

Desarrollo, recalibración y validación del ArgenSCORE en el reemplazo valvular aórtico. ArgenSCORE ajustado al centro en base a un análisis del registro multicéntrico CONAREC XVI: Realidad de la cirugía cardíaca en Argentina: Mundo real

Development, recalibration and validation of ArgenSCORE in aortic valve replacement

Victorio Carlos Carosella¹

RESUMEN

El ArgenSCORE, como modelo de predicción de riesgo en cirugía cardíaca, tiene una versión I de desarrollo y una versión II de recalibración del modelo, con el que se obtiene una sustancial reducción de la mortalidad esperada. El objetivo de este estudio fue evaluar una metodología de ajuste utilizando estas escalas, que permita estimar el riesgo operatorio real en un centro quirúrgico, considerando su perfil de riesgo preoperatorio y su mortalidad observada.

Se analizaron 2548 pacientes de 44 centros de un registro prospectivo y multicéntrico en cirugía cardíaca, CONAREC XVI. Para definir la complejidad del centro se calculó el valor absoluto medio del ArgenSCORE (VAMA) de cada institución. Se evaluó el rendimiento de ambas versiones I y II, de acuerdo con la relación entre el VAMA y la mortalidad observada (VAMA/MO) para cada centro.

La mortalidad intrahospitalaria fue 7,69%. Identificamos como punto de corte de la relación VAMA/MO un valor de 4,54 para utilizar la versión I o II del ArgenSCORE. La versión I mostró mejor capacidad para asignar el riesgo en los centros con una relación VAMA/MO inferior a 4,54 y la versión II en instituciones con una relación VAMA/MO igual o superior a 4,54.

La relación VAMA/MO permite ajustar la estimación de riesgo a la realidad de cada centro logrando una mejor relación mortalidad esperada / observada. En instituciones con relación VAMA/MO < 4,54 es recomendable utilizar el ArgenSCORE I y en centros con relación $\geq 4,54$ el ArgenSCORE II.

Palabras claves: cirugía cardiovascular, mortalidad, evaluación de riesgo, factores de riesgo, riesgo ajustado.

ABSTRACT

ArgenSCORE, as a risk stratification system for cardiac surgery, has a developmental version I and a recalibration version II of the model, with which a substantial reduction in expected mortality is obtained. The objective of this study was to evaluate an adjustment methodology using these scales, which allows to estimate the actual operative risk in a surgical center, considering its preoperative risk profile and its observed mortality.

We analyzed 2548 patients from 44 centers of a prospective and multicentric registry in cardiac surgery, CONAREC XVI. To define the complexity of the center, the mean absolute value of ArgenSCORE (MAVA) of each institution was calculated. The performance of both versions I and II was evaluated according to the relationship between VAMA and observed mortality (MAVA / OM) for each center.

In-hospital mortality was 7.69%. We identified as a cut-off point for the MAVA / OM relation a value of 4.54 to use version I or II of ArgenSCORE. Version I showed a better ability to assign risk in centers with a MAVA / OM ratio < 4.54 and version II in institutions with a MAVA / OM ratio ≥ 4.54 .

The MAVA / OM relationship allows to adjust the risk estimate to the reality of each center achieving a better expected / observed mortality ratio. In institutions with relation to MAVA / OM < 4.54 it is advisable to use ArgenSCORE I and in centers with relation ≥ 4.54 ArgenSCORE II.

Key words: cardiovascular surgery, mortality, risk assessment, risk factors, adjusted risk.

Revista Argentina de Cardioangiología Intervencionista 2017;8(3):125-130

Diversas publicaciones han demostrado que la aplicación de los modelos de predicción de riesgo preoperatorio en cirugía cardíaca en poblaciones diferentes de aquellas en las cuales fueron elaborados provoca una pérdida de sus rendimientos¹⁻³. Esta limitación puede es-

tar relacionada con el hecho de que existen diferencias regionales en las características de las poblaciones, así como en las tomas de decisiones y en los resultados de los procedimientos quirúrgicos²⁻⁶. En particular, estas diferencias podrían ser de relevancia clínica cuando se comparan poblaciones de Latinoamérica con las de Estados Unidos o de Europa, donde los modelos de riesgo comúnmente utilizados fueron desarrollados en décadas pasadas.

Por otra parte, para que un modelo de predicción de riesgo mantenga confiabilidad y fortaleza se requieren evaluaciones periódicas en su rendimiento mediante

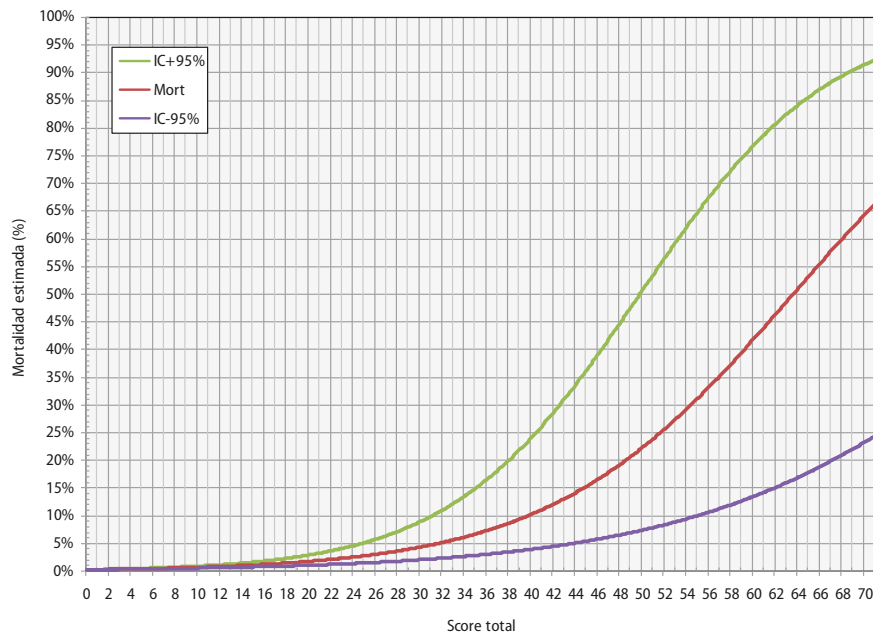
1. Consultor en Bioestadística. Servicio de Cirugía Cardiovascular. Sanatorio Finocchio, Avenida Córdoba 2678, (1187) CABA, Argentina.

✉ Correspondencia: drarosella@hotmail.com

Los autores no declaran conflictos de intereses

Recibido: 16/05/2017 | Aceptado: 30/08/2017

Factor de riesgo	Score
60-69 años	7,0
70-79 años	11,0
80 o más años	21,5
Mujer	5,0
Insulinodependencia	11,0
Insuficiencia renal	6,0
Vasculopatía periférica	6,5
Reoperación	7,5
Cirugía de urgencia	8,5
Cirugía de emergencia	14,5
Cirugía de salvataje	32,5
BCIA preoperatorio	6,0
Reemplazo valvular aórtico	2,5
Reemplazo valvular mitral	15,5
Plástica valvular aórtica	5,5
Plástica valvular mitral	5,5
Reemplazo aorta torácica	13,0
Disección aórtica aguda	5,5
Transplante cardíaco	15,0
Cirugía combinada	5,0
Enfermedad de 1 vaso	4,0
Enfermedad de 2 vasos	7,5
Enfermedad de 3 vasos	14,0
Función VI moderada	2,5
Función VI severa	9,0
Score total	



BCIA: Balón de contrapulsación intraaórtico, VI: Ventriculo izquierdo, Mort: Mortalidad predicha, IC: Intervalo de confianza

Figura 1.

procesos de validación tanto en poblaciones diferentes de las del desarrollo como en poblaciones más contemporáneas (validación externa y temporal)⁷⁻¹⁰.

DESARROLLO Y RECALIBRACIÓN DEL ARGENSCORE

Se incluyeron en forma consecutiva 5268 pacientes adultos que fueron intervenidos con algún procedimiento quirúrgico cardíaco e incluidos en una base de datos prospectiva, monitorizada y auditada, desde junio de 1994 hasta diciembre de 2009. Nuestra base de datos se desarrolló a semejanza de la base de la *Society of Thoracic Surgeons* (STS)¹¹ y, por lo tanto, las variables de riesgo y los eventos se definieron según la STS (<http://www.sts.org>). El ArgenSCORE es un modelo de riesgo simple, aditivo y con representación gráfica, desarrollado mediante un análisis de 2903 pacientes intervenidos con cirugía cardíaca en forma consecutiva en el Instituto de Cardiología del Hospital Español de Buenos Aires, desde junio de 1994 hasta diciembre de 1999.

La metodología del desarrollo y recalibración del modelo se han publicado en detalle^{12,13}. Se consideraron 49 variables del registro preoperatorio en un análisis univariado, estimándose las diferencias de proporciones mediante chi cuadrado de Pearson o la prueba exacta de Fisher, según el caso. Las variables continuas se transformaron en categóricas mediante puntos de corte según la bibliografía¹⁴, en tanto que las variables categóricas se expresaron en porcentajes y las variables continuas como media \pm desviación estándar. Las variables preoperatorias se ingresaron a un modelo de regresión logística múltiple cuando presentaron significación estadística en

el análisis univariado o bien siguiendo un criterio de importancia clínica. Posteriormente se modificó el ingreso de las variables hasta lograr el modelo con mejor ajuste. Con los valores obtenidos de constantes y de coeficientes de las variables de la regresión logística múltiple se calcula la mortalidad predicha sumando los coeficientes positivos (o presentes) y la constante de la regresión; a este valor se le calcula el *logit* para estimar la mortalidad predicha. Se identificaron 18 variables predictoras independientes de mortalidad intrahospitalaria^{12,13}. Se desarrolló además un método gráfico que permite estimar el riesgo de muerte en forma simple y que sólo requiere una grilla preimpresa. A cada variable incluida en el análisis se le asigna un puntaje, el cual resulta de multiplicar por 10 el coeficiente de la regresión logística; el riesgo estimado corresponde a la suma de los puntajes individuales de cada una de las variables positivas detectadas en el paciente. Finalmente, se graficó una curva de distribución que permitió correlacionar los valores absolutos del puntaje (*score*) con los riesgos predichos de la regresión logística múltiple. La evaluación inicial del rendimiento del modelo se realizó mediante una validación interna prospectiva desde enero de 2000 a junio de 2001 en 708 pacientes operados en el mismo centro, observándose un área bajo la curva *receiver operating characteristic* (ROC)¹⁵ de 0,77 (IC95%: 0,74-0,80). Posteriormente se realizó una primera validación prospectiva, externa y temporal del modelo en 1087 pacientes operados en tres centros de Buenos Aires entre febrero de 2000 y enero de 2007. Aunque el modelo demostró un buen poder de discriminación, área bajo la curva ROC de 0,81 (IC95%: 0,75-0,87), se observó una mala calibración debido a una mortalidad observada menor

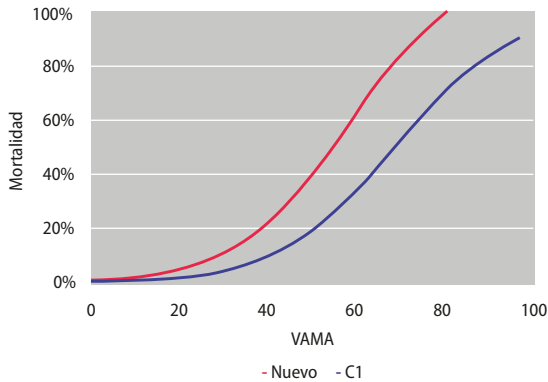


Figura 2. Explicación de lo que es el VAMA: Por la recalibración del modelo: A similar VAMA: ArgenSCORE I estima un riesgo mayor y ArgenSCORE II estima un riesgo menor.

que la predicha (3,96% vs. 8,20%; $p < 0,0001$)¹⁵. Para optimizar su rendimiento, el modelo original-(1999) fue recalibrado^{1,9,10,16} haciendo correr una nueva regresión logística para mortalidad intrahospitalaria en donde el modelo original-(1999) es la variable independiente y la mortalidad intrahospitalaria la variable dependiente^{16,17}. El ArgenSCORE recalibrado-(2007) mostró un área bajo la curva ROC de 0,81 (IC95%: 0,75-0,87), la prueba de Hosmer-Lemeshow²¹ fue no significativa (chi cuadrado=1,51; $p=0,68$) y se comprobó un nivel adecuado de relación entre mortalidad observada / mortalidad predicha en todos los pacientes ($p=0,92$). En la **Figura 1** se muestra el modelo recalibrado-(2007), donde la estimación de la mortalidad se grafica con sus correspondientes IC.

VALIDACIÓN DEL ARGENSCORE EN EL REEMPLAZO VALVULAR AÓRTICO

De un total de 1098 pacientes adultos consecutivos intervenidos con cirugía cardíaca desde febrero de 2008 hasta diciembre de 2012 en cuatro centros de la ciudad de Buenos Aires (Instituto FLENI, Sanatorio Dr. Julio Méndez, Sanatorio de la Trinidad y Sanatorio de los Arcos) se incluyeron para este análisis solamente los pacientes con RVA aislado y con RVA asociado con cirugía de revascularización miocárdica (CRM). Se excluyeron los pacientes con RVA asociado con reemplazo de la aorta ascendente, con otras intervenciones valvulares o con cualquier otra intervención cardíaca y todos los pacientes operados con otros tipos de cirugía cardíaca. Se evaluó como punto final la mortalidad intrahospitalaria, definida como la que ocurre hasta el alta del paciente.

En esta población multicéntrica se incluyeron 250 pacientes consecutivos; se validó en forma externa y temporal el ArgenSCORE y se comparó su rendimiento con el EuroSCORE en su versión I y con el EuroSCORE en su versión II. Se evaluó la discriminación de los tres modelos calculando el área bajo la curva ROC²⁰.

TABLA 1. Grupos de centros de acuerdo con la mortalidad observada y el VAMA.

Grupo de centros (Nº)	1	2	3
Mortalidad (%)	< 5	5 - 7	> 7
Nº de pacientes	597	952	999
Cantidad de centros	7	9	28
Mayor de 50 cirugías/año (%)	85,7	50	21,4
Centros con VAMA: 23 a 25	1	2	9
Centros con VAMA: 26 a 28	4	3	9
Centros con VAMA: 29 a 31	2	4	10

VAMA: valor absoluto medio del ArgenSCORE de cada centro.

Se evaluó además el poder de calibración de los modelos comparando la relación entre la mortalidad observada *versus* la mortalidad predicha en todos los pacientes y a través de los tres terciles de riesgo^{4,9,18}. La diferencia entre la media de la mortalidad observada y la media de la mortalidad predicha se calculó mediante prueba de la t ¹⁹. Se consideró significativo un valor de $p < 0,05$. En todos los análisis se utilizó el programa estadístico SPSS en su versión 21.0 (SPSS Inc., Chicago, III).

La población de validación externa y temporal consistió en 250 pacientes intervenidos en forma consecutiva con RVA aislado o asociado con CRM. La edad media de los pacientes fue de $68,62 \pm 13,3$ años, el 36,8% fueron mujeres y, en relación con el procedimiento quirúrgico, se realizó RVA aislado en el 67,2% de los pacientes y RVA asociado con CRM en el 32,8%; la mortalidad intrahospitalaria global fue del 3,6%. Las características de la población de validación y su mortalidad se muestran en la **Tabla 1**. La enfermedad valvular corregida fue estenosis aórtica predominante en el 86,4% de los pacientes e insuficiencia aórtica predominante en el 13,6%.

La validación externa y temporal mostró que el ArgenSCORE recalibrado presentó una muy buena capacidad para discriminar mortalidad intrahospitalaria en la cirugía de RVA. El área bajo la curva ROC fue de 0,82 (IC95%: 0,74-0,91; $p < 0,001$).

El modelo también mostró una muy buena capacidad para asignar riesgo operatorio (poder de calibración) en la población global, comprobando una excelente relación entre la mortalidad observada del 3,6% *versus* una mortalidad predicha del 3,39% ($p=0,471$). El análisis de acuerdo con los terciles de riesgo evidenció una capacidad más irregular en la calibración.

La evaluación del rendimiento del EuroSCORE en su versión I en esta población de pacientes intervenidos con RVA mostró una muy baja capacidad de este modelo para discriminar el riesgo de mortalidad quirúrgica, área bajo la curva ROC de 0,62 (IC95%: 0,43-0,82; $p=0,186$). Además, su capacidad para asignar riesgo de mortalidad fue inadecuada debido a que sobrevaloró en forma muy significativa el riesgo estimado con una relación entre la mortalidad observada del 3,6% *versus* una mortalidad predicha del 5,58% ($p < 0,0001$)²⁰.

Con respecto al EuroSCORE en su nueva versión II, este modelo mostró una aceptable capacidad para dis-

criminar el riesgo de mortalidad quirúrgica, área bajo la curva ROC de 0,76 (IC95%: 0,65-0,87; $p=0,007$), aunque este poder de discriminación fue menor (si bien no significativo) que el del ArgenSCORE. Sin embargo, a diferencia de los otros modelos, el EuroSCORE II mostró una muy pobre capacidad para asignar riesgo de mortalidad en esta población, ya que subvaloró en forma muy significativa el riesgo estimado, con una relación entre la mortalidad observada del 3,6% versus una mortalidad predicha del 1,64% ($p<0,0001$).

Los resultados de este trabajo muestran un excelente rendimiento del ArgenSCORE cuando se aplica en una población de pacientes intervenidos con RVA. Se observó un buen poder de discriminación para mortalidad y una buena capacidad para asignar el riesgo en toda la población, mostrando una excelente relación entre la mortalidad observada (3,6%) y la mortalidad predicha (3,39%), aunque los terciles inferiores sobrevaloran ligeramente la mortalidad de manera significativa.

En forma diferente, la aplicación del EuroSCORE I mostró una baja capacidad tanto para discriminar el riesgo de mortalidad quirúrgica como para asignar riesgo operatorio en la población total al sobrevalorar el riesgo quirúrgico estimado (relación mortalidad observada/mortalidad predicha 3,6% versus 5,58%). A su vez, el EuroSCORE II mostró una aceptable capacidad para discriminar el riesgo de mortalidad intrahospitalaria, aunque menor que la del ArgenSCORE. Sin embargo, en relación con su poder de calibración, el EuroSCORE II mostró una falencia en su rendimiento debido a la pobre capacidad para asignar riesgo de mortalidad quirúrgica, con una subvaloración significativa del riesgo estimado (relación mortalidad observada/mortalidad predicha 3,6% versus 1,64%).

Experiencias previas que evaluaron el comportamiento de diferentes modelos de riesgo internacionales en subpoblaciones con RVA aislado, mostraron pobres rendimientos en sus poderes predictivos. El EuroSCORE I, tanto en su versión logística como en su versión aditiva, sobreestimó el riesgo en pacientes con RVA aislado, mientras que el modelo de la *Society of Thoracic Surgeons* en general subestimó el riesgo en estas poblaciones²¹⁻²⁶. Por otra parte, un metaanálisis de todas las publicaciones que buscaron validar el EuroSCORE I en pacientes con cirugía valvular mostró un mal rendimiento de este modelo tanto en la población global como en los pacientes con RVA²⁷.

El recientemente publicado EuroSCORE II fue desarrollado con el objetivo de mejorar el rendimiento del EuroSCORE I. Si bien, a diferencia de su predecesor, permite estratificar el riesgo según el tipo de procedimiento a realizar, este nuevo modelo tanto en su publicación original²⁸ como en recientes validaciones externas³⁰⁻³² mostró que subestima el riesgo preoperatorio. De la misma forma, también se observó en el presente trabajo una subvaloración del riesgo operatorio estima-

TABLA 2. Centros con mortalidad observada intermedia (entre 5 y 7%).

Centro (Nº)	6	20	12	24	43	7	11	25	38
Complejidad	Baja			Intermedia			Alta		
VAMA	25,0	25,6	26,3	27,4	26,1	31,2	29,9	30,4	29,7
VAMA/Mort.	4,54	3,71	4,20	4,21	4,14	5,42	5,12	4,82	4,71
	Corte								
Mort. obs. (%)	5,5	6,9	6,25	6,5	6,3	5,7	5,7	6,3	6,3
ArgenSCORE I (mort. estim) (%)	7,59	7,93	8,35	9,05	8,23	11,9	10,8	11,5	10,7
ArgenSCORE II (mort. estim) (%)	2,79	2,94	3,13	3,45	3,07	4,83	4,31	4,66	4,23

VAMA: valor absoluto medio del ArgenSCORE de cada centro. Mort.: mortalidad. obs.: observada. estim.: estimada.

do. Estos hallazgos sustentan las ventajas de desarrollar y utilizar modelos locales en la estratificación del riesgo preoperatorio²⁻⁴.

La aplicación de modelos de riesgo preoperatorio internacionales de uso generalizado ha mostrado limitaciones en sus rendimientos cuando dichos modelos fueron utilizados en poblaciones diferentes de aquellas sobre las que fueron desarrollados^{2,3,6,29}. Esto puede estar relacionado con diferencias geográficas y epidemiológicas en el perfil de riesgo, en las estrategias quirúrgicas utilizadas y en la toma de decisión de una eventual cirugía cardíaca entre diferentes países e incluso entre centros quirúrgicos de un mismo país^{2,3,29,30}. Estas diferencias epidemiológicas significativas también fueron observadas en nuestra experiencia, al comparar una población local con la población de desarrollo del EuroSCORE I^{12,13}.

ARGENSCORE AJUSTADO AL CENTRO EN BASE A UN ANÁLISIS DEL REGISTRO MULTICÉNTRICO CONAREC XVI: REALIDAD DE LA CIRUGÍA CARDÍACA EN ARGENTINA: MUNDO REAL

Como observamos anteriormente, el ArgenSCORE es un modelo de riesgo local que tiene dos versiones, su versión original desarrollada en 1999 sobre una población con mortalidad operatoria media de 8,2% (ArgenSCORE I) y una segunda versión que surge de la recalibración del modelo original realizada en 2007 sobre una población con una mortalidad operatoria media de 3,96% (ArgenSCORE II).

En 42º Congreso Argentino de Cardiología, llevado a cabo en octubre de 2016, se presentó este reciente estudio donde se analizaron 49 centros del registro CONAREC XVI (2548 pacientes) y se excluyeron 5 centros porque cada uno de ellos solo ingresaron un paciente al registro, por lo tanto se incluyeron en el análisis 2543 pacientes de 44 centros del registro. Para mantener la confidencialidad los centros se describen con un número. Se analizaron las variables de riesgo preoperatorias, la complejidad de las intervenciones y la mortalidad intrahospitalaria en cada centro.

La mortalidad intrahospitalaria en todo el registro fue de 7,69%. Como la mortalidad en los centros fue muy heterogénea, los 44 centros se dividieron en 3 grupos, Gru-

TABLA 3. Centros con relación VAMA / mortalidad observada $\geq 4,54$. Mejor estimación con ArgenSCORE II

Centro (N°)	22	29	35	42	7	11	23	32
Complejidad	Baja	Intermedia			Alta			
VAMA	25,0	27,9	28,0	28,8	31,2	29,9	29,44	31,75
VAMA/Mort.	14,75	15,51	7,37	11,84	5,42	5,12	11,32	9,33
Mort. obs. (%)	1,70	2,40	3,84	2,43	5,7	5,7	2,6	3,4
ArgenSCORE I (mort. estim.) (%)	7,59	9,38	9,45	10,0	11,9	10,8	10,45	12,3
ArgenSCORE II (mort. estim.) (%)	2,79	3,61	3,64	3,91	4,83	4,31	4,12	5,0

VAMA: valor absoluto medio del ArgenSCORE de cada centro. Mort.: mortalidad. obs.: observada. estim.: estimada.

po 1: mortalidad $<5\%$ (7 centros: 597 pacientes), Grupo 2: mortalidad entre 5 y 7% (9 centros: 952 pacientes), y Grupo 3: mortalidad $>7\%$ (28 centros: 999 pacientes). La mortalidad operatoria se relacionó con el volumen de intervenciones quirúrgicas realizadas por año en cada centro; en el Grupo 1, el 83,3% de los centros intervinieron más de 50 cirugías por año, en el Grupo 2, el 50%, mientras que en el Grupo 3 solo el 27,3% de los centros. (Tabla 1). De acuerdo a la complejidad del centro determinado por el VAMA se identificaron 16 centros con VAMA entre 29 y 31 (alto riesgo preoperatorio y complejidad quirúrgica), 16 centros con VAMA entre 26 y 28 (riesgo y complejidad intermedia) y 12 centros con VAMA entre: 23 y 25 (riesgo y complejidad baja).

VALORACIÓN DEL PUNTO DE CORTE DE LA RELACIÓN VAMA/MO

Con el cálculo del VAMA de cada centro se elaboró la relación VAMA/MO. Para definir el punto de corte que nos permita discernir qué versión del ArgenSCORE aplicar, identificamos en el grupo de centros con mortalidad intermedia al centro N°6. Este centro con mortalidad del 5,5%, tuvo un VAMA de 25 y una relación VAMA/MO de 4,54. En la Tabla 2 observamos que, en este centro, el ArgenSCORE I sobrestima la mortalidad en 2,09% mientras que el ArgenSCORE II subestima la mortalidad en 2,71% siendo de esta manera el centro de este grupo intermedio que presentó la menor dispersión entre ambos valores de estimación de riesgo. Además como este centro es el que presenta la menor mortalidad del grupo intermedio debería aplicar el ArgenSCORE II. Por lo tanto, en centros con una relación VAMA/MO $\geq 4,54$ se debería aplicar el ArgenSCORE II y en centros con una relación $<4,54$ el ArgenSCORE I. La capacidad de este punto de corte de la relación VAMA/MO (centro N°6) se evaluó mediante el rendimiento de ambos modelos comparando ArgenSCORE I versus ArgenSCORE II. En la Tabla 2 se describe el grupo de centros con mortalidad intermedia, donde se observa como en aquellos con relación VAMA/MO $\geq 4,54$ el ArgenSCORE II estima mejor el riesgo, en cambio en centros con relación VAMA/MO $<4,54$ el ArgenSCORE I presenta un mejor rendimiento. De esta manera, aplicando el ArgenSCORE II en los centros del grupo intermedio que tengan una relación VAMA/MO

TABLA 4. Centros con relación VAMA / mortalidad observada $< 4,54$. Mejor estimación con ArgenSCORE I.

Centro (N°)	2	9	20	1	3	28	8	21	39
Complejidad	Baja			Intermedia			Alta		
VAMA	25,7	25,9	25,6	26,5	27,7	27,8	31,4	31,6	31,4
VAMA/Mort.	2,82	2,87	3,71	3,19	2,91	2,69	2,83	2,32	2,90
Mort. obs. (%)	9,1	9,0	6,9	8,3	9,5	10,3	11,1	13,6	10,8
ArgenSCORE I (mort. estim.) (%)	7,99	8,17	7,93	8,47	9,25	9,31	12,0	12,2	11,96
ArgenSCORE II (mort. estim.) (%)	2,97	3,05	2,94	3,19	3,54	3,58	4,91	5,0	4,87

VAMA: valor absoluto medio del ArgenSCORE de cada centro. Mort.: mortalidad. obs.: observada. estim.: estimada.

$\geq 4,54$ (centros N° 7, 11, 25, 38 y el centro N°6 inclusive) y el ArgenSCORE I en los centros que tengan una relación VAMA/MO $<4,54$, (centros N° 12, 20, 24 Y 43) (Tabla 2), el ArgenSCORE II mostró un área bajo la curva ROC de 0,75 (intervalo de confianza del 95% [IC95%]: 0,67-0,83) y prueba de Hosmer-Lemeshow no significativa ($p=0,07$) y el ArgenSCORE I un área bajo la curva ROC de 0,79 (IC95%: 0,65-0,92) y prueba de Hosmer-Lemeshow no significativa ($p=0,23$). Ambos modelos evidenciaron una adecuada relación mortalidad observada / mortalidad predicha en todos los decilos de riesgo. En cambio, el comportamiento fue muy diferente cuando aplicamos el ArgenSCORE I al centro N°6, es decir aplicando el ArgenSCORE II en los centros del grupo intermedio con relación VAMA/MO solo $>4,54$ (sin el centro N°6) y aplicando el ArgenSCORE I en centros con una relación VAMA/MO $\leq 4,54$, (incluido el centro N°6) (Tabla 2). En este escenario diferente el ArgenSCORE II, aunque mostró un área bajo la curva ROC de 0,77 (IC95%: 0,69-0,84), presentó una prueba de Hosmer-Lemeshow con diferencia significativa ($p=0,03$), mostrando una clara tendencia a subestimar el riesgo predicho. En cambio, el ArgenSCORE I mostró una menor área bajo la curva ROC, 0,71 (IC95%: 0,57-0,86) (menor discriminación), con prueba de Hosmer-Lemeshow no significativa ($p=0,41$). Al aplicar estos criterios para elegir el modelo a utilizar, de acuerdo a la mortalidad y complejidad por institución, se observó una adecuada estimación del riesgo en todos los centros de este registro. La capacidad para asignar el riesgo (relación entre mortalidad observada versus mortalidad predicha) en todo el registro fue mejor con el ArgenSCORE II en los centros que presentan una relación VAMA/MO $\geq 4,54$ (Tabla 3) y con el ArgenSCORE I en los centros con una relación VAMA/MO $<4,54$ (Tabla 4).

PARA QUÉ SIRVE LA RELACIÓN VAMA/MORTALIDAD OBSERVADA

Obviamente, esta relación VAMA / mortalidad observada es más útil en los casos de centros grises o que tienen una mortalidad intermedia, es decir con una mortalidad observada entre el 5% y el 7% (Tabla 2). Sin duda, en los grupos de centros que se presentan con extremos

en su mortalidad, sería más seguro aplicar una versión u otra del ArgenSCORE. Es decir en centros con mortalidades mayores al 7% estaríamos más seguros de aplicar el ArgenSCORE I, como podemos observar en la **Tabla 4**, mientras que en los centros con mortalidades menores al 5% estaríamos más seguros de aplicar el ArgenSCORE II (**Tabla 3**). Es decir esta relación VAMA

/ mortalidad observada nos daría más seguridad en el caso que nuestro centro sea una institución con mortalidad intermedia (entre el 5% y el 7%). Al evaluar ambas versiones tanto en centros con mortalidad menor al 5% como en centros con mortalidad mayor al 7%, ambos se comportaron correctamente tanto en curva ROC como con pruebas de Hosmer - Lemeshow.

BIBLIOGRAFÍA

- Ivanov J, Tu JV, Naylor CD. Ready-made, recalibrated, or remodeled? Issues in the use of risk indexes for assessing mortality after coronary artery bypass graft surgery. *Circulation* 1999;99:2098-104.
- Yap CH, Reid C, Yip M et al. Validation of the EuroSCORE model in Australia. *Eur J Cardiothorac Surg* 2006;29:441-6.
- Asimakopoulos G, Al-Ruzzeh S, Ambler G et al. An evaluation of existing risk stratification models as a tool for comparison of surgical performances for coronary artery bypass grafting between institutions. *Eur J Cardiothorac Surg* 2003;23:935-42.
- Al-Ruzzeh S, Asimakopoulos G, Ambler G et al. Validation of four different risk stratification systems in patients undergoing off-pump coronary bypass graft surgery: a UK multicentre analysis of 2223 patients. *Heart* 2003;89:432-5.
- Nashef SA, Roques F, Michel P et al. Coronary surgery in Europe: comparison of the national subsets of the European System for Cardiac Operative Risk Evaluation database. *Eur J Cardiothorac Surg* 2000;17:396-9.
- Wynne-Jones K, Jackson M, Grotte G, Bridgewater B. Limitations of the Parsonnet score for measuring risk stratified mortality in the north west of England. The North West Regional Cardiac Surgery Audit Steering Group. *Heart* 2000;84:71-8.
- Anderson RP, Jin R, Grunkemeier GL. Understanding logistic regression analysis in clinical reports: an introduction. *Ann Thorac Surg* 2003;75:753-7.
- Harrell FE. *Regression Modeling Strategies*. New York, NY: Springer-Verlag; 2001.
- Omar RZ, Ambler G, Royston P, Eliahoo J, Taylor KM. Cardiac surgery risk modeling for mortality: a review of current practice and suggestions for improvement. *Ann Thorac Surg* 2004;77:2232-7.
- Shahian DM, Blackstone EH, Edwards FH et al. Cardiac surgery risk models: a position article. *Ann Thorac Surg* 2004;78:1868-77.
- Edwards FH, Grover FL, Shroyer AL, Schwartz M, Bero JW. The Society of Thoracic Surgeons National Cardiac Surgery Database: Current risk assessment. *Ann Thorac Surg* 1997;63:903-8.
- Carosella VC, Navia JL, Al-Ruzzeh S et al. The first Latin-American risk stratification system for cardiac surgery: can be used as a graphic pocket-card score. *Interact CardioVasc Thorac Surg* 2009;9:203-8.
- Carosella VC, Grancelli H, Rodríguez W et al. Primer puntaje latinoamericano en cirugía cardíaca (ArgenSCORE): validación externa y temporal a 10 años de su desarrollo. *Rev Argent Cardiol* 2011;79:500-7.
- Hannan EL, Kilburn H, Jr, O'Donnell JF, Lukacic G, Shields EP. Adult open heart surgery in New York State. An analysis of risk factors and hospital mortality rates. *JAMA* 1990;264:2768-74.
- Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29-36.
- Jin R, Grunkemeier GL, Starr A. Validation and refinement of mortality risk models for heart valve surgery. *Ann Thorac Surg* 2005;80:471-9.
- Metnitz PGH, Lang T, Vesely H, Valentin A, Le Gall JR. Ratios of observed to expected mortality are affected by differences in case mix and quality of care. *Intensive Care Med* 2000;26:1466-72.
- Shahian DM, Normand SL, Torchiana DF, Lewis SM, Pastore JO, Kuntz RE, et al. Cardiac surgery report cards: comprehensive review and statistical critique. *Ann Thorac Surg* 2001;72:2155-68.
- Beck DH, Smith GB, Pappachan JV, Millar B. External validation of the SAPS II, APACHE II and APACHE III prognostic models in South England: a multicentre study. *Intensive Care Med* 2003;29:249-56.
- Carosella VC, Mastantuono C, Golovonevsky V et al. Validación prospectiva y multicéntrica del ArgenSCORE en la cirugía de reemplazo valvular aórtico. Comparación con el EuroSCORE I y el EuroSCORE II. *Rev Argent Cardiol* 2014; 82: 6-12.
- Bates ER. Treatment options in severe aortic stenosis. *Circulation* 2011;124:355-9. <http://doi.org/c8zw8t>
- Osswald BR, Gegouskov V, Badowski-Zyla D et al. Overestimation of aortic valve replacement risk by EuroSCORE: implications for percutaneous valve replacement. *Eur Heart J* 2009;30:74-80. <http://doi.org/djzj4p>
- Dewey TM, Brown D, Ryan WH, Herbert MA, Prince SL, Mack MJ. Reliability of risk algorithms in predicting early and late operative outcomes in high risk patients undergoing aortic valve replacement. *J Thorac Cardiovasc Surg* 2008;135:180-7. <http://doi.org/ddxf8>
- Wendt D, Osswald BR, Kayser K et al. Society of Thoracic Surgeons Score is superior to the EuroSCORE determining mortality in high risk patients undergoing isolated aortic valve replacement. *Ann Thorac Surg* 2009;88:468-74. <http://doi.org/bmh7xr>
- Basraon J, Chandrashekar YS, John R et al. Comparison of risk scores to estimate perioperative mortality in aortic valve replacement surgery. *Ann Thorac Surg* 2011;92:535-40. <http://doi.org/dpqpch>
- Sündermann S, Dademasch A, Praetorius J et al. Comprehensive assessment of frailty for elderly high-risk patients undergoing cardiac surgery. *Eur J Cardiothorac Surg* 2011;39:33-7.
- Parolari A, Pesce LL, Trezzi M et al. EuroSCORE performance in valve surgery: a meta-analysis. *Ann Thorac Surg* 2010;89:787-93.
- Nashef SA, Roques F, Sharples LD et al. EuroSCORE II. *Eur J Cardiothorac Surg* 2012;41:734-44; discussion 44-5.
- Bridgewater B, Neve H, Moat N, Hooper T, Jones M. Predicting operative risk for coronary artery surgery risk in the United Kingdom: a comparison of various prediction algorithms. *Heart* 1998;79:350-5.
- Nashef SAM, Roques F, Michel P et al. Coronary surgery in Europe: comparison of the national subsets of the European System for Cardiac Operative Risk Evaluation database. *Eur J Cardiothorac Surg* 2000;17:396-9.