

## ArgenSCORE versus EuroSCORE II

### Al Director

Hemos leído en el número anterior de la *Revista* el artículo de Carosella y colaboradores. (1) Los autores trabajan en puntajes (*scores*) de riesgo desde hace años y son investigadores de referencia en el tema. En este artículo comparan los resultados del ArgenSCORE con el EuroSCORE I y II entre pacientes sometidos a reemplazo valvular aórtico (aislado o combinado) y concluyen con las ventajas del puntaje local. Aunque el desarrollo de modelos locales es importante, existen limitaciones de este análisis que merecen comentarios. La calibración se realizó comparando la diferencia entre mortalidades observadas y esperadas, globalmente y en terciles de riesgo. Los autores no usaron la prueba habitual de calibración de Hosmer-Lemeshow que permite una evaluación global y regional (por terciles). Pese a sus limitaciones, la mayoría de los estudios refieren este estadístico. Aunque la capacidad del ArgenSCORE para asignar riesgo parecería globalmente adecuada (3,6% vs. 3,4%;  $p = 0,471$ ), el análisis por terciles de la Tabla 3 muestra que en dos de los tres terciles la mortalidad predicha por el ArgenSCORE difirió significativamente de la observada; o sea, solo fue adecuada en los pacientes con riesgo mayor (tercil superior). Otra duda de la separación por terciles es por qué el segundo tercil tiene solo 77 operados y el superior 89, cuando por definición cada tercil debería tener entre 83 y 84 individuos para un  $n = 250$ . Otro aspecto controversial es el relativo a la comparación "entre" modelos, debido a que no se hacen comparaciones formales entre curvas ROC o razones de verosimilitud, sino una comparación informal. El método de Hanley-McNeil hubiese permitido evaluar diferencias entre ellos. Sin embargo, la comparación de intervalos de confianza entre áreas ROC del ArgenSCORE y el EuroSCORE II sugiere diferencias no significativas, indicando que los modelos son equivalentes y que solo podría dirimirse una ventaja aumentando el tamaño de la muestra.

También hay que destacar que el uso de un puntaje internacional tiene ventajas comparativas y de inserción en el mundo, cosa más difícil de lograr con el ArgenSCORE. Más allá de la sobrestimación del riesgo del EuroSCORE I, la validez de las predicciones del EuroSCORE II está en estudio; sin embargo, las validaciones de este último registran áreas ROC de entre 0,760 y 0,990. (2-5)

A los fines comparativos calculamos los resultados prospectivos con el EuroSCORE II de 192 reemplazos aórticos aislados o combinados realizados en 2012-2013. El área ROC fue de 0,808 y la prueba de Hosmer-Lemeshow mostró una buena calibración ( $p = 0,656$ ).

Más allá de que el EuroSCORE II es más preciso que su antecesor en los riesgos altos, es excesivamente exigente con los riesgos bajos. Así, pacientes jóvenes

sin morbilidades sometidos a cirugía coronaria tendrán un riesgo esperado cercano al 0,5%. Este nuevo estándar de calidad exige enfrentar el problema de mejorar progresivamente los resultados quirúrgicos en nuestro medio. En la situación de tener que medir el desempeño de puntajes de riesgo, el límite de la predicción de los eventos de probabilidad sigma-3 ( $\approx 0,5\%$ ) se halla en la presunción de que a este nivel de probabilidad esperada, la muerte se transforma en un hecho prácticamente impredecible.

**Raúl A. Borracci**<sup>MTSAC</sup>

Universidad Austral, Argentina

**Javier Mariani**<sup>MTSAC</sup>

Hospital El Cruce, Florencio Varela,  
Buenos Aires, Argentina

### BIBLIOGRAFÍA

1. Carosella VC, Mastantuono C, Golovonevsky V, Cohen V, Grancelli H, Rodríguez W, 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;82:5-11. <http://doi.org/rr7>
2. 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;1-5. Epub ahead of print.
3. 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* 2012;98:1568-72. <http://doi.org/pvt>
4. Chalmers J, Pullan M, Fabri B, McShane J, Shaw M, Mediratta N, et al. Validation of EuroSCORE II in a modern cohort of patients undergoing cardiac surgery. *Eur J Cardiothorac Surg* 2013;43:688-94. <http://doi.org/rh9>
5. Carnero-Alcázar M, Silva Guisasaola JA, Reguillo Lacruz FJ, Maroto Castellanos LC, Cobiella Carnicer J, Villagrán Medinilla E, et al. Validation of EuroSCORE II on a single-centre 3800 patient cohort. *Interact CardioVasc Thorac Surg* 2013;16:293-300. <http://doi.org/pvv>

### Respuesta de los autores

Agradecemos los comentarios de los Dres. Borracci y Mariani sobre nuestra experiencia y nos permitimos hacer las siguientes observaciones.

En relación con el uso de la prueba de Hosmer-Lemeshow, similar a nuestro trabajo, múltiples publicaciones evaluaron la calibración para validar modelos en cirugía cardíaca y terapia intensiva comparando la mortalidad observada y estimada (MO/ME) globalmente y en diferentes grupos de riesgo mediante otras pruebas. (1, 2) La prueba de Hosmer-Lemeshow no evalúa por terciles (como se señala), sino por deciles de riesgo. Aunque se utiliza frecuentemente, tiene limitaciones y otros investigadores han alertado sobre sus problemas para calibrar modelos, destacando la importancia de evaluar la relación MO/ME en grupos de riesgo mediante otras pruebas.

La capacidad del ArgenSCORE para asignar riesgo fue muy adecuada en la población global (3,6% vs. 3,39%;  $p = 0,471$ ); el EuroSCORE I sobrevaloró el riesgo (3,6% vs. 5,58%;  $p < 0,0001$ ) y el EuroSCORE II lo subvaloró (3,6% vs. 1,64%;  $p < 0,0001$ ). En el análisis por terciles, los tres modelos mostraron diferencias significativas en dos de tres terciles, pero con la ventaja de que el ArgenSCORE mostró una buena predicción del riesgo global, hallazgo no observado con los otros dos modelos.

Llama la atención la observación sobre el número de casos por terciles; estas categorías se construyen ordenando la muestra por su valor (mortalidad estimada), se calcula qué valor corresponde a la partición deseada (33,33%) y se observa a qué valor corresponde en la distribución. Como la muestra tiene muchos números repetidos y dado que un mismo valor no puede estar en dos categorías diferentes, es perfectamente posible la diferencia observada en el número de casos por terciles.

Es evidente cierta confusión en la interpretación de nuestros resultados y conclusiones. En el artículo se describió que el EuroSCORE II mostró una aceptable capacidad para discriminar el riesgo (curva ROC: 0,76, IC: 0,65-0,87), algo menor que el ArgenSCORE (curva ROC: 0,82, IC: 0,74-0,91), aunque esta diferencia sugiere ser no significativa por los intervalos de confianza. Si concluimos, como comentamos previamente, que el EuroSCORE II subvaloró significativamente el riesgo estimado.

Aunque el EuroSCORE II mejoró el rendimiento de su predecesor, requiere futuras validaciones que determinen su real capacidad. Los Dres. Borracci y Mariani citan estudios que evaluaron la validación del EuroSCORE II, pero solo describen su capacidad de discriminación (área ROC), que no debe confundirse con su calibración (asignación de riesgo). Casualmente, tres de los artículos citados (y otros no citados), evidenciaron inadecuada calibración del EuroSCORE II porque subestimó el riesgo estimado, similar a nuestra experiencia. (3-5)

Coincidimos en que estos modelos constituyen un “estándar de calidad para mejorar resultados quirúrgicos”, pero otra utilidad importante es poder estimar el riesgo operatorio en la mayor parte de la población, justamente la de riesgo bajo. No consideramos esta comparación entre modelos como una competencia (“ArgenSCORE vs. EuroSCORE II”); solo se busca evaluar cómo funciona nuestro puntaje en relación con modelos ya establecidos. Aunque un puntaje internacional como el EuroSCORE puede tener “mayor inserción global”, el desarrollo y la validación de modelos locales como el ArgenSCORE pueden contribuir a mejorar la estratificación de riesgo en nuestra población.

Victorio C. Carosella, César Cárdenas y  
Hugo Grancelli<sup>MTSAC</sup>  
Instituto FLENI

## BIBLIOGRAFÍA

1. Bhatti F, Grayson AD, Grotte G, Fabri BM, Au J, Jones M, et al; North West Quality Improvement Programme in Cardiac Interventions. The logistic EuroSCORE in cardiac surgery: how well does it predict operative risk? *Heart* 2006;92:1817-20. <http://doi.org/d6wxpr>
2. Metnitz PG, 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. <http://doi.org/cws9wf>
3. Carnero-Alcázar M, Silva Guisasaola JA, Reguillo Lacruz FJ, Maroto Castellanos LC, Cobiella Carnicer J, Villagrán Medinilla E, et al. Validation of EuroSCORE II on a single-centre 3800 patient cohort. *Interact Cardiovasc Thorac Surg* 2013;16:293-300. <http://doi.org/pvv>
4. 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 23740 patients from the Society for Cardiothoracic Surgery in Great Britain and Ireland National Database. *Heart* 2012;98:1568-72. <http://doi.org/pvt>
5. Chalmers J, Pullan M, Fabri B, McShane J, Shaw M, Mediratta N, et al. Validation of EuroSCORE II in a modern cohort of patients undergoing cardiac surgery. *Eur J Cardiothorac Surg* 2013;43:688-94. <http://doi.org/rh9>

REV ARGENT CARDIOL 2014;82:167-168. <http://dx.doi.org/10.7775/rac.es.v82.i2.3936>

VÉASE CONTENIDO RELACIONADO: *Rev Argent Cardiol* 2013;82:6-12. <http://dx.doi.org/10.7775/rac.es.v82.i1.3388>

## Período preeyectivo en el tilt test

“Hay tres formas de mentira, la real,  
la falsa y las estadísticas.”

Atribuido a BENJAMIN DISRAELI

Felicito a los autores del artículo “Fast tilt test: tras un nuevo paradigma en el manejo del síncope reflejo” por la capacidad e ingenio para encontrar soluciones fáciles a problemas difíciles. (1) El período preeyectivo (PEy) del ventrículo izquierdo nos recuerda nuestros esfuerzos para evaluarlo en los años sesenta. Comento algunas experiencias obtenidas esperando que puedan ser de utilidad para esta metodología.

El estímulo eléctrico requiere reclutar una cantidad razonable de miocitos que ejerzan una fuerza que eleve la presión intraventricular, cierre la mitral y después abra la aórtica. Todo ello constituye el PEy que incluye una fase *electromecánica*, desde el QRS hasta el cierre de la válvula mitral, y otra *isovolumétrica (IS)* sistólica, desde allí hasta la apertura de las sigmoideas aórticas. (2)

El retardo *electromecánico* es habitualmente constante, luego las variaciones del PEy son provocadas generalmente por la fase IS. Es aconsejable la derivación electrocardiográfica que exprese el inicio del QRS. Corrientemente se utiliza DII, pero no siempre es la indicada. (3)

El IS depende de múltiples variables, que se pueden resumir en precarga, poscarga, frecuencia cardíaca (FC) y estado del miocardio. (3)

Reducciones de la precarga, manteniendo FC y poscarga fijas, lo aumenta y viceversa.

Incrementos de la poscarga, con precarga y FC fijas, lo prolonga y viceversa.

Mayor FC, con precarga y poscarga estables, lo reducen.

Depresiones del estado inotrópico, con las otras variables fijas, lo prolongan.

Considerar variables fisiológicas como la respiración, por lo cual se deben promediar dos ciclos, variaciones circadianas, posición del paciente (de pie reduce la precarga), prolongaciones etarias y las del sexo femenino. Se acorta con drogas simpaticomiméticas e inotrópicas y se prolonga con betabloqueantes y vasodilatadores. Las variaciones patológicas son múltiples. (4)

Los nitritos provocan caída de precarga, presión arterial y taquicardia refleja, técnica con demasiadas variables para establecer constantes.

Los controles, hacerlos en condiciones basales, ambientes adecuados, con ayuno de 12 horas, sin tabaco ni cafeína. Los transductores deben tener validación y considerarse amortiguación y resonancia. En el estudio no se mencionan características del sistema ni su validación. (1) El pulso de la Figura 2 no muestra la pequeña onda previa al pie del latido, quizás por baja sensibilidad, y el pulso de la Figura 4 es técnicamente deficiente.

El PEy se ha evaluado después de ejercicios isométricos y ergométricos, variaciones posicionales, patologías cardíacas, y con diferentes drogas. También se midió con ecocardiografía. (5)

Los síncope vasogénicos suceden por diferentes causas que no ocurren en condiciones normales durante el examen, hecho que genera desánimo respecto de la reproducibilidad del *tilt test*.

En nuestra experiencia, basada en varios miles de registros, la medición del PEy ha probado utilidad en el seguimiento individual y algo en la evaluación grupal, pero no así en el paciente aislado por la amplia gama de variables que lo afectan.

En el estudio (1) se ha encontrado respaldo estadístico a los resultados que generan lógicas conclusiones e ilusiones, pero no siempre los números son expresión de la realidad, lo que lleva a la necesidad de un análisis más amplio y tecnología más rigurosa para obtener aseveraciones concluyentes y cuasi lapidarias.

Ricardo J. Esper<sup>MTSAC</sup>

## BIBLIOGRAFÍA

1. Villamil AM, Torres Y, Mariani JA, Perona C, Tajer CD, Piccinini JM y cols. Fast tilt test: tras un nuevo paradigma en el manejo del síncope reflejo. *Rev Argent Cardiol* 2014;82:42-9. <http://doi.org/rr8>
2. Esper RJ. Tiempos sistólicos del ventrículo izquierdo. *Rev Argent Cardiol* 1973;41:451-6.
3. Esper RJ, Madoery RJ. Evaluación de la función ventricular mediante el fonomecanocardiograma. En: Esper RJ, Madoery RJ, editores. *Progresos en Auscultación y Fonomecanocardiografía*. Buenos Aires: López Libreros Editores; 1974. p. 279-91.

4. Weissler AM. Interpreting systolic time intervals in man. *J Am Coll Cardiol* 1983;2:1019.

5. Esper RJ, Madoery RJ. El pulso arterial. Los tiempos sistólicos del ventrículo izquierdo. En: Esper RJ, Madoery RJ, editores. *Examen físico cardiovascular*. Buenos Aires: Ed Promedicina; 1987. p. 42-59.

## Respuesta de los autores

Agradecemos al Dr. Ricardo Esper su interés en nuestro trabajo (1) y sus aportes críticos. Creemos, sin embargo, que su lectura no ha captado la contribución de nuestra exploración, y sorprende un poco su cuestionamiento al aporte de la estadística en la investigación científica.

Hemos utilizado la mejor tecnología disponible, extremado las reiteraciones y validaciones del método y aplicado la mejor estadística recomendada. Desde los trabajos de Pierre Alexander Louis en el siglo XIX en Francia, que permitieron por un estudio controlado descartar el uso de la sangría, el método estadístico constituye una herramienta básica de la metodología científica. No encontramos en su carta observaciones sobre nuestra metodología estadística aplicable a métodos diagnósticos (sensibilidad, especificidad, área bajo la curva ROC) y no parece que el pensamiento estadístico requiera nuestra defensa.

Respondiendo a sus cuestionamientos metodológicos, en 2009 publicamos nuestro primer trabajo sobre el comportamiento del parámetro (2) y recibimos su primera carta de lector (3) en la que se explayó en su experiencia sobre la evaluación del período preeyectivo mediante fonomecanocardiogramas en las décadas de los sesenta y setenta, hecho que hoy ha reiterado con igual énfasis y similares planteos.

En ese entonces respondimos (4): “Analizamos el **comportamiento dinámico** del intervalo adquirido en decúbito dorsal y durante inclinación a 70 grados, tomando el inicio del QRS en el mismo punto del registro de monitorización en ambas condiciones”. Ello quita relevancia al criterio del inicio de la actividad eléctrica, mientras sea el mismo, ya que el parámetro surge de la diferencia numérica y no como valor estático absoluto. Le reiteramos que el transductor carotídeo posee (4) “un ancho de banda desde el sentido mecánico a la digitalización de la señal medida a caída de -3 db de 1 Hz a 330 Hz”, y se encuentra convalidado.

En referencia a la fisiopatología del parámetro, ya expusimos en 2009 (4) y ahora ampliamos, (1) que la evaluación simultánea con Doppler cardíaco durante el *tilt test* en un grupo de pacientes con síncope vasovagal y prolongación del parámetro, mostró que el **período preeyectivo no se modificó significativamente tras la inclinación**. Ello sugiere que el fenómeno se debe a **prolongación del tiempo eyectivo**, reflejando probablemente una respuesta de distensibilidad particular en pacientes con crisis reflejas.

Más allá de las consideraciones fisiopatológicas, nuestro planteo es pragmático en el sentido de que el parámetro medido como fue medido predice bastante

bien el resultado del *tilt test*, con la ventaja de ser sencillo, reproducible, no requerir camilla basculante y ser independiente del operador, en línea con “una solución fácil a un problema difícil”.

PD/ La cita atribuida a Benjamin Disraeli, aunque acuñada por el humorista norteamericano Mark Twain, está mal traducida, y originalmente refiere tres formas de mentiras en progresivo descrédito: las mentiras, las malditas mentiras y las estadísticas, y no “falsas, reales y estadísticas”, como cita. Si hubiera mentiras falsas como propone su errónea traducción, resultarían así verdades, lo que nos introduciría en el laberinto de las paradojas autorreferenciales (ejemplo: “esta frase es falsa”, no puede ser verdadera ni falsa), un verdadero sinsentido.

**Alejandro Villamil<sup>MTSAC</sup>, Javier Mariani<sup>MTSAC</sup>,  
Carlos Tajer<sup>MTSAC</sup>**

## BIBLIOGRAFÍA

1. Villamil AM, Torres Y, Mariani JA, Perona C, Tajer CD, Piccinini JM y cols. Fast tilt test: tras un nuevo paradigma en el manejo del síncope reflejo. *Rev Argent Cardiol* 2014;82:42-9. <http://doi.org/rr8>
2. Villamil AM, Perona C, Carnero GS, Torres Y, Mariani JS, Tajer CD y cols. Evaluación del comportamiento del intervalo desde el inicio del QRS al inicio de la onda del pulso radial y carotídeo con el resultado del tilt test. *Rev Argent Cardiol* 2009;77:347-53.
3. Esper R. Evaluación del comportamiento del intervalo desde el inicio del QRS al inicio de la onda de pulso radial y carotídeo con el resultado del tilt test. Carta de Lectores. *Rev Argent Cardiol* 2009;78:185-6.
4. Villamil A, Torres Y, Tajer C. Evaluación del comportamiento del intervalo desde el inicio del QRS al inicio de la onda de pulso radial y carotídeo con el resultado del tilt test. Respuesta de los autores. *Rev Argent Cardiol* 2009;78:186.

---

REV ARGENT CARDIOL 2014;82:168-170. <http://dx.doi.org/10.7775/rac.es.v82.i2.4053>