿Funciona en nuestro medio? *Cirugía Cardiovascular*. 2013;20:2.
12. Altman DG, Royston P. What do we mean by validating a prognostic model? *Statistics in medicine*. 2000;19:453-73.
13. Harrell FE. *Regression Modeling Strategies*. Nueva York: Springer-Verlag; 2001.
14. Di Dedda U, Pelissero G, Agnelli B, De Vincentiis C, Castelvécchio S, Ranucci M. Accuracy, calibration and clinical performance of the new EuroSCORE II risk stratification system. *European journal of cardio-thoracic surgery: official journal of the European Association for Cardio-thoracic Surgery*. 2013;43:27-32.
15. Grant SW, Hickey GL, Dimarakis I, Trivedi U, Bryan A, Treasure T, et al. How does EuroSCORE II perform in UK cardiac surgery: an analysis of 23 740 patients from the Society for Cardiothoracic Surgery in Great Britain and Ireland National Database. *Heart (British Cardiac Society)*. 2012;98:1568-72.
16. Zhang G-X, Wang C, Wang L, Lu FL, Li BL, Han L, et al. Validation of EuroSCORE II in Chinese Patients Undergoing Heart Valve Surgery. *Heart, Lung Circ*. Publicado primero online el 31 Ene 2013. doi:10.1016/j.hlc.2012.12.012
17. Kunt AG, Kurtcephe M, Hidiröglü M, Cetin L, Kucuker A, Bakuy V, et al. Comparison of original EuroSCORE, EuroSCORE II and STS risk models in a Turkish cardiac surgical cohort. *Interactive cardiovascular and thoracic surgery*. 2013;16:625-9.
18. Howell NJ, Head SJ, Freemantle N, van der Meulen TA, Senanayake E, Menon A, et al. The new EuroSCORE II does not improve prediction of mortality in high-risk patients undergoing cardiac surgery: a collaborative analysis of two European centres. *European journal of cardio-thoracic surgery: official journal of the European Association for Cardio-thoracic Surgery*. 2013.
19. Barili F, Pacini D, Capo A, Rasovic O, Grossi C, Alamanni F, et al. Does EuroSCORE II perform better than its original versions? A multicentre validation study. *European Heart Journal*. 2013;34:22-9.
20. Bridgewater B, Gummert J. Fourth EACTS Adult Cardiac Surgical Database Report 2010. Londres: Dendrite Clinical Systems Ltd; 2010.
21. Shroyer AL, London MJ, Villanueva CB, Sethi GK, Marshall GMT. Relationships Between Patient-Related Factors, Processes, Structures, and Outcomes of Cardiac Surgical Care. *Conceptual Models*. *Medical Care*. 1995;33:26-34.



BIOMED



unidix

Especialistas en cirugía cardiovascular

desde 1977 al cuidado de tu salud



91 803 28 02



info@biomed.es