

## Concordancia intercentros en la interpretación de la ecocardiografía de ejercicio

Jesús Peteiro<sup>a</sup>, Ángel M. Alonso<sup>b</sup>, Rafael Florenciano<sup>c</sup>, Carlos González Juanatey<sup>d</sup>, Gonzalo de la Morena<sup>c</sup>, Ignacio Iglesias<sup>f</sup>, Mar Moreno<sup>e</sup> y Miguel A. Rodríguez<sup>f</sup>

<sup>a</sup>Unidad de Ecocardiografía. Hospital Juan Canalejo. A Coruña. España.

<sup>b</sup>Unidad de Ecocardiografía. Hospital Txagorritxu. Vitoria. Álava. España.

<sup>c</sup>Unidad de Ecocardiografía. Virgen de la Arrixaca. Murcia. España.

<sup>d</sup>Unidad de Ecocardiografía. Hospital Xeral. Lugo. España.

<sup>e</sup>Unidad de Ecocardiografía. Hospital Gregorio Marañón. Madrid. España.

<sup>f</sup>Unidad de Ecocardiografía. Hospital de León. León. España.

**Introducción y objetivos.** Se ha demostrado un acuerdo bajo-moderado en la interpretación de la ecocardiografía con dobutamina, pero no se ha realizado un estudio similar con ejercicio. El objetivo fue evaluar la concordancia intercentros de la ecocardiografía de ejercicio y la precisión de la técnica realizada de forma «ciega».

**Pacientes y método.** Cada uno de los 6 centros remitió 25 casos a los demás centros: 15 eran pacientes consecutivos en los que se había realizado una coronariografía y 10 eran no diabéticos, asintomáticos o con dolor precordial atípico y con una probabilidad pretest < 10%. Cada centro corrigió 150 casos: 125 de forma «ciega», y 25 de su propio centro conociendo los datos clínicos.

**Resultados.** Cuatro o más de los 5 centros que evaluaban de forma ciega cada caso coincidieron en un diagnóstico positivo o negativo en 116 pacientes (78%). La media de los valores kappa obtenidos era 0,48 (rango intercentros, 0,45-0,52). El grado de acuerdo era mayor en la enfermedad de 3 vasos (93%; rango, 85-95%) y de la arteria descendente anterior (83%; rango, 80-86%), y cuando el remitente informaba disineria basal (86%; rango, 82-90%), disineria en el territorio de la descendente anterior (81%; rango, 76-84%) o índice de motilidad segmentaria pico > 1,50 (88%; rango, 85-90%). La sensibilidad, la especificidad y la precisión diagnóstica media de la técnica «a ciegas» para estenosis coronaria  $\geq 50\%$  en  $\geq 1$  vaso fue del 68, el 66 y el 67%, respectivamente, pero con una amplia variabilidad intercentros.

**Conclusiones.** La concordancia intercentros de la ecocardiografía de ejercicio es moderada. La precisión diagnóstica «a ciegas» es inferior a la comunicada cuando se conocen los datos clínicos.

**Palabras clave:** *Ecocardiografía de ejercicio. Acuerdo intercentros. Precisión diagnóstica.*

VÉASE EDITORIAL EN PÁGS. 9-11

Estudio financiado por la Red Cardiovascular RECAVA.

Correspondencia: Dr. J.C. Peteiro.  
P.º Ronda, 5, 4.º izqda. 15011 A Coruña. España.  
Correo electrónico: pete@canalejo.org

Recibido el 9 de junio de 2005.

Aceptado para su publicación el 18 de octubre de 2005.

## Agreement Between Centers on the Interpretation of Exercise Echocardiography

**Introduction and objectives.** A low-to-moderate level of agreement on the interpretation of dobutamine echocardiography has been reported, but there are no similar findings on exercise echocardiography. The objectives of this study were to assess the level of agreement between centers on the use of exercise echocardiography and to evaluate the accuracy of the technique when used in a blinded manner.

**Patients and method.** Six institutions with experience in exercise echocardiography each sent 25 study results to the other centers. Of these, 15 were positive or negative studies on consecutive patients undergoing coronary angiography, and 10 were on non-diabetic patients who had non-coronary chest pain or were asymptomatic and whose pretest probability of coronary artery disease was < 10%. Each institution evaluated 150 studies: 125 blinded and 25 of their own with knowledge of clinical data.

**Results.** For 116 patients (78%), four or more of the five centers blindly evaluating each study agreed with the positive or negative result. The average kappa coefficient was 0.48 (intercenter range 0.45-0.52). The percentage agreement was higher with three-vessel disease (93%, range 85%-95%), with left anterior descending coronary artery disease (83%, range 80%-86%), and when the referring institution reported baseline dyssynergy (86%, range 82%-90%), dyssynergy in left anterior descending coronary artery territory (81%, range 76%-84%), or a peak wall motion score index >1.50 (88%, range 85%-90%). When the technique was used blinded to detect =50% coronary narrowing in =1 vessel, its sensitivity, specificity and accuracy were 68%, 66% and 67%, respectively, with wide variability between centers.

**Conclusions.** There was moderate agreement between centers on the interpretation of exercise echocardiography. When used blinded, the technique's accuracy was lower than that reported when clinical data is known.

**Key words:** *Exercise echocardiography. Intercenter agreement. Accuracy.*

Full English text available at: [www.revespcardiol.org](http://www.revespcardiol.org)

## ABREVIATURAS

EAC: enfermedad arterial coronaria.

ECG: electrocardiograma.

EE: estrés con ejercicio.

## INTRODUCCIÓN

Una de las principales limitaciones de la ecocardiografía de estrés es su variabilidad, sobre todo desde que en el primer estudio de Hoffmann, aunque realizado con imagen fundamental y sin criterios de lectura comunes, se encontrara una concordancia intercentros sólo ligera en la interpretación de la ecocardiografía de estrés con dobutamina<sup>1</sup>, concordancia que mejoró en un estudio subsiguiente del mismo autor, esta vez utilizando imagen armónica y criterios de lectura comunes<sup>2</sup>.

Sin embargo y sorprendentemente, aunque la ecocardiografía de estrés con ejercicio (EE) es la forma de estrés más antigua<sup>3</sup>, sensible y segura<sup>4,5</sup>, así como la más utilizada en nuestro medio<sup>6</sup>, no se ha llevado a cabo ningún estudio que investigue la concordancia intercentros con esta técnica.

Por lo tanto, el propósito de este estudio fue evaluar: *a)* la concordancia intercentros de la EE, y *b)* la sensibilidad, la especificidad y la precisión diagnóstica de la técnica realizada de forma «ciega».

## PACIENTES Y MÉTODO

Participaron 6 centros con amplia experiencia en ecocardiografía de estrés y, en particular, en EE (entre 1.000 y 7.000 EE realizados). Cada uno de los 6 centros envió 25 casos. Se exigía que 15 de ellos correspondieran a pacientes consecutivos con EE positivo o negativo y que tuvieran realizada una coronariografía dentro de los 3 meses de la EE; y que los otros 10 estudios correspondieran a estudios realizados también consecutivamente en pacientes no diabéticos, asintomáticos o con dolor precordial atípico y con una probabilidad pretest de enfermedad arterial coronaria (EAC) < 10% según el sexo, la edad y los factores de riesgo<sup>7</sup>. Por lo tanto, cada centro corregía 150 casos: 125 de forma «ciega» correspondientes a los casos de los otros centros, y los 25 de su propio centro conociendo los datos clínicos.

Se utilizaron aparatos de última generación con imagen de segundo armónico y paquete de estrés (Sonos-5500, Phillips por 4 centros y Vivid-5, GE por 2 centros). Cada estudio se envió al centro coordinador en disco óptico, desde donde se distribuyó a los demás centros, bien en el mismo formato o bien en formato de vídeo, según la disponibilidad de lectura de los distintos centros. Para la lectura se comparaban los planos apicales de 4 y de 2 cámaras y los paraesternales

longitudinal y transversal, en reposo y en ejercicio en formato «cuad».

## Criterios de lectura

Se utilizaron criterios de lectura comunes<sup>8</sup>. Se definía un EE positivo cuando había al menos 1 segmento anormal en reposo o ejercicio, o bien tardocinesia en caso de que no hubiera alteraciones de la conducción y EE negativo cuando ningún segmento era anormal en reposo o ejercicio, o había hipocinesia aislada del segmento posterobasal y/o del septobasal, salvo que se acompañaran de disinerxia de 1 segmento adyacente.

Cada centro categorizaba cada resultado positivo como necrosis (alteración regional de la motilidad que persistía o mejoraba con el ejercicio), isquemia (alteración de la motilidad con el ejercicio), isquemia + necrosis en el mismo territorio (alteración de la motilidad basal que empeoraba con el ejercicio en el mismo territorio), o isquemia a distancia (alteración de la motilidad en uno o más territorios en situación basal, con aparición de nueva alteración de la motilidad con el ejercicio en un territorio distinto). Se calculó el índice de motilidad segmentaria en reposo y ejercicio en cada lectura dividiendo el ventrículo izquierdo en 16 segmentos<sup>9</sup>. Se determinaban los territorios afectados en cada estudio según fueran dependientes de la arteria descendente anterior (DA), la circunfleja (Cx), la coronaria derecha (CD) o una combinación de las anteriores.

Además, cada centro calificó objetiva y subjetivamente la calidad de cada estudio. De forma objetiva se utilizó un índice de calidad segmentaria en el que se asignaba una puntuación de 3 a cada segmento con buena visibilidad (engrosamiento y desplazamiento), 2 a los que tenían una visibilidad regular, 1 a los que tenían mala visibilidad y 0 a los no visibles. De forma subjetiva, cada estudio se calificó como de calidad buena, regular, mala o no interpretable.

## Análisis estadístico

Se utilizó el paquete estadístico SPSS 12.0. Las variables continuas se describen como media  $\pm$  DE. Las variables categóricas se describen como porcentajes. La comparación entre pacientes con y sin EAC se realizó mediante el test de la  $\chi^2$  para variables categóricas y mediante el test de la t de Student para variables continuas. La concordancia entre 2 centros se estimaba por el grado de acuerdo (EE negativo o positivo) al analizar los estudios de los demás centros sin incluir los casos de los propios centros (150 – 50 casos = 100 casos). Se exponen los porcentajes de concordancia y los valores kappa ( $\kappa$ ) (proporción de acuerdo superior a la esperada por el azar). Un valor de  $\kappa$  de entre 0 y 0,20 se considera muy bajo; entre 0,21 y 0,40 bajo; entre 0,41 y 0,60, moderado; entre 0,61 y 0,80, bueno, y

**TABLA 1. Características clínicas, medicaciones y electrocardiograma basal**

	EAC (n = 58)	Sin EAC (n = 91)	p
Edad, años, media ± DE	63 ± 10	57 ± 13	< 0,01
Sexo varón, n (%)	47 (81)	48 (53)	< 0,0001
Diabetes mellitus, n (%)	20 (34)	6 (7)	< 0,0001
HTA, n (%)	30 (52)	38 (42)	NS
Colesterol, n (%)	33 (58)	20 (22)	< 0,0001
Tabaquismo, n (%)	28 (48)	11 (12)	< 0,0001
Historia familiar, n (%)	2 (3)	7 (8)	NS
IAM previo, n (%)	24 (41)	12 (13)	< 0,0001
ACTP/cirugía, n (%)	15 (26)	14 (15)	NS
Nitritos, n (%)	21 (36)	10 (11)	< 0,0001
Antagonistas del calcio, n (%)	8 (14)	10 (11)	NS
Bloqueadores beta, n (%)	25 (43)	13 (14)	< 0,0001
IECA, n (%)	21 (36)	21 (23)	NS
Digoxina, n (%)	1 (2)	0 (0)	NS
Fibrilación auricular, n (%)	1 (2)	2 (2)	NS
BRIHH, n (%)	2 (3)	5 (5)	NS

ACTP: angioplastia coronaria transluminal percutánea; BRIHH: bloqueo completo de rama izquierda del haz de His; DE: desviación estándar; EAC: enfermedad arterial coronaria; HTA: hipertensión arterial; IAM: infarto agudo de miocardio; IECA: inhibidores de la enzima de conversión de la angiotensina.

entre 0,80 y 1,0, excelente<sup>10</sup>. Se calcularon la sensibilidad, la especificidad y la precisión diagnóstica de cada centro en la corrección de sus propios casos, así como en la corrección a «ciegas» de los casos de los otros centros. La sensibilidad se define como el porcentaje de casos con EE positivo entre los pacientes con estenosis coronaria significativa en al menos 1 vaso. La especificidad se define como el porcentaje de casos con EE negativo entre los pacientes sin lesiones coronarias demostradas angiográficamente o bien con una baja probabilidad pretest. La precisión diagnóstica se define como el porcentaje de aciertos (casos con EE positivo y EAC, más casos con EE negativo y ausencia de EAC) entre el global de pacientes.

## RESULTADOS

Se dispuso para análisis de 149 estudios (1 estudio fue excluido por imágenes deficientes). Se utilizaron agentes de contraste para la opacificación del ventrículo izquierdo en 9 estudios (6%) y el estudio de estrés se realizó con imagen en pico de ejercicio en 124 casos (83%).

### Características clínicas basales y respuesta al ejercicio

En 58 pacientes (39%) se observaba una EAC significativa definida por estenosis  $\geq 50\%$  en  $\geq 1$  arteria coronaria, rama principal o puente coronario, mientras que 91 pacientes (61%) tenían EAC no significativa angiográficamente demostrada (n = 37), o bien baja

**TABLA 2. Respuesta al ejercicio**

	EAC	Sin EAC	p
MET, media ± DE	8 ± 2	10 ± 3	< 0,0001
Frecuencia cardíaca máxima, lat/min, media ± DE	127 ± 24	144 ± 20	< 0,0001
Clínica (+), n (%)	14 (24%)	7 (8%)	< 0,01
ECG (+), n (%)	20 (34%)	14 (15%)	< 0,001
ECG no diagnóstico, n (%)	18 (31%)	16 (18%)	< 0,001
< 85% frecuencia cardíaca máxima teórica, n (%)	32 (55%)	28 (31%)	< 0,01

DE: desviación estándar; MET: equivalentes metabólicos; EAC: enfermedad arterial coronaria; ECG: electrocardiograma.

probabilidad pretest según la definición previa (n = 54). En los pacientes con EAC había enfermedad de 1 vaso en 24, de 2 vasos en 18, y de 3 vasos en 16. La arteria DA estaba estenosada en 40 pacientes, la CD en 39 y la Cx en 29. En la tabla 1 se muestran las características clínicas basales, la medicación y los datos del electrocardiograma (ECG) basal en los pacientes con y sin EAC. En la tabla 2 se muestran los datos de respuesta al ejercicio en los pacientes con y sin EAC.

### Calidad de las imágenes

La calificación subjetiva de la calidad de los estudios difería significativamente entre los distintos centros. Había centros que calificaban un alto porcentaje de estudios como buenos ( $\geq 80\%$  de estudios), mientras que otros sólo los consideran buenos en menos de la mitad de los casos, así como no interpretables en un porcentaje de ellos que osciló entre el 0 y el 8% (fig. 1). Se observaban las mismas diferencias en cuanto al índice de calidad segmentaria calculado por los distintos centros (fig. 2). En general, los centros que calificaban peor a los demás tendían a tener las imágenes de mejor calificación según los otros centros.

### Concordancia

Cuatro o más de los 5 centros que evaluaban de forma ciega cada caso coincidieron en un diagnóstico positivo de EAC en 51 pacientes y en un diagnóstico negativo en 65 pacientes, por lo que hubo concordancia en un total de 116 de los 149 pacientes (78%). En 4,1 ± 0,9 centros de los 5 centros hubo coincidencia en un diagnóstico positivo o negativo de EAC. La media de los valores  $\kappa$  obtenidos entre los diferentes centros era de 0,48, con un rango de media de  $\kappa$  intercentros de 0,45 a 0,52. El porcentaje de acuerdo y los valores  $\kappa$  en distintos escenarios se muestran en la tabla 3. El porcentaje de acuerdo y el índice  $\kappa$  diferían según las anomalías regionales de la contractilidad del diagnóstico del centro remitente, y el grado de acuerdo era mayor cuando el centro remitente evidenciaba anomalías basales de la contractilidad regional en algún terri-

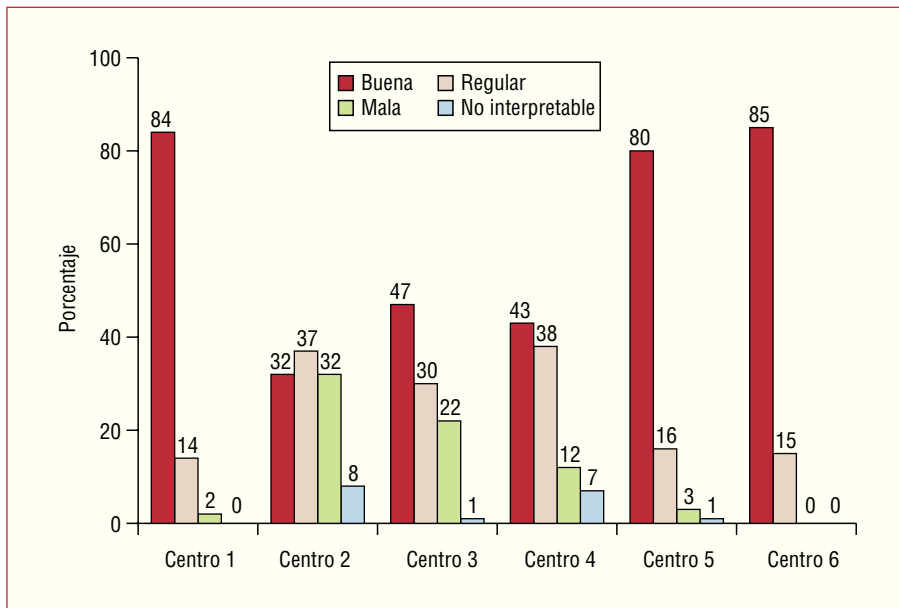


Fig. 1. Porcentaje de estudios calificados como buenos, regulares, malos y no interpretables según los diferentes centros.

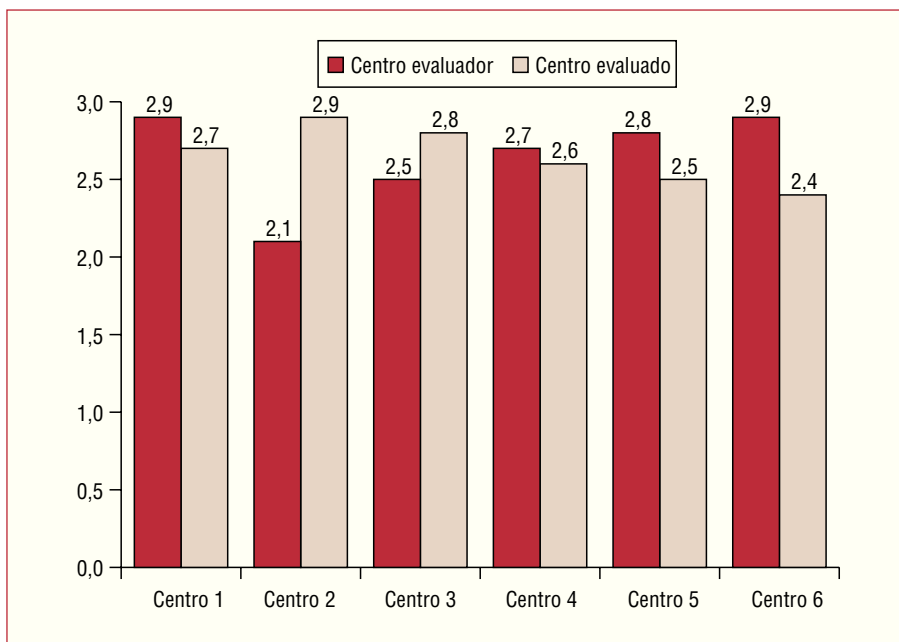


Fig. 2. Puntuación de la calidad de los estudios de los otros centros según el centro remitente (columnas en blanco) y puntuación de la calidad de los estudios de cada centro según los demás centros (columnas en negro).

torio, anomalías de la contractilidad en reposo y/o ejercicio en el territorio de la arteria DA, o cuando informaba peores índices de motilidad segmentaria en ejercicio (tabla 4).

### Sensibilidad, especificidad y precisión diagnóstica

El porcentaje de lecturas positivas y negativas, así como la sensibilidad, la especificidad y la precisión diagnóstica «a ciegas» diferían entre los distintos centros (fig. 3). Había 2 centros con alta sensibilidad pero baja especificidad y 1 en el que ocurría lo contrario.

La media de sensibilidad, especificidad y precisión diagnóstica de los 6 centros para estenosis en al menos 1 vaso  $\geq 50\%$  (según estimación visual) era del 68, el 66 y el 67%, respectivamente. La media de la sensibilidad y especificidad de los distintos centros era similar en pruebas superiores o inferiores a las submáximas (el 68 frente al 64% y el 66 frente al 65%, respectivamente). Estos datos contrastan con la media de sensibilidad, especificidad y precisión diagnóstica de cada centro cuando corregía sus propios casos (fig. 4). Si consideráramos que los casos positivos para enfermedad coronaria eran los que tenían una estenosis  $\geq 50\%$  en al menos 1 vaso, o bien antecedentes

**TABLA 3. Porcentaje de acuerdo e índices kappa según el resultado de la coronariografía y la calidad de las imágenes ecocardiográficas**

	Porcentaje de acuerdo (rango de la media intercentros)	$\kappa$ (rango de la media intercentros)
3 vasos (n = 16)	93 (85-95)	<sup>b</sup>
2 vasos (n = 18)	63 (48-71)	0,32 (0,20-0,46)
1 vaso (n = 24)	76 (67-82)	0,52 (0,37-0,62)
0 vasos (n = 37) <sup>a</sup>	71 (69-78)	0,42 (0,32-0,57)
Enfermedad de la DA (± otros vasos) (n = 40)	83 (80-86)	0,57 (0,46-0,63)
Enfermedad de la CD/Cx (± DA) (n = 46)	74 (72-76)	0,39 (0,31-0,45)
Sin EAC (n = 91)	71 (68-75)	0,37 (0,32-0,40)
Buena calidad del estudio (n = 68)	73 (69-76)	0,46 (0,41-0,53)
Calidad subóptima (n = 81)	75 (74-77)	0,51 (0,48-0,53)

CD: arteria coronaria derecha; Cx: arteria circunfleja; DA: arteria descendente anterior; EAC: enfermedad arterial coronaria.

<sup>a</sup> Demostrados en la coronariografía.

<sup>b</sup> No se expone el valor de  $\kappa$ , porque es incalculable en alguna de las comparaciones de un centro con otro, como sucede cuando uno o ambos centros dan sólo 1 de los 2 diagnósticos posibles (EAC o normal) para todos los casos analizados.

de infarto agudo de miocardio (IAM) y disinergia basal según el diagnóstico del centro remitente, la sensibilidad, la especificidad y la precisión diagnóstica en la lectura ciega eran similares a las obtenidas con el único criterio de estenosis  $\geq 50\%$ : el 69% (rango intercentros, 53-82), el 70% (rango, 49-89) y el 69% (rango, 64-78). La sensibilidad, la especificidad y la precisión diagnóstica según la decisión mayoritaria (4 o más centros cuando corregían 5 centros; 3 o más cuando corregían 4 centros) eran similares para estenosis  $\geq 50\%$  en 1 vaso y para el criterio de estenosis  $\geq 50\%$  o antecedentes de IAM y disinergia basal según el centro remitente (el 72 frente al 73%; el 74 frente al 80% y el 73 frente al 77%, respectivamente;  $p = \text{NS}$ ).

**TABLA 4. Porcentaje de acuerdo e índice kappa según el diagnóstico del centro remitente**

	Porcentaje de acuerdo (rango de la media intercentros)	$\kappa$ (rango de la media intercentros)
Necrosis en la ecocardiografía basal (n = 33)	86 (82-90)	<sup>a</sup>
Ecocardiografía basal normal (n = 116)	71 (66-74)	0,36 (0,28-0,40)
EE: territorio de la DA (± otro territorio) (n = 45)	81 (76-84)	0,45 (0,37-0,57)
EE: territorio de la CD/Cx (± otro territorio) (n = 45)	73 (65-77)	0,39 (0,29-0,47)
EE: sólo territorio de CD/Cx (n = 24)	70 (65-74)	0,35 (0,19-0,43)
IMS ejercicio = 1,00 (n = 80)	72 (64-79)	<sup>a</sup>
IMS ejercicio = 1,06-1,25 (n = 19)	64 (52-68)	<sup>a</sup>
IMS ejercicio = 1,26-1,50 (n = 22)	75 (62-85)	<sup>a</sup>
IMS ejercicio > 1,50 (n = 28)	88 (85-90)	<sup>a</sup>

CD: arteria coronaria derecha; Cx: arteria circunfleja; DA: arteria descendente anterior; EAC: enfermedad arterial coronaria; EE: estrés con ejercicio; IMS: índice de motilidad segmentaria.

<sup>a</sup> No se expone el valor de  $\kappa$  porque es incalculable en alguna de las comparaciones de un centro con otro, como sucede cuando uno o ambos centros dan sólo 1 de los 2 diagnósticos posibles (EAC o normal) para todos los casos analizados.

### Lecturas falsas positivas

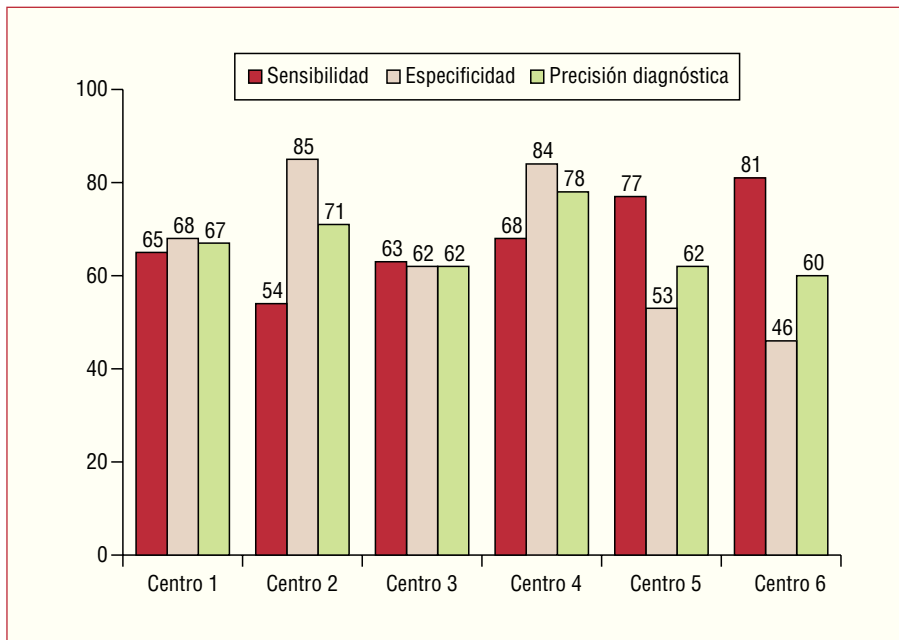
En 403 lecturas correspondientes a 83 pacientes sin estenosis coronaria, IAM previo ni disinergia basal según el centro remitente hubo 124 interpretaciones falsas (+) que corresponden a 31% de las correcciones sin EAC, con un amplio rango intercentros (11-51%). Estas lecturas falsas (+) lo eran en su mayoría por asignar alteraciones contráctiles al territorio de la CD (el 36% de las lecturas) o de la DA (el 35%), y menos frecuentemente al territorio de la Cx (el 10%) o a varios territorios (el 19%). El índice de motilidad segmentaria (IMS) medido por los centros correctores en estos casos era de  $1,1 \pm 0,2$  en reposo y de  $1,3 \pm 0,2$  en ejercicio.

### Lecturas falsas negativas

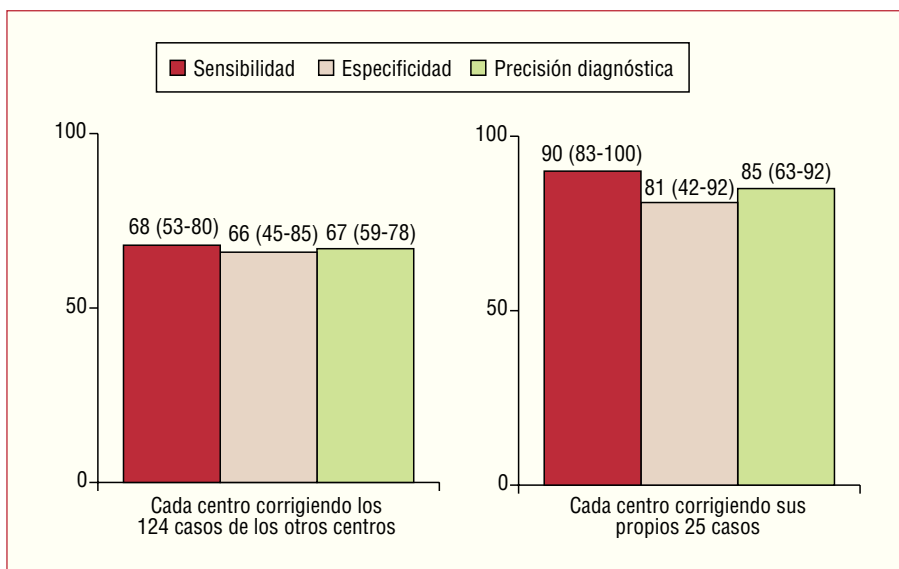
En 319 lecturas correspondientes a 66 pacientes que tenían estenosis coronarias demostradas en la angiografía, o ausencia de CG o no lesiones coronarias en ésta, pero con antecedentes de IAM por la historia clínica y disinergia según el centro remitente, hubo 102 interpretaciones falsas (-) que corresponden al 32% de las correcciones con EAC (rango intercentros, 18-47%). En la mayoría de estos casos había enfermedad de sólo 1 vaso (45%; enfermedad de la DA en 23 de ellos) o de 2 vasos (34%), y en menos ocasiones se trataba de enfermedad de 3 vasos (9%) o de 0 vasos (13%). El centro remitente informaba disinergia en 32 de estos 66 pacientes (48%), que era severa (IMS > 1,50) en 12 de ellos (18%).

### DISCUSIÓN

El principal interés de este estudio radica en que es el primero en que se ha evaluado el acuerdo intercentros de la ecocardiografía de ejercicio. Los principales hallazgos fueron los siguientes: a) la concordancia intercentros de la ecocardiografía de ejercicio es mode-



**Fig. 3.** Sensibilidad, especificidad y precisión diagnóstica de cada uno de los centros que evaluaban, de forma ciega, los casos referidos por los demás centros.



**Fig. 4.** Sensibilidad, especificidad y precisión diagnóstica para detectar lesiones coronarias con  $\geq 50\%$  de estenosis de la ecocardiografía de ejercicio interpretada a ciegas y de la ecocardiografía de ejercicio interpretada con el conocimiento de los datos clínicos y la respuesta al ejercicio. Se muestra la media de los centros, así como el rango intercentros.

rada, y b) la sensibilidad, la especificidad y la precisión diagnóstica de la técnica realizada de forma «ciega» son inferiores a las comúnmente comunicadas cuando se conocen las características basales y la respuesta al ejercicio de los pacientes.

### Acuerdo intercentros de la ecocardiografía de estrés

Aunque Hoffmann et al estudiaron el acuerdo intercentros de la ecocardiografía de estrés con dobutamina<sup>1,2</sup>, no hay estudios similares sobre la ecocardiografía de estrés con ejercicio a pesar de ser más utilizada, sensible y segura<sup>4,6</sup>. En el primer estudio de

Hoffmann, realizado con imagen fundamental y sin criterios de lectura comunes, se observó un acuerdo ligero ( $\kappa = 0,37$ )<sup>1</sup>, mientras que en el segundo estudio, realizado con imagen armónica y criterios de lectura comunes, el acuerdo era moderado ( $\kappa = 0,55$ )<sup>2</sup>. La mejoría del acuerdo parecía ser debida tanto a la imagen armónica como a la unificación de los criterios de lectura, puesto que el grado de acuerdo sobre los mismos pacientes estudiados con imagen fundamental era mayor que en el primer estudio de Hoffmann. Nosotros hemos utilizado los mismos criterios de lectura que en el segundo estudio de Hoffmann<sup>2</sup>, lo que en general no ha supuesto ninguna modificación de la práctica clínica habitualmente seguida en cada centro. Cabría pen-

sar que el grado de acuerdo de la ecocardiografía de ejercicio debería ser menor que en la realizada con dobutamina, toda vez que las imágenes deberían ser de mayor calidad con la segunda técnica. Sin embargo, con criterios de lecturas comunes e imagen armónica, el grado de acuerdo fue moderado, con un valor medio de  $\kappa = 0,48$ , mejor que el del primer estudio de Hoffmann y similar al del segundo estudio de este autor. El grado de acuerdo fue mayor en la enfermedad de 3 vasos, en la enfermedad de la arteria descendente anterior, cuando había alteraciones basales de la contractilidad regional y cuando el centro remitente informaba disinerxia en el territorio de la DA o bien disinerxia importante. La existencia de un grado de acuerdo superior en estas circunstancias tiene importancia clínica diagnóstica y pronóstica, pues son los pacientes con estas características los que tienen peor pronóstico<sup>11-13</sup>.

### Acuerdo intercentros de otras técnicas diagnósticas

La preocupación por el grado de acuerdo no es exclusiva de la ecocardiografía de estrés. Con otras técnicas se han observado distintos grados de variabilidad en la interpretación. Así, se han comunicado niveles muy bajos de acuerdo en la interpretación del supra-desnivel ( $\kappa = 0,05$ ) o del infradesnivel del segmento ST ( $\kappa = 0,38$ ) entre 2 centros en pacientes con síndrome coronario agudo<sup>14</sup>. Los estudios de perfusión miocárdica con medicina nuclear también presentan dificultades de interpretación, puesto que esta técnica es asimismo subjetiva y, como en la ecocardiografía de estrés, en ella pueden influir la experiencia del observador y la calidad de las imágenes. Con imagen de talio se ha descrito una concordancia moderada-alta, con valores de  $\kappa$  que oscilaban entre 0,56 y 0,74 en 2 estudios<sup>15,16</sup>. Sin embargo, en un estudio multicéntrico en el que participaron 25 hospitales, la concordancia de los distintos centros sin uniformidad de criterios de lectura fue baja ( $\kappa = 0,27$ )<sup>17</sup>. En un estudio de Candell-Riera et al<sup>18</sup> se observó una concordancia buena con tomografía de perfusión de ejercicio con tetrofosmina, con índices  $\kappa$  entre 0,62 y 0,70 según se evaluaran imágenes topográficas o mapas polares<sup>18</sup>. En este mismo estudio también se observó que la sensibilidad del informe a «ciegas» fue significativamente inferior que la comunicada cuando se conocían los datos clínicos de los pacientes.

### Sensibilidad, especificidad y precisión diagnóstica

Aunque damos valores de sensibilidad y especificidad media de los distintos centros, resulta más informativa como hecho la variabilidad entre ellos en la interpretación «a ciegas». Sin embargo, para cada centro, la sensibilidad, la especificidad y la precisión

diagnóstica de la técnica realizada de forma «ciega» fueron inferiores a las que había cuando cada centro corregía sus propios casos conociendo las características basales y de respuesta al ejercicio de los pacientes. Éste no fue un hallazgo sorprendente, pero nos da una idea de hasta dónde llega la técnica cuando no se conocen la probabilidad pretest, la respuesta hemodinámica, clínica y ECG al ejercicio. Está claro que la EE puede y debe usarse en la práctica clínica, pero no de forma ciega.

### Limitaciones

El formato de lectura de los estudios fue igual para los distintos centros (imágenes apicales basales y de ejercicio, e imágenes paraesternales basales y de ejercicio), pero no así la calidad pues, dependiendo de la posibilidad de lectura de discos ópticos de los distintos centros, se grababa el estudio en vídeo o se enviaba en el propio disco óptico. Sin embargo, el porcentaje de acuerdo y el índice  $\kappa$  fueron similares para estudios de calidad óptima y subóptima. No se envió el estudio completo grabado en vídeo como se hizo en el estudio de Hoffmann et al<sup>1</sup>. Este hecho podría sobrestimar la concordancia, toda vez que el explorador tiende a adquirir y almacenar las imágenes que considera más representativas teniendo en cuenta otras características de la prueba distintas de la imagen. Un 25% de los pacientes recibía tratamiento con bloqueadores beta y hasta 40% de las pruebas fueron inferiores a las submáximas. Este hecho podría haber subestimado la sensibilidad, aunque en este estudio no hemos observado una mayor sensibilidad en las pruebas superiores a las submáximas.

### BIBLIOGRAFÍA

- Hoffmann R, Lethen H, Marwick T, Arnesi M, Fioretti P, Pingitore A, et al. Analysis of interinstitutional observer agreement in interpretation of dobutamine stress echocardiograms. *J Am Coll Cardiol*. 1996;27:330-6.
- Hoffmann R, Marwick TH, Poldermans D, Lethen H, Ciani R, Van der Meer P, et al. Refinements in stress echocardiographic techniques improve interinstitutional agreement in interpretation of dobutamine stress echocardiograms. *Eur Heart J*. 2002;23:821-9.
- Wann LS, Faris JV, Childress RH, Weyman AE, Feigenbaum H. Exercise cross-sectional echocardiography in ischemic heart disease. *Circulation*. 1979;60:1300-8.
- Beleslin BD, Ostojic M, Stepanovic J, Djordjevic-Dikic A, Stojkovic S, Nedeljkovic M, et al. Stress echocardiography in the detection of myocardial ischemia. Head-to-head comparison of exercise, dobutamine, and dipyridamole tests. *Circulation*. 1994; 90:1168-76.
- Acquatella H. Ecocardiografía de estrés en Latinoamérica. Revisión de 5 años (1997-2002). *Rev Esp Cardiol*. 2003;56:21-8.
- Rodríguez García MA, Iglesias-Garriz I, Corral Fernández F, Garrrote Coloma C, Alonso-Orcajo N, Branco L, et al. Evaluación de la seguridad de la ecocardiografía de estrés en España y Portugal. *Rev Esp Cardiol*. 2001;54:941-8.

7. Prevention of coronary heart disease in clinical practice. Recommendations of the Second Joint Task Force of European and other Societies on Coronary Prevention. *Eur Heart J.* 1998;19:1434-503.
8. Hoffmann R, Lethen H, Marwick T, Rambaldi R, Fioretti P, Pingitore A, et al. Standardized guidelines for the interpretation of dobutamine echocardiography reduce inter-institutional variance in interpretation. *Am J Cardiol.* 1998;82:1520-4.
9. Bourdillon PD, Broderick TM, Sawada SG, Armstrong WF, Ryan T, Dillon JC, et al. Regional wall motion index for infarct and noninfarct regions after reperfusion in acute myocardial infarction: comparison with global wall motion index. *J Am Soc Echocardiogr.* 1989;9:398-407.
10. Fleiss JL. The measurement of interrater agreement. En: Fleiss JL, editor. *Statistical methods for rates and proportions.* New York: John & Sons; 1981. p. 212-36.
11. Arruda AM, Das MK, Roger VL, Klarich KW, Mahoney DW, Pellikka PA. Prognostic value of exercise echocardiography in 2,632 patients  $\geq$  65 years of age. *J Am Coll Cardiol.* 2001;37:1036-41.
12. Elhendy A, Mahoney DW, Khandheria BK, Paterick TE, Burger KN, Pellikka PA. Prognostic significance of the location of wall motion abnormalities during exercise echocardiography. *J Am Coll Cardiol.* 2002;40:1623-9.
13. Peteiro J, Monserrat L, Mariñas J, Garrido I, Bouzas M, Muñiz J, et al. Valor pronóstico de la ecocardiografía de ejercicio en cinta rodante. *Rev Esp Cardiol.* 2005;58:924-33.
14. Holmvang L, Hasbak P, Clemmensen P, Wagner G, Grande P. Differences between local investigator and core laboratory interpretation of the admission electrocardiogram in patients with unstable angina pectoris or non-Q-wave myocardial infarction (a Thrombin Inhibition in Myocardial Ischemia [TRIM] substudy). *Am J Cardiol.* 1998;82:54-60.
15. Okada RD, Boucher CA, Kirshenbaum HK, Kushner FG, Strauss HW, Block PC, et al. Improved diagnostic accuracy of thallium-201 stress test using multiple observers and criteria derived from interobserver analysis of variance. *Am J Cardiol.* 1980;46:619-24.
16. Atwood JE, Jensen D, Froelicher V, Witzum K, Gerber K, Gilpin E, et al. Agreement in human interpretation of analog thallium myocardial perfusion images. *Circulation.* 1981;64:601-9.
17. Wackers FJ, Bodenheimer M, Fleiss JL, Brown M. Factors affecting uniformity in interpretation of planar thallium-201 imaging in a multicenter trial. The Multicenter Study on Silent Myocardial Ischemia (MSSMI) Thallium-201 Investigators. *J Am Coll Cardiol.* 1993;21:1064-74.
18. Candell-Riera J, Santana-Boado C, Bermejo B, Armadans L, Castell J, Casás I, et al. Impacto de los datos clínicos y concordancia interhospitalaria en la interpretación de la tomografía miocárdica de perfusión. *Rev Esp Cardiol.* 1999;52:892-7.