

Sensibilidad y especificidad de los criterios OARSI de mejoría para artrosis: el efecto de la utilización de tres diferentes medidas de dolor

I. Villanueva^a, M.M. Guzmán^a, F.J. Toyos^b, R. Ariza^b y F. Navarro^b

^aUniversity of Arizona Arthritis Center. Tucson. EE.UU. ^bServicio de Reumatología. Hospital Universitario Virgen Macarena. Sevilla. España.

Objetivo: Evaluar la validez de los criterios recientemente propuestos de mejoría para artrosis (OARSI-RCI) en comparación con el grado de mejoría percibido por el paciente, analizando el efecto del uso de tres medidas diferentes de dolor.
Métodos: Se seleccionó a 73 pacientes con artrosis de rodilla para participar en un ensayo clínico aleatorizado, para la comparación de dos antiinflamatorios no esteroideos (AINE). Se aplicaron las siguientes medidas: subescalas de dolor y función de los índices WOMAC y Lequesne y valoración global del dolor por parte del paciente en una escala visual analógica (EVA). En la visita final los pacientes evaluaron el grado de mejoría percibido (MPP) en una escala categórica (mucho peor, peor, algo peor, igual, algo mejor, mejor, mucho mejor). Se analizó la sensibilidad (SE) y especificidad (ES) de los criterios mediante su comparación con la mejoría percibida por el paciente utilizando tres medidas diferentes de dolor.

Resultados: Cuando el estándar de medida del dolor fue la EVA, la SE y ES fueron: *a)* SE = 0,60, ES = 0,74 frente a MPP de al menos «algo mejor»; *b)* SE = 0,61, ES = 0,58 frente a MPP de al menos «mejor», y *c)* SE = 0,83, ES = 0,56 frente a MPP de «mucho mejor». Cuando se utilizó la subescala de dolor del índice WOMAC la SE y ES fueron: *a)* SE = 0,47, ES = 0,83 frente a MPP de al menos «algo mejor»; *b)* SE = 0,61, ES = 0,74 frente a MPP de al menos «mejor», y *c)* SE = 0,83, ES = 0,69 frente a MPP de «mucho mejor». Y finalmente, cuando se utilizó el índice de Lequesne la SE y ES fueron: *a)* SE = 0,45, ES = 0,83 frente a MPP de al menos

«algo mejor»; *b)* SE = 0,67, ES = 0,79 frente a MPP de al menos «mejor», y *c)* SE = 0,83, ES = 0,71 frente a MPP de «mucho mejor».

Conclusión: En general los criterios mostraron valores de SE y ES de moderados a buenos independientemente de la medida de dolor utilizada.

Palabras clave: Criterios de respuesta. Ensayos clínicos. Artrosis.

Sensitivity and specificity of the Osteoarthritis Research Society International Responder Criteria Initiative (OARSI-RCI) for osteoarthritis: the effect of three different pain measures

Aim: To evaluate the validity of the recently proposed OARSI-RCI for osteoarthritis (OA) compared with patient's perceived improvement (PPI), using three different pain measures.

Methods: Seventy-three patients with knee OA were recruited in a double-blind, randomized, clinical trial comparing two non-steroidal anti-inflammatory drugs (NSAIDs). The following measures were applied at baseline and after treatment: pain and physical function subscales of the WOMAC and the Lequesne Index, pain on a visual analogue scale (VAS) and overall patient assessment on a VAS. At the end of treatment, PPI was evaluated using a 7-point scale (much worse, worse, somewhat worse, the same, somewhat better, better, much better). The sensitivity (SE) and specificity (SP) of the OARSI criteria were evaluated by comparison with the PPI using the three different pain measures.

Results: When the VAS was used: *a)* SE = 0.60, SP = 0.74 when PPI was at least «somewhat better»; *b)* SE = 0.61, SP = 0.58 when PPI was «better»; and *c)* SE = 0.83, SP = 0.56 when PPI was «much better». When the WOMAC subscale was used: *a)* SE = 0.47, SP = 0.83 when PPI was at least «somewhat better»; *b)* SE = 0.61, SP = 0.74 when PPI was at

Correspondencia: Dr. I. Villanueva-Torrecillas.
Assistant Research Scientist.
University of Arizona Arthritis Center.
University Medical Center. Room 7341A.
1501 N Campbell Ave. Tucson, AZ 85724-5093. EE.UU.
Correo electrónico: isidro@u.arizona.edu

Manuscrito recibido el 30-9-2002 y aceptado el 19-12-2002.

least «better»; and *c*) SE = 0.83, SP = 0.69 when PPI was «much better». Finally, for the Lequesne subscale: *a*) SE = 0.45, SP = 0.83 when PPI was at least «somewhat better»; *b*) SE = 0.67, SP = 0.79 when PPI was at least «better»; and *c*) SE = 0.83, SP = 0.71 when PPI was «much better».

Conclusion: In general, the OARSI criteria show moderate-to-high sensitivity and specificity in detecting different degrees of improvement in patients with knee OA, independently of the pain measures used.

Key words: Response criteria. Clinical trials. Osteoarthritis.

Introducción

La artrosis es la enfermedad reumática más frecuente y una de las enfermedades crónicas que producen mayor grado de incapacidad funcional¹⁻⁴. En mujeres mayores de 65 años puede ser la enfermedad crónica más frecuente⁴; aun así, se espera que la prevalencia de la enfermedad aumente debido al envejecimiento progresivo de la población. En décadas pasadas ha habido gran variabilidad en las medidas utilizadas en los ensayos clínicos para la valoración de la eficacia de los fármacos antirreumáticos, no sólo en la selección de las variables sino también en el tipo de escalas e instrumentos utilizados para su medición⁵. En la conferencia OMERACT III (Outcome Measures in Rheumatology, 1997) se propuso un núcleo mínimo de medidas que debería ser incluido, junto con otras a juicio del investigador, en todos los ensayos clínicos de artrosis: dolor, función física, valoración global por parte del paciente y, en estudios de al menos un año de duración, valoración radiográfica⁶. Otras medidas recomendadas son: calidad de vida y valoración global por el médico.

Sin embargo, hasta ahora no se ha dispuesto de ninguna definición de mejoría o de respuesta al tratamiento para artrosis similar a los criterios ACR20, ACR50 o ACR70 definidos para artritis reumatoide y basados en el núcleo de medidas propuestas en OMERACT⁷. Recientemente el comité para ensayos clínicos de la Osteoarthritis Research Society International (OARSI) ha propuesto una serie de criterios para la definición de respuesta en el contexto de ensayos clínicos de artrosis, los llamados OARSI Response Criteria Initiative (RCI)⁸. Según estos criterios de respuesta, un paciente con artrosis de rodilla y en tratamiento con un AINE es clasificado como respondedor cuando se observa ≥ 45% de mejoría en dolor (siempre y cuando se alcance mejoría ≥ 20 unidades normalizadas [UN] en la escala de dolor) o cuando se cumplen dos de las tres siguientes condiciones: *a*) mejoría del dolor ≥ 15%

TABLA 1. Características de los participantes en visita basal

	(n = 73)
Edad*	57,3 (35-73)
Sexo (varón/mujer)	8/65
Duración artrosis (meses) ^a	74 (3-264)
Escalas del índice WOMAC	
Dolor (0 a 50) ^b	26,82 (7,65)
Rigidez (0 a 20) ^b	11,74 (4,02)
Función física (0 a 170) ^b	86,87 (28,03)
Dolor (EVA) ^b	66,68 (20,81)
Valoración global del paciente (EVA) ^b	68,01 (15,39)
Valoración global del médico (EVA) ^b	58,67 (14,18)

*Media (intervalo); ^bmedia (desviación estándar).

(≥ 10 UN); *b*) mejoría de la función física ≥ 30% (≥ 15 UN); *c*) mejoría de la valoración global por parte del paciente ≥ 35% (≥ 10 UN).

El objetivo de este estudio fue evaluar la validez de los criterios OARSI-RCI, utilizando como patrón el grado de mejoría expresado por el paciente al final del período de tratamiento. Otro objetivo fue evaluar el impacto de la utilización de tres medidas diferentes de dolor en la definición de los criterios de mejoría.

Material y métodos

Participantes y diseño del estudio

Se seleccionó a 73 pacientes con diagnóstico de artrosis de rodilla, según los criterios del American College of Rheumatology (ACR)⁹, sintomáticos en el momento del cribado, para participar en un ensayo clínico multicéntrico, doble ciego, aleatorizado, de 6 semanas de duración, para la comparación de la seguridad y eficacia de 2 AINE (ketoprofeno frente a dexketoprofeno). Los datos utilizados para este análisis corresponden a los pacientes seleccionados en el Hospital Universitario Virgen Macarena de Sevilla. Las características de los pacientes se exponen en la tabla 1. El estudio fue aprobado por el comité ético de investigación del hospital.

Criterios de inclusión

Los pacientes debían tener una edad comprendida entre los 35 y 75 años; dolor en una rodilla de al menos 40 en una escala visual analógica (EVA) de 0 a 100; una puntuación de 4 a 15 en el índice de Lequesne (siendo 24 el máximo posible para dicho instrumento)¹⁰; evidencia radiográfica de artrosis de rodilla definida como al menos grado 2 según la clasificación de Kellgren y Lawrence¹¹; dolor persistente en rodilla durante un período de reposo de la toma de AINE de dos semanas de duración y com-

TABLA 2. Comparación entre las evaluaciones basal y final

	Basal		6 semanas		p
	Media	DE	Media	DE	
Escalas del índice WOMAC:					
Dolor (0 a 500)	26,82	7,65	20,82	10,82	< 0,001
Rigidez (0 a 200)	11,74	4,02	8,63	5,17	0,002
Función física (0 a 1700)	97,38	28,02	78,25	34,02	< 0,001
Dolor (EVA de 0 a 100)	66,63	20,76	51,24	25,68	< 0,001
Valoración global paciente (EVA)	58,67	14,17	49,60	20,76	< 0,001
Valoración global médico (EVA)	68,01	15,39	51,49	23,49	< 0,001

DE: desviación estándar.

prensión, capacidad y disposición para cumplimentar los cuestionarios. Además de cumplir los criterios de inclusión y exclusión que se citan más abajo, los participantes dieron su consentimiento por escrito, u oral con testigo, para participar en el estudio.

Criterios de exclusión

Fueron excluidos los pacientes que presentaron alguna de las siguientes características: artropatías inflamatorias, metabólicas o neuropáticas; necesidad de bastones o muletas para caminar; enfermedad vascular o neurológica grave que afectara a las extremidades inferiores; diagnóstico de fibromialgia; enfermedades concomitantes graves que pudieran interferir con la evaluación clínica del paciente; historia reciente de trastornos psiquiátricos, y cualquier antecedente que contraindicara la toma de AINE, como hipersensibilidad, enfermedad digestiva potencialmente hemorrágica, úlcera gástrica o duodenal activa, anemia importante, elevación de las enzimas hepáticas de más del doble de la normalidad, bronquitis asmática, hipertensión arterial (HTA) no controlada o insuficiencia renal grave.

Medidas

Se realizó una evaluación clínica basal y al finalizar el período de tratamiento (6 semanas). Fueron determinadas las siguientes medidas de desenlace: las subescalas de dolor, rigidez y dificultad del índice WOMAC (Western Ontario and McMaster Universities Osteoarthritis Index¹²) en formato de escala visual analógica (EVA) de 0 a 100, todas ellas referidas a las 72 h previas a la evaluación; el índice de Lequesne¹⁰ (escalas de dolor, rigidez y dificultad); dolor de rodilla en EVA de 0 a 100; valoración global del paciente y del médico en EVA de 0 a 100 y, finalmente, el grado de mejoría percibido por el paciente al final del período de tratamiento evaluado mediante una escala categórica (tipo Likert) compuesta por 7 puntos nominados con los adjetivos «mucho peor», «peor», «un poco peor», «igual», «un poco mejor», «mejor», «mucho me-

jor». Se pidió al paciente que identificara la rodilla más sintomática en el momento basal y que refiriera a esa rodilla sus evaluaciones posteriores.

Análisis

Tras un proceso de preselección, se realizó una visita de cribado, una basal, una intermedia de seguridad y una final tras las 6 semanas de tratamiento. A efectos de este análisis se consideraron las valoraciones clínicas obtenidas en el momento basal y después del período de tratamiento (6 semanas). Como requisito indispensable para la definición de los criterios de mejoría, se realizó un procedimiento de normalización de todas las puntuaciones de modo que todas se expresaran en una escala de 0 a 100. Como se indicó más arriba, los criterios OARSI definen la mejoría en función de los cambios relativo (porcentaje de mejoría con respecto a evaluación basal) y absoluto (mejoría en unidades absolutas sobre una escala de 0 a 100). Cada paciente del estudio fue clasificado como respondedor o no respondedor en función de la definición de mejoría. Los criterios OARSI-RCI fueron definidos independientemente con cada una de las medidas de dolor citadas más arriba. La validez (de criterio) de los criterios OARSI-RCI se evaluó estimando su sensibilidad (SE) y su especificidad (ES) en comparación con el grado de mejoría percibido por el paciente (MPP) expresado en la escala tipo Likert. Finalmente, se estimó la SE y ES de los criterios para cada una de las medidas de dolor (dolor en EVA, escala de dolor del índice WOMAC, y escala de dolor del índice de Lequesne).

Resultados

Se seleccionó a 73 pacientes (8 varones y 65 mujeres) con una edad media de 57,3 años (intervalo, 35-73) y una duración media de la artrosis de 74 meses desde el momento del diagnóstico. Las características clínicas basales de los participantes se exponen en la tabla 1. Sesenta y tres pacientes completaron la visita final tras el período de seguimiento de seis semanas. De los pacientes que

TABLA 3. Sensibilidad de los criterios en comparación con la mejoría percibida por el paciente en función de las tres medidas de dolor

MPP	EVA dolor	WOMAC dolor	Lequesne dolor
Primer criterio			
Un poco mejor	0,37	0,37	0,42
Mejor	0,50	0,50	0,72
Mucho mejor	1,0	0,83	0,83
Serie completa			
Un poco mejor	0,60	0,47	0,45
Mejor	0,61	0,61	0,67
Mucho mejor	0,83	0,83	0,83

MPP: mejoría percibida por el paciente.

abandonaron, 7 pacientes lo hicieron debido a efectos secundarios del AINE, en todos los casos de tipo gastrointestinal, y 3 por pérdida en el seguimiento tras la visita inicial o la visita de seguridad. Se utilizó el test de rangos de Wilcoxon para la comparación de los valores medios observados en el momento basal y a las 6 semanas. Al final del tratamiento se observó una mejoría estadísticamente significativa en todas las subescalas de dolor, rigidez y dificultad del índice WOMAC, dolor de rodilla en EVA y valoraciones globales de médico y paciente en EVA (tabla 2). A las 6 semanas, 20 (31,7%) participantes consideraron que estaban «un poco mejor», 12 (19%) que estaban «mejor» y 6 (9,5%) «mucho mejor»; de modo que, 38 (60,3%) participantes percibieron algún grado de mejoría con el tratamiento.

Cuando se utilizaron los criterios OARSI-RCI para definir mejoría, las cifras porcentuales correspondientes a pacientes que «respondieron» fueron las siguientes: 17 (27,9%) pacientes cumplieron el primer criterio ($\geq 45\%$ mejoría en dolor) y 29 (47,5%) cumplieron el primero o dos de los tres alternativos. La SE y ES difirieron dependiendo del criterio de mejoría aplicado (primer criterio OARSI frente a la serie completa), del grado de mejoría percibido por el paciente (MMP) y del tipo de medida utilizada para medir el dolor: EVA de 0 a 100, escala de dolor del índice WOMAC o escala de dolor del índice de Lequesne.

De este modo, en pacientes que habían evaluado la mejoría como «un poco mejor», la SE y ES del primer criterio OARSI de respuesta (disminución en dolor $\geq 45\%$) fueron de 0,37 y 0,87, respectivamente, cuando se usó como medida de dolor la EVA de 0 a 100 mm; en pacientes con nivel de MPP de «mejor» los valores fueron: SE = 0,50 y ES = 0,81, y en el grupo «mucho mejor» fueron: SE = 1,0 y ES = 0,80.

Cuando se consideró la definición completa de respuesta (tanto el primer criterio como los alternativos), los valores de SE y ES fueron los siguientes: cuando MPP fue definido como al menos «un poco mejor», SE = 0,60 y ES = 0,74; cuando fue definido

TABLA 4. Especificidad de los criterios en comparación con la mejoría percibida por el paciente en función de las tres medidas de dolor

MPP	EVA dolor	WOMAC dolor	Lequesne dolor
Primer criterio			
Un poco mejor	0,87	0,83	0,78
Mejor	0,81	0,79	0,81
Mucho mejor	0,80	0,76	0,71
Serie completa			
Un poco mejor	0,74	0,83	0,83
Mejor	0,58	0,74	0,79
Mucho mejor	0,56	0,69	0,71

MPP: mejoría percibida por el paciente.

como al menos «mejor», SE = 0,61 y ES = 0,58, y cuando MPP fue definido como «mucho mejor», SE = 0,83 y ES = 0,56,

Cuando se cambió la medida de dolor utilizada para la definición de mejoría (EVA de 0 a 100 frente a subescala de dolor del índice WOMAC o del índice de Lequesne) la SE y ES de los criterios también se modificaron (tablas 3 y 4). La SE y ES fueron muy similares entre sí cuando se usaron las subescalas de dolor de los índices WOMAC y de Lequesne; sin embargo la ES fue ligeramente inferior cuando se utilizó la EVA de dolor.

Discusión

Que sepamos, éste es el primer estudio relacionado con enfermedades reumáticas en el que se compara la validez de unos criterios de respuesta a tratamiento con la percepción de mejoría por parte del paciente. En esta muestra, y en comparación con el grado de mejoría percibido por el paciente, los criterios de respuesta OARSI-RCI mostraron valores de moderados a elevados de sensibilidad y especificidad para detectar diferentes grados de mejoría en pacientes con artrosis de rodilla. Por otra parte, los valores obtenidos de sensibilidad y especificidad fueron, en general, consistentes independientemente del tipo de medida de dolor empleada para la definición de los criterios.

En general, la sensibilidad de los criterios fue aceptable para detectar grados de mejoría más elevados pero relativamente baja para clasificar como respondedores a los pacientes con niveles inferiores de mejoría. Esto podría ser indicativo de que los criterios son relativamente exigentes, si bien la valoración de mejoría por parte del paciente, usada aquí como patrón, no debe ser considerada infalible. Otra lectura posible de estos resultados es que los criterios no abarcan todos los aspectos involucrados en la percepción de bienestar del paciente, que sí estaría mejor reflejada en sus valoraciones propias de mejoría. Lo que sugiere que los criterios podrían beneficiarse de la incorporación de una

medida de percepción de mejoría como la que se ha descrito en este estudio. En detrimento de la validez de una medida de este tipo se puede argumentar que tal valoración puede verse afectada por un sesgo de aquiescencia por parte del paciente, pero es razonable pensar que el potencial sesgo introducido sería de tipo no diferencial (entre los grupos en estudio) y que por tanto no afectaría a los resultados obtenidos.

En conclusión, se ha demostrado la validez de los criterios OARSI-RCI para la valoración de mejoría en pacientes con artrosis de rodilla en tratamiento con AINE, y la consistencia de los criterios independientemente del tipo de medida de dolor utilizada. Son necesarios estudios adicionales para analizar el valor de la incorporación de medidas directas de mejoría en la definición de los criterios.

Bibliografía

1. Felson DT. The epidemiology of osteoarthritis: results from the Framingham osteoarthritis Study. *Semin Arthritis Rheum* 1990;20(Suppl 1):42-50.
2. Felson DT. Epidemiology of hip and knee osteoarthritis. *Epidemiol Rev* 1988;10:1-28.
3. Davis MA, Ettinger WH, Neuhaus JM, Mallon KP. Knee osteoarthritis and physical functioning: evidence from the NHANES I epidemiologic follow-up study. *J Rheumatol* 1991;18: 591-8.
4. Hochberg MC, Kasper J, Williamson J, Skinner A, Fried LP. The contribution of Osteoarthritis to disability: preliminary data from the Women's Health and Aging Study. *J Rheumatol* 1995;(Suppl 43):22:16-8.
5. Bellamy N. Outcome measurement in osteoarthritis clinical trials. *J Rheumatol* 1995;(Suppl 43):22:49-51.
6. Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, et al. Recommendations for a core set of outcome measures for future phase III clinical trials in Knee, hip and hand osteoarthritis. Consensus development at OMERACT III. *J Rheumatol* 1997;24:799-802.
7. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology preliminary criteria definition of improvement in RA. *Arthritis Rheum* 1995;38:727-35.
8. Dougados M, Leclaire P, van der Heijde D, Bloch DA, Bellamy N, Altman RD. A report of the Osteoarthritis Research Society International Standing Committee for Clinical Trials Response Criteria Initiative. *Osteoarthritis Cartilage* 2000;8: 395-403.
9. Altman R, Asch E, Bloch D, Bole G, Bosenstein D, Brandt K, et al. Development criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. *Arthritis Rheum* 1986;29:1039-50.
10. Lequesne MG, Mery C, Samson M, Gerard P. Indexes of severity for osteoarthritis of the hip and knee. Validation value in comparison with other assessment tests. *Scand J Rheumatol* 1987;65(Suppl):85-9.
11. Kellgren JH, Lawrence JS. Radiological assessment of osteoarthritis. *Ann Rheum Dis* 1957;16:494-501.
12. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Still LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;15: 1833-40.