

GUIPCAR: una guía de práctica clínica para el manejo de la artritis reumatoide en España

C. Hernández-García^a y P. Lázaro y de Mercado^b

^aServicio de Reumatología. Hospital Clínico San Carlos. Madrid.

^bTécnicas Avanzadas de Investigación en Servicios de Salud (TAISS). Madrid.

Como ocurre con la mayor parte de las enfermedades, en el manejo de la artritis reumatoide (AR) existe una enorme variabilidad entre países, entre instituciones, entre médicos, e incluso intramédico¹⁻³. Uno de los instrumentos más desarrollados para disminuir la variabilidad y mejorar la práctica médica son las guías de práctica clínica. Las GPC son «aseveraciones desarrolladas sistemáticamente para ayudar al médico y al paciente a tomar decisiones apropiadas en circunstancias clínicas específicas»⁴. Las GPC forman parte de un cambio cultural que consiste en desplazarnos desde la confianza no evaluada en los juicios de los profesionales hacia un soporte más estructurado y fundamentado de las decisiones clínicas.

La Sociedad Española de Reumatología (SER) se propuso en 1998 el reto de crear una guía de práctica clínica que ayudase a los clínicos a tomar decisiones en el manejo de pacientes con AR. Fruto del esfuerzo colectivo de un buen número de personas e instituciones (anexo), en junio de 2001 se publicó la Guía de Práctica Clínica para el Manejo de la Artritis Reumatoide en España (GUIPCAR)⁵. Con esta guía se pretende mejorar la calidad de la asistencia, reducir la variabilidad en el manejo de la AR, avanzar en el manejo integrador de la AR y acercar la práctica clínica a la mejor evidencia científica disponible.

Metodología en el desarrollo de GUIPCAR

En la literatura médica existen algunas directrices sobre distintos tópicos relacionados con el manejo de la AR⁶⁻¹⁰, pero la mayoría de ellas han abordado esta tarea desde una metodología de consenso o panel de expertos. Lo que hace a GUIPCAR distinta de otras directrices o guías es la metodología empleada en su desarrollo. Las guías de práctica clínica deben basarse en la evidencia científica, pero cuando falta evidencia o es contradictoria puede

completarse mediante técnicas de manejo no sesgado del conocimiento de expertos. En el desarrollo de GUIPCAR se han utilizado ambas técnicas.

Síntesis de la evidencia científica

Los propósitos de la síntesis de la evidencia en una guía de práctica clínica son apoyar las recomendaciones en la mejor evidencia posible, obtener la información en la que se apoyan las recomendaciones y permitir la evaluación de la calidad de la guía. Para la elaboración de GUIPCAR, tras una estimación inicial del volumen de publicaciones sobre el tema, se decidió, por una cuestión de eficiencia, centrar la revisión sistemática de la bibliografía en estudios que analizasen la eficacia –ensayos clínicos– de 10 fármacos modificadores de la enfermedad (FAME) en la AR. En la tabla 1 se exponen los FAME sobre los que se centró la revisión sistemática, la estrategia de búsqueda bibliográfica y los criterios de inclusión de los artículos. En la figura 1 se expone un esquema del proceso de selección de artículos. Cada uno de los artículos leídos de manera independiente por 2 revisores (un reumatólogo y un médico con formación epidemiológica). En caso de discrepancia dentro de cada equipo, el artículo fue discutido entre los 2 revisores. Si no se alcanzaba un acuerdo, el registro era leído por otra pareja de revisores. Se presentaron discrepancias entre los revisores en el 4% de los registros y se alcanzó acuerdo en el 100% de los casos. Finalmente, fueron incluidos en la revisión un total de 103 artículos correspondientes a 91 ensayos clínicos diferentes, que permiten realizar 140 comparaciones entre distintas estrategias terapéuticas (fármacos aislados o combinaciones de ellos). En consecuencia, la base de datos de la evidencia de la eficacia de los FAME tiene 140 registros.

La información relevante de cada artículo fue recogida en un formulario diseñado específicamente para esta guía (disponible para el lector que lo solicite en la dirección milena@taiss.com), que permitía identificar la información bibliográfica de los artículos incluidos en la revisión, verificar que el artículo –ya en texto completo– cumplía los criterios de inclusión de la revisión, recopilar la información clínica relevante de cada artículo aceptado

GUIPCAR ha sido un trabajo financiado por Novartis Farmacéutica S.A.

Correspondencia: Dr. C. Hernández-García.
Servicio de Reumatología. Hospital Clínico San Carlos.
28040 Madrid.

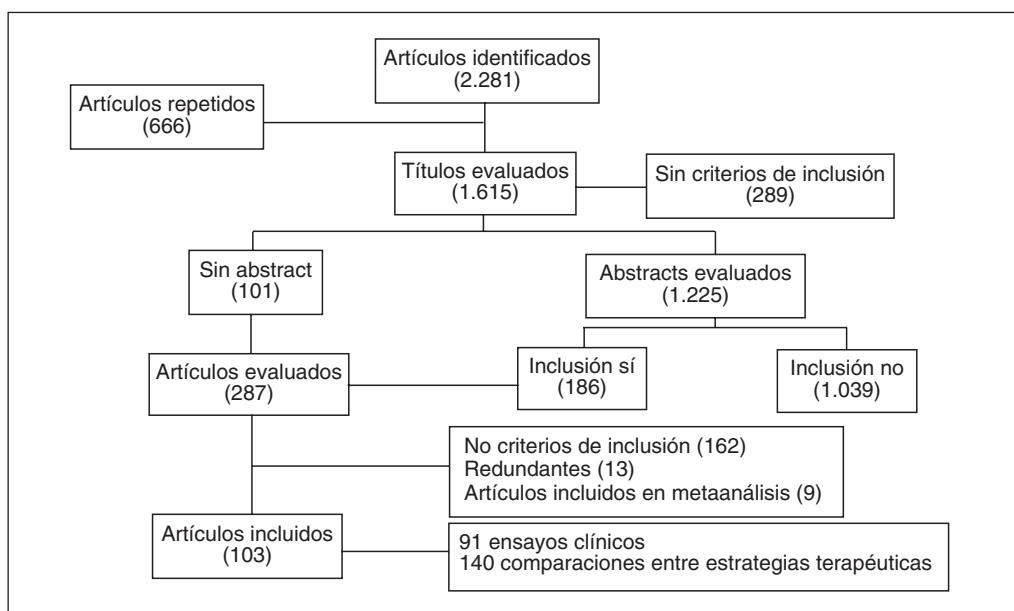


Figura 1. Esquema resumido de la selección de artículos para GUIPCAR.

TABLA 1. Fármacos modificadores de la enfermedad incluidos en la revisión sistemática, estrategia de búsqueda bibliográfica y criterios de inclusión de los artículos

Fármacos modificadores de la enfermedad	Antipalúdicos (cloroquina e hidroxicloroquina) Azatioprina Ciclofosfamida Ciclosporina A D-penicilamina Fármacos anti-TNF (infliximab y etanercept) Leflunomida Metotrexato Sales de oro (vía oral e inyectable) Sulfasalazina MEDLINE, de 1966 a 2000 EMBASE (sección <i>Drugs and Pharmacology</i>) de 1984 a 2000 Índice Médico Español (IME), de 1971 a 2000 Cochrane Library (versión año 2000)
Estrategia de búsqueda bibliográfica	
Criterios de inclusión	<ol style="list-style-type: none"> 1. Ensayo clínico controlado, metaanálisis o revisión sistemática 2. Referirse a uno o más de los fármacos modificadores de la enfermedad seleccionados 3. Comparar la eficacia de un medicamento o terapia combinada frente a otro/a con al menos uno de los fármacos seleccionados, o comparar leflunomida o anti-TNF con placebo 4. Si se trataba de una revisión sistemática o metaanálisis, comparar un medicamento o terapia combinada frente a otro/a con al menos uno de los fármacos seleccionados, o con placebo 5. Haberse realizado en pacientes con AR 6. Haberse realizado en humanos 7. Estar publicado en inglés o castellano

(número y características de los participantes, intervenciones, medidas de resultado evaluadas, retiradas y abandonos, fiabilidad de los instrumentos de medición, seguimiento de los grupos y análisis estadístico), y evaluar la calidad metodológica y el nivel de evidencia de los ensayos. La calidad metodológica de los ensayos fue valorada con la escala de evaluación de la calidad de ensayos clínicos de Jadad¹¹, y el nivel de evidencia fue establecido si-

guiendo los criterios de Hadorn diseñados para evaluar la calidad de la evidencia de las publicaciones usadas para elaborar guías de práctica clínica¹². En la tabla 2 se exponen los ensayos clínicos y las comparaciones según los niveles de calidad de la evidencia de Jadad y de Hadorn. Una primera conclusión es que la evidencia científica de calidad en el tratamiento de la AR es muy escasa, de modo que gran parte de las recomendaciones que se ha-

TABLA 2. Ensayos clínicos y comparaciones según los niveles de calidad de la evidencia de Jadad y los niveles de calidad de la evidencia de Hadorn

Nivel de calidad Escala de Jadad ^a	Ensayos clínicos	Comparaciones
1	7 (8%)	7 (5%)
2	19 (21%)	24 (17%)
3	18 (20%)	31 (22%)
4	32 (35%)	53 (38%)
5	15 (16%)	25 (18%)
Total	91 (100%)	140 (100%)

Nivel de evidencia Escala de Hadorn ^b	Ensayos clínicos	Comparaciones
A1	13 (14%)	25 (18%)
A2	13 (14%)	20 (14%)
B	65 (72%)	95 (68%)
Total	91 (100%)	140 (100%)

^aLa puntuación de un artículo en la escala de Jadad puede adquirir un valor entre 0 y 5. Se considera que un artículo es de buena calidad si la puntuación es de 3 o mayor, y de baja calidad si la puntuación es menor de 3.

^bEl nivel A1 incluye evidencia apoyada por ensayos controlados, aleatorizados, multicéntricos, bien realizados que incluyan 100 o más pacientes; el nivel A2 incluye evidencia respaldada por ensayos aleatorizados controlados bien realizados con menos de 100 pacientes, en una o más instituciones; el nivel B incluye evidencia respaldada por estudios pobremente controlados o no controlados, es decir, evidencia a partir de ensayos clínicos aleatorizados con suficientes defectos metodológicos que puedan invalidar los resultados. Los niveles A1 y A2 se refieren a buena calidad de la evidencia; el nivel B se refiere a calidad de la evidencia con sesgos potenciales que pueden invalidar los resultados.

cen en la guía están apoyadas en un nivel de evidencia científica C, es decir, en la opinión de expertos. Idealmente debería existir evidencia con la suficiente validez, por ejemplo, buenos ensayos clínicos como para apoyar todas las decisiones médicas. El problema es que no existen estos estudios y, por tanto, la evidencia es escasa¹³. Varios trabajos han estimado que sólo entre el 15 y el 20% de las decisiones médicas tomadas en la práctica diaria tienen una justificación en estudios científicos rigurosos^{14,15}. En la AR abunda este tipo de paradigmas; por ejemplo, se han estudiado diferentes estrategias de tratamiento en pacientes en los que fracasó el metotrexato (MTX) en monoterapia, pero sólo existe un ensayo clínico que haya comparado dos de las posibles alternativas a este fármaco en monoterapia entre sí y, desde luego, no existe en la bibliografía ningún estudio que haya comparado 2 alternativas clínicamente relevantes para la práctica diaria⁵. En resumen, no se realizan los ensayos clínicos que podrían ofrecer respuestas relevantes para el manejo de la AR. El enfoque comercial de la investigación clínica y la ausencia de incentivos –o en muchos casos la existencia de desincentivos– para la investigación independiente son algunas de las causas. En la producción de conocimiento sobre la eficacia de las estrategias terapéuticas debería desempeñar un papel relevante el primer financiador en España, que es el Sistema Nacional de Salud, que debería impulsar estudios sobre el uso apropiado o inapropiado de procedimientos. También deberían existir mecanismos para estimular que los investigadores (institucionales o individuales) identifiquen las áreas del conocimiento en las que se carece de estudios relevantes (esta guía puede contribuir notablemente a ello) e investigar

en estas áreas. Por supuesto, las sociedades científicas y la propia industria farmacéutica pueden patrocinar y fomentar estudios independientes de intereses corporativos o comerciales, de los que esta guía puede considerarse como un buen ejemplo. Pero mientras eso ocurre, la vida sigue. Aunque exista poca evidencia científica, los médicos tienen que tomar decisiones. Si no hay evidencia, la única alternativa es que las recomendaciones estén basadas en el manejo no sesgado del juicio colectivo de un panel de expertos que, a su vez, deben haber sido elegidos con criterios como la diversidad de ámbitos, el respeto profesional, o la ausencia de conflictos de interés, entre otros. Se ha demostrado que existe una buena concordancia entre los juicios de un panel de expertos y los resultados de ensayos clínicos posteriores¹⁶. Es decir, los resultados de ensayos clínicos han cambiado en poca o ninguna medida las recomendaciones de un panel de expertos manejado adecuadamente.

Elaboración de las recomendaciones

El panel de expertos estuvo mayoritariamente –pero no sólo– compuesto por reumatólogos propuestos por la SER, siguiendo una metodología diseñada por TAISS. Los criterios que debía reunir el panel para contribuir a la validez de la guía fueron: multidisciplinariedad, diversidad geográfica, diversidad en el nivel asistencial, diversidad en el nivel académico y representación de varones y mujeres en el panel. Individualmente, cada experto debía aportar un *curriculum vitae*, la ausencia explícita de conflictos de interés y la disponibilidad del tiempo necesario para abordar las tareas a desarrollar en el panel.

El panel definió el contenido de la guía y elaboró las recomendaciones y definiciones revisando el texto escrito por los investigadores. Se pidió a los expertos que la redacción de sus recomendaciones proporcionara consejos prácticos y específicos sobre los temas que abordasen, y que sus recomendaciones estuviesen basadas en el balance riesgo/beneficio para el paciente, sin tener en cuenta los costes en sus consideraciones. Es decir, la recomendación debía hacerse siempre considerando qué sería lo más adecuado para el paciente, consistentemente con el objetivo de mejorar la calidad asistencial. Si las recomendaciones no eran coincidentes, cada experto era libre de expresar sus argumentos y la evidencia a favor de su opinión para intentar convencer al resto del panel, pero no se forzó el consenso. Es decir, GUIPCAR no está basada en el consenso, sino en la mejor evidencia científica disponible, y cuando no existe o es contradictoria, se basa en el juicio del panel de expertos sin obligarles a llegar a un consenso. Dicho en otras palabras, como parte de la variabilidad de la práctica clínica puede ser debida al hecho de que la evidencia científica sobre la eficacia de los distintos FAME es insuficiente, esta guía reconoce este hecho y considera admisible que, en estos casos, distintos profesionales puedan tener diferentes opiniones. Cuando la recomendación no está basada en evidencia científica, la recomendación que aparece en GUIPCAR es la votada por la mayoría de los expertos.

Contenido y utilización de GUIPCAR

GUIPCAR abarca aspectos sobre el diagnóstico, la evaluación inicial, la clasificación de la AR, el tratamiento médico, rehabilitador y quirúrgico de la enfermedad, criterios de respuesta al tratamiento, los efectos adversos de los FAME y las complicaciones extraarticulares de la AR. La síntesis de la evidencia sobre la eficacia de los FAME se llevó a cabo mediante una revisión sistemática. En los demás casos (p. ej., eficacia de la cirugía o de la rehabilitación) se han utilizado publicaciones aportadas por los expertos, pero sin realizar una búsqueda sistemática. La guía está dividida en 8 capítulos, que contienen 44 recomendaciones con sus respectivas explicaciones, justificaciones y una extensa revisión bibliográfica. En su edición en papel, cada recomendación se diferencia en su tipografía del resto del texto y tiene una estructura concisa. Se ha intentado que las recomendaciones sean flexibles, de modo que puedan ser utilizadas en función de la experiencia del médico, las características locales, las preferencias personales u otras circunstancias. GUIPCAR incluye, además, una considerable cantidad de información relevante en documentos anexos. En primer lugar, la descripción metodológica de cómo se ha elaborado la guía. En segundo lugar, los instrumentos de recogida de información

para la evaluación inicial y seguimiento de los pacientes, así como las instrucciones para el uso correcto de estos instrumentos de evaluación. En tercer lugar, los algoritmos de decisión desarrollados. Las recomendaciones se presentan en forma de algoritmo anotado, donde cada paso constituye un nivel de decisión diagnóstica o terapéutica, que es seguido por un nuevo nivel dependiendo de los resultados del nivel anterior. Cada nodo del algoritmo lleva un número que se corresponde con una recomendación concreta. Dado que existía una gran variabilidad entre expertos en la clasificación y en la aproximación terapéutica al manejo de la AR, para clasificar a los pacientes según las características de su enfermedad y poder agruparlos con sentido clínico se utilizaron determinadas variables clínicas. Combinando las categorías de estas variables clínicas se configuraron 144 tipos de pacientes distintos en los que el tratamiento podría ser diferente. Los expertos emitieron, de forma anónima e independiente, su recomendación de tratamiento inicial para cada uno de estos 144 pacientes tipo. Sus recomendaciones se analizaron matemáticamente, en especial la proporción de expertos que sugerían cada recomendación terapéutica. Como en muchos de los 144 pacientes diferentes el tratamiento recomendado era similar, fueron agrupados en 44 tipos de pacientes en los que las decisiones terapéuticas eran idénticas (tanto las votaciones anónimas de los expertos como su análisis matemático están a disposición del lector que lo solicite en TAISS: milena@taiss.com). La toma de decisiones para el tratamiento se simplifica aún más en el texto de las recomendaciones, donde se ha separado el uso de corticoides y/o AINE, que sólo se utilizan aisladamente en casos muy concretos, del tratamiento con FAME. De esta manera pueden tenerse sólo en consideración dos parámetros objetivos (articulaciones tumefactas y presencia o no de erosiones) que, tras el análisis, demostraron ser los más relevantes para realizar una clasificación de la enfermedad por su implicación en el tratamiento. En cuarto lugar, se hace explícita la calidad de la evidencia científica que apoya el uso de los FAME y los artículos revisados. Al pie de cada una de las recomendaciones de tratamiento el lector encontrará qué grado de calidad de evidencia científica apoya el uso de cada uno de los FAME y sus combinaciones. Es decir, en qué comparaciones la eficacia del FAME elegido es superior, igual o mejor a otro FAME y qué tipo de calidad de evidencia existe para cada una de ellas. Algunas de las opciones recomendadas no han sido comparadas en ningún ensayo clínico porque no existe, pero otras opciones con suficiente calidad de la evidencia no han sido recomendadas por el panel. Una de las razones es que el panel intentó desarrollar recomendaciones prácticas en el ámbito asistencial español y, en ocasiones, las recomendaciones académicas

ideales pueden no ser aplicables. Por estas razones, una de las características de GUIPCAR es que, con independencia de las recomendaciones, aporta al lector toda la información relevante para que ejerza su libertad en aquellas circunstancias en las que la evidencia científica no existe o es contradictoria sin forzar una u otra elección.

Algunas reflexiones finales

GUIPCAR está dirigida a reumatólogos y, por tanto, excluye aspectos que podrían tener interés para otros especialistas, profesionales o el público en general. La síntesis de la evidencia se centra en la eficacia de los FAME, por lo que una buena parte de las recomendaciones se basan en evidencias científicas sin revisión sistemática u opinión de expertos. Finalmente, el formato en el que está editada no es el ideal y la hace más vulnerable a la obsolescencia que otros formatos. Es decir, GUIPCAR no está exenta de limitaciones pero, pese a ello, creemos que constituye un importante paso adelante en la generación y el ordenamiento del conocimiento en torno al manejo de la AR. El objetivo final de esta guía es ayudar a la toma de decisiones clínicas, al ofrecer recomendaciones para el manejo terapéutico de la AR que se basan en la mejor evidencia disponible. ¿Quiere eso decir que GUIPCAR va a cambiar el manejo de la AR en España? Probablemente no, o no por lo menos en su actual formato. Una vez que una guía de práctica clínica ha sido desarrollada es necesario implantarla y posteriormente evaluarla. Las guías de práctica clínica no se autoimplantan. Requieren liderazgo, diseminación activa, contacto con el usuario y financiar el coste adicional que supone sumarse a las recomendaciones, incentivación y eliminación de trabas. GUIPCAR supone un primer paso y el equipo investigador sigue trabajando para mejorárla, facilitar su difusión (p. ej., en formato electrónico), implantación y evaluación. Si algún usuario de la guía quiere hacer sugerencias para mejorárla, los autores (y sobre todo los usuarios de GUIPCAR) lo agradecerán. Esperamos que nuestro esfuerzo resulte de utilidad para los que tienen la difícil tarea de decidir sobre diferentes cursos de tratamiento y que, al final, redunde en mejores resultados para los pacientes con AR.

ANEXO. Equipo investigador y expertos que han trabajado en la elaboración de la GUIPCAR (por orden alfabético)

TAISS: Miguel Ángel Abad, Santiago Alonso, José Manuel Estrada, Sandra García Armesto, Milena Gobbo, Hildegarda Godoy, Pablo Lázaro y de Mercado, Setefilla Luengo, Yesenia Tordecillas, Yira Tordecillas.

Panel de expertos: José Luis Andreu, Enrique Batlle Gualda, Loreto Carmona Ortells, Federico Díaz-González, Juan José Domínguez Reboiras, Ángel Elena Ibáñez, Juan Gómez-Reino Carnota, Agustín Gómez de la Cámara, Nuria Guñaabens, César Hernández-García, María Victoria Irigoyen, José Luis Marenco, Víctor Martínez-Taboada, Cristina Riera Riazu y José María Salazar Vallinas.

Bibliografía

1. Hernández-García C, González-Álvaro I. El estudio emAR. Variabilidad en el uso de recursos asistenciales. En: Hernández-García C, editor. Estudio sobre el manejo de la artritis reumatoide en España. Madrid: Sociedad Española de Reumatología, 2001; p. 52-89.
2. González-Álvaro I, Hernández-García C. El estudio emAR. Variabilidad en el uso de procedimientos terapéuticos. En: Hernández-García C, editor. Estudio sobre el manejo de la artritis reumatoide en España. Madrid: Sociedad Española de Reumatología, 2001; p. 92-128.
3. Villaverde García V, Hernández-García C. El estudio emAR. Variabilidad en el seguimiento de los pacientes con artritis reumatoide. En: Hernández-García C, editor. Estudio sobre el manejo de la artritis reumatoide en España. Madrid: Sociedad Española de Reumatología, 2001; p. 1 30-9.
4. Field MJ, Lohr KN, editors. Clinical Practice Guidelines: Directions for a New Program. Washington, DC: National Academy Press, 1990.
5. Grupo GUIPCAR. Guía de Práctica Clínica para el Manejo de la Artritis Reumatoide en España. Madrid: Sociedad Española de Reumatología, 2001.
6. American College of Rheumatology. Guidelines for the management of rheumatoid arthritis. *Arthritis Rheum* 1996; 39:713-22.
7. Management of early rheumatoid arthritis. A national clinical guideline. Edinburgh: Scottish Intercollegiate Guidelines Network (SIGN), 2000.
8. Emery P, Reginster JY, Appelboom T, Bredeweld FC, Edelmann E, Kekow S, et al. WHO Collaborating Centre consensus meeting on anti-cytokine therapy in rheumatoid arthritis. *Rheumatology (Oxford)* 2001;40:699-702.
9. Furst DE, Breedveld FC, Burmester GR, Crofford JJ, Emery P, Feldman M, et al. Updated consensus statement on tumour necrosis factor blocking agents for the treatment of rheumatoid arthritis and other rheumatic diseases (April 2001). *Ann Rheum Dis* 2000;59(Suppl 1):1-2.
10. Miembros del Panel (Comité de Expertos de la SER). Consenso de la Sociedad Española de Reumatología sobre la terapia con inhibidores del TNF y otros fármacos inductores de remisión en la artritis reumatoide. *Rev Esp Reumatol* 2000;27:352-4.
11. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Ganavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blindly necessary? *Control Clin Trials* 1996;17:1-12.
12. Hadorn DC, Baker D, Hodges JS, Hicks N. Rating the quality of evidence for clinical practice guidelines. *J Clin Epidemiol* 1996;49:749-54.
13. Fink A, Brook RH, Kosecoff J, Chassin MR, Solomon DH. Sufficiency of clinical literature on the appropriate uses of six medical and surgical procedures. *West J Med* 1987;147: 609-14.
14. Institute of Medicine. Assessing medical technologies. Washington DC: National Academy Press, 1985.
15. Dubinsky M, Ferguson JH. Analysis of the National Institutes of Health Medicare coverage assessment. *Int J Technol Assess Health Care* 1990;6:480-8.
16. Shekelle PG, Chassin MR, Park RE. Assessing the predictive validity of the RAND/UCLA appropriateness method criteria for performing carotidendarterectomy. *Int J Technol Assess Health Care* 1998;14:707-27.