

## Revisión

## Evaluación del uso apropiado de procedimientos sanitarios mediante el método RAND: revisión de su aplicación en la literatura biomédica (1999–2004)

Nerea González<sup>a,d,\*</sup>, José M. Quintana<sup>a,d</sup>, Juan Ramón Lacalle<sup>b,d</sup>, Susana Chic<sup>c</sup> y David Maroto<sup>a,d</sup><sup>a</sup> Unidad de Investigación, Hospital de Galdakao, Vizcaya, España<sup>b</sup> Unidad de Bioestadística, Universidad de Sevilla, España<sup>c</sup> Servicio de Neumología, Hospital de Calahorra, La Rioja, España<sup>d</sup> Ciber de Epidemiología y Salud Pública (CIBERESP), España

## INFORMACIÓN DEL ARTÍCULO

## Historia del artículo:

Recibido el 27 de diciembre de 2006

Aceptado el 5 de junio de 2007

On-line el 9 de marzo de 2009

## Palabras clave:

Método RAND

Revisión

Uso apropiado

## RESUMEN

**Objetivos:** Identificar y describir los estudios de evaluación del uso apropiado de procedimientos sanitarios en los cuales se haya empleado el método RAND/UCLA. Se trata de una técnica de consenso con varias fases para el desarrollo de criterios de uso apropiado.

**Métodos:** La búsqueda bibliográfica se realizó en 2005, consultando las bases de datos OVID-Medline, ISI Web of Knowledge, Índice Médico Español y Highwire. Se seleccionaron los artículos publicados entre 1999 y 2004 que tuviesen las palabras clave «appropriateness», «utilization review» y «physician practice patterns». Se incluyeron los estudios en que se hubiese aplicado el método RAND y se excluyeron aquellos cuya metodología no estuviese suficientemente explicada. De cada artículo se extrajo información sobre el procedimiento estudiado, el lugar y el año de publicación, y las características de la revista.

**Resultados y discusión:** Se identificaron 5.092 referencias y se seleccionaron 205. Algo más de la mitad analizaban procedimientos quirúrgicos o médicos, mientras que el 16,5% se centraba en la evaluación de la calidad asistencial. Más del 50% fueron trabajos publicados en revistas de salud pública, medicina general, y gastroenterología y hepatología. La media del factor de impacto era de 4,07. El 25,4% de los artículos habían sido publicados en 1999.

**Conclusiones y perspectivas:** El método RAND se sigue utilizando ampliamente. Los criterios de uso apropiado pueden emplearse para la revisión de la utilización de procedimientos, como base para elaborar guías o como apoyo para la toma de decisiones. Estas herramientas han de ser sometidas a revisiones para obtener resultados cada vez más válidos y fiables.

© 2006 SESPAS. Publicado por Elsevier España, S.L. Todos los derechos reservados.

## Review of the utilization of the RAND appropriateness method in the biomedical literature (1999–2004)

## A B S T R A C T

## Keywords:

RAND method

Review

Appropriateness

**Objectives:** To identify and describe studies using the RAND/UCLA method to evaluate the appropriateness of health procedures. This method is a consensus technique that involves several phases to develop appropriateness criteria.

**Methods:** We performed a literature search in 2005. OVIDMedline, ISI Web of Knowledge, IME and Highwire were consulted. Articles published between 1999 and 2004 and using the key words «appropriateness», «utilization review» and «physician practice patterns» were selected. Studies using the RAND method were included and those that did not explain the methodology in sufficient detail were excluded. Information on the procedure studied, the place and year of publication, and the characteristics of the journal were extracted from each article.

**Results and discussion:** A total of 5092 articles were identified and 205 were selected. Slightly more than half analyzed surgical or medical procedures, while 16.5% evaluated healthcare quality. More than 50% were published in journals of public health, general medicine, and gastroenterology and hepatology. The mean impact factor was 4.07. A quarter (25.4%) of the articles was published in 1999.

**Conclusions and perspective:** The RAND method is still widely used. Appropriateness criteria can be used to review utilization of procedures, to design guidelines, or to support for decision making. These tools should be reviewed to obtain evermore valid and reliable results.

© 2006 SESPAS. Published by Elsevier España, S.L. All rights reserved.

## Introducción

La rápida incorporación de nuevas tecnologías, el aumento en la demanda de servicios y la ausencia de evidencia científica de

calidad, entre otros factores, han llevado a un aumento de la variabilidad en los criterios para la utilización de determinados procedimientos<sup>1,2</sup>. Esta variabilidad puede provocar una sobreutilización de dichos procedimientos en algunos lugares, así como su infrautilización en otros. Además, el aumento de la frecuencia de determinadas intervenciones y de la incidencia de ciertas dolencias o enfermedades, dan lugar a listas de espera y costes económicos difíciles de gestionar<sup>3,4</sup>.

\* Autor para correspondencia.

Correo electrónico: nerea.gonzalezhernandez@osakidetza.net (N. González).

Todos estos hechos generan dudas sobre la calidad de la asistencia que se dispensa a los pacientes, y suscitan la necesidad de buscar estrategias y métodos para desarrollar criterios consensuados que ayuden en la toma de decisiones sobre la utilización de determinados procedimientos en la práctica clínica<sup>2,5,6</sup>. Entre tales estrategias se encuentran los métodos de consenso, que combinan la opinión de expertos junto con la síntesis de la evidencia<sup>7</sup>. Una de las técnicas más utilizadas es el método RAND/UCLA, desarrollado por la Corporación RAND y la Universidad de California en Los Ángeles, pensado para la evaluación del uso apropiado de las tecnologías médicas<sup>8</sup>. Este método fue desarrollado en los años ochenta y se ha empleado en diversos países para la evaluación de distintos tipos de procedimientos<sup>9-15</sup>.

El método RAND tiene varias fases: revisión crítica de la bibliografía sobre el tema, creación de escenarios comprensibles y mutuamente excluyentes basados en las principales variables que participan en la toma de decisiones, selección de los expertos nacionales (de diversas especialidades si es factible), que formarán parte del panel, y puntuación de los escenarios en dos rondas, una de forma individual y otra presencial, siguiendo un proceso Delphi modificado.

Tanto el RAND como otros métodos de evaluación del uso apropiado pretenden, entre otras cosas, aportar herramientas aplicables en la práctica asistencial. En este contexto surgió hace 5 años la RedIRYSS (Red de Investigación en Resultados y Servicios Sanitarios), compuesta por 9 líneas de investigación, de las cuales una tiene como objetivo principal analizar el uso apropiado de las tecnologías médicas.

Una de las primeras actividades planteadas en esta línea de investigación consistió en realizar una revisión de la bibliografía para conocer otras experiencias de evaluación del uso apropiado con el método RAND, con el fin de obtener información de los tipos de procedimientos evaluados, los resultados obtenidos y las aplicaciones metodológicas. Anteriormente, investigadores de RAND Europe y otros grupos asociados ya habían realizado una revisión de lo publicado hasta junio de 1999<sup>16</sup>, y en esta ocasión se ha completado dicha revisión desde esa fecha hasta diciembre de 2004. Se pretendía identificar los procedimientos estudiados en este período, la metodología empleada y las mejoras introducidas en ella, así como los criterios desarrollados. También se evaluó la aplicabilidad de esta técnica y su relevancia actual para la evaluación de la calidad asistencial.

En el presente artículo se describen los métodos y los resultados de la revisión, que tenía como objetivo identificar y describir las características de las experiencias de evaluación del uso apropiado de la indicación de procedimientos, en las cuales se hubiese empleado el método RAND/UCLA.

## Métodos

Se realizó una revisión de la bibliografía sobre la utilización del método RAND, que incluyó una búsqueda bibliográfica sistemática, una selección de los estudios basada en criterios predefinidos y una descripción de éstos basada en una serie de variables asociadas con el tipo de estudio, el tipo de tecnología y algunos aspectos relacionados con la revista en que se publicaron.

La búsqueda bibliográfica se realizó entre febrero y abril de 2005, consultando las bases de datos OVID-Medline, ISI Web of Knowledge, IME (Índice Médico Español) y Highwire de la Universidad de California. En dicha búsqueda se incluyó cualquier artículo que hubiese sido publicado entre junio de 1999 y diciembre de 2004.

Las palabras clave empleadas fueron «*appropriateness*», «*utilization review*» y «*physician practice patterns*», como texto libre en

el título o el resumen, o bien como palabras clave (MeSH Major Topic). Para búsquedas más específicas se acotó con los términos «*RAND Appropriateness Method*» y «*Modified Delphi Technique*». Los resultados de la búsqueda se completaron con los proporcionados por el profesor Hermann Stoevelaar, provenientes de bases de datos que se habían ido completando desde que el Grupo RAND Europe iniciara la revisión.

Se incluyeron en la revisión todos los estudios que hubiesen aplicado el método RAND, que versasen sobre la evaluación del uso apropiado de una tecnología, el desarrollo o la validación de un instrumento de medida (método RAND) de la adecuación de una tecnología, o el desarrollo metodológico del método RAND y su aplicación práctica. Se empleó un concepto amplio del término «evaluación» de tecnología, que incluía la valoración de fármacos, procedimientos quirúrgicos, pruebas diagnósticas, dispositivos asistenciales, atención médica en general, servicios sanitarios, etc.

Se excluyeron los estudios en que se aplicasen métodos diferentes del RAND/UCLA, así como aquellos cuya metodología no estuviese suficientemente explicada y pudiera identificarse de forma inequívoca como la que se emplea en el método RAND. Los elementos necesarios para identificar la metodología RAND de uso apropiado incluían la realización de una revisión bibliográfica, la configuración de un panel de expertos y la puntuación de unos escenarios en una o más rondas, empleando un algoritmo similar al del RAND en cuanto a la definición de acuerdo y uso apropiado<sup>8</sup>.

Para poder decidir qué artículos se incluirían en el informe, sus resúmenes fueron revisados por cuatro personas, por pares, y las discordancias se evaluaron de nuevo conjuntamente y se consensaron. Cuando el resumen no tenía información suficiente para valorar los criterios de inclusión y exclusión, se accedió al texto completo.

Los artículos seleccionados se clasificaron en 4 grandes clases: a) artículos teóricos y metodológicos; b) estudios en que se hubiese aplicado la metodología RAND para la evaluación de tecnologías o procedimientos concretos; c) comparaciones de estudios que emplearan el método RAND para la evaluación de un mismo procedimiento, y d) estudios sobre nuevas aplicaciones del método RAND más allá de la evaluación de la indicación de uso apropiado, como son la elaboración de indicadores de calidad o la evaluación de los procesos de cuidados. También se recogió información sobre el tipo de procedimiento estudiado, el país de origen de los autores, el año de publicación, el tipo de revista, su factor de impacto y el número de veces que el artículo había sido citado. El tipo de procedimiento fue clasificado en médico, quirúrgico, diagnóstico, elaboración de indicadores de calidad, evaluación de proceso de cuidados, y otros (en esta última categoría se incluían todos los procedimientos que no podían clasificarse en las anteriores). La variable «tipo de revista» fue categorizada teniendo en cuenta el criterio empleado en la base de datos ISI Web of Knowledge, que clasifica las revistas en distintos grupos en función del contenido, y de esa misma base de datos se extrajo el factor de impacto de las revistas.

## Análisis

Los artículos identificados se introdujeron en una base de datos del programa Reference Manager v.11<sup>17</sup>. Con el fin de obtener datos descriptivos de los artículos incluidos en la revisión, se creó una base de datos en el programa SPSS<sup>18</sup>.

Como indicadores descriptivos de los datos se calcularon las frecuencias y los porcentajes para las variables cualitativas, y las medias y la desviación estándar (DE) para las cuantitativas. Además, en algunos casos se utilizó la prueba de la  $\chi^2$  para analizar la homogeneidad de la distribución entre las distintas categorías.

Por último, para determinar la relación entre algunas de las variables cualitativas se utilizó la prueba  $\chi^2$ . En todos los casos se consideró significación estadística un valor de  $p < 0,05$ .

## Resultados y discusión

La revisión bibliográfica aquí presentada se ha centrado en los estudios publicados desde 1999, dado que la Corporación RAND ya había realizado su propia revisión desde que se creó el método, en los años ochenta, hasta mediados de 1999<sup>16</sup>. Se identificaron 5.092 referencias, de las cuales se seleccionaron 205 (4%) que cumplieran los criterios de inclusión. En el período anterior, hasta 1999, se habían incluido 211 referencias. En ambos períodos se encontró un número importante de estudios que reflexionan sobre el método o bien lo emplean para desarrollar criterios aplicables en la práctica clínica. Se puede observar que desde que se creó el método hasta el final de la primera revisión se recopiló un número similar de estudios a los incluidos en la siguiente revisión, lo que indica que se trata de un método que se sigue empleando. En la primera revisión, muchos de los artículos se centraron en el estudio del uso apropiado de procedimientos cardiovasculares, mientras que en la que aquí se presenta parece haber una mayor dispersión en el tipo de procedimiento. El listado completo de las referencias incluidas en la presente revisión, así como un informe más extenso, pueden solicitarse a los autores del artículo.

Los artículos se clasificaron según su contenido. Así, 47 de los estudios trataban aspectos teóricos y metodológicos del RAND, 100 aplicaban criterios de uso adecuado en la evaluación de tecnologías o procedimientos, 16 llevaban a cabo una comparación de los resultados obtenidos por distintos paneles en un mismo procedimiento, y los 42 restantes fueron clasificados como «nuevas aplicaciones del método», como la elaboración de indicadores de calidad<sup>19,20</sup> y la evaluación de los procesos de cuidados<sup>21,22</sup>.

Los procedimientos medicoquirúrgicos constituyeron el objeto de análisis de más de la mitad de los artículos (56,1%), mientras que en el otro extremo nos encontramos con un 2,4% de estudios centrados en la evaluación de los procesos de cuidados. Además, el 10,7% de las referencias no pudieron clasificarse dentro de los grupos establecidos, por tratarse de discusiones metodológicas sobre el RAND o por no centrarse en un único procedimiento (tablas 1 y 2). Se encontraron diferencias estadísticamente significativas entre el tipo de procedimiento y el tipo de estudio, que se explicaban porque el 81% de los estudios clasificados como «nuevas aplicaciones» analizaban procedimientos médicos y de creación de indicadores de calidad, y porque más del 60% de los que comparaban criterios desarrollados por distintos paneles estaban centrados en procedimientos quirúrgicos (tabla 3).

También se analizaron las diferencias entre el tipo de procedimiento y la clase de revista en que había sido publicado el estudio. Se encontraron diferencias determinadas por el hecho de que la mayoría de los estudios centrados en procedimientos quirúrgicos se publicaron en revistas de salud pública y servicios de salud, y los centrados en diagnósticos aparecían en revistas

**Tabla 2**

Procedimientos evaluados en cada categoría de la variable «tipo de procedimiento»

Categoría	Tipo de procedimiento
Quirúrgico	Endarterectomía carotídea. Revascularización coronaria. Histerectomía. Cirugía de catarata. Prótesis de cadera. Cateterismo. Angiografía coronaria. Angioplastia coronaria. Bypass coronario. Angiograma coronario. Laminectomía. Colectomía. Trasplante renal. Cirugía para ciática. Prótesis de rodilla
Diagnóstico	Colonoscopia. Endoscopia. Evaluación diagnóstica de las enfermedades de la arteria coronaria. Evaluación diagnóstica de síntomas de obstrucción benigna de próstata. Gastroscopia. Evaluación diagnóstica de síndrome de colon irritable
Médico	Tratamiento de las urgencias. Depresión. Metadona. Tratamientos gastroprotectores. Angina de pecho. Asma. Diabetes mellitus tipo 2. Trastorno bipolar. Alcohol en ancianos. Derivación de pacientes. Hiperplasia benigna de próstata. Síndromes mieloplásicos. Trastornos psicóticos. Prescripción de fármacos en ancianos. Prescripción de antiinflamatorios. Antipsicóticos. Hemofilia. Trastornos cardíacos. Eritropoyetina recombinante. Neumotórax. Manipulación espinal. Síndrome de colon irritable. Síndrome de fatiga crónica. Dolor de espalda crónico. Latigazos. Infarto. Esquizofrenia. Leucemia. Infección respiratoria. Quimioterapia y autotrasplante en cáncer de mama. Mieloma. Artritis reumatoide
Calidad	Indicadores de calidad en cáncer, asistencia sanitaria en urgencias, cuidado de ancianos, atención primaria, pacientes oncológicos, pacientes infectados por el virus de la inmunodeficiencia humana, trastornos cardiopulmonares, niños y adolescentes, y mujeres. Criterios de calidad de manejo de la angina y el fallo cardíaco. Desarrollo de herramientas de evaluación de la calidad asistencial. Intervenciones en cardiología y ginecología
Proceso de cuidados	Biopsia de nodo linfático e interferón alfa-2b. Medida de la infratilización de los cuidados necesarios en ancianos. Definición y medida de la calidad asistencial. Síntomas del tracto urinario. Cuidado de pacientes no agudos en urgencias
Otros	Actividades de laboratorio. Perfeccionamiento de los procesos de paneles de expertos. Directrices farmacéuticas. Influencia de la medicina basada en la evidencia en la práctica. Disfunción sexual femenina. Metodología de uso apropiado. Técnicas de consenso. Desarrollo y validación de sistemas informáticos. Uso del método RAND para detectar la infratilización. Estándares para la práctica asistencial basados en la evidencia. Definición de calidad asistencial. Criterios de uso apropiado para la mejora de la salud. Quiropraxis. Desarrollo y validación de guías de práctica clínica

relacionadas con la gastroenterología, pues los procedimientos diagnósticos más estudiados con el método RAND durante este período fueron la gastroscopia y la colonoscopia (tabla 4). El hecho de que la mayoría de los estudios sobre procedimientos quirúrgicos se hubieran publicado en revistas de salud pública y servicios de salud podría estar motivado por la relación entre el desarrollo de criterios de uso apropiado y la evaluación de la calidad asistencial, así como por la menor disposición de revistas de especialidades médicas para publicar artículos sobre el desarrollo de criterios explícitos, como los del RAND.

Estados Unidos, país donde se desarrolló el método, ya presentaba un gran volumen de estudios en los que se empleaba el RAND<sup>23-25</sup>, y continúa siendo el lugar que más producción científica tiene, con 92 referencias identificadas (tabla 5). Sin embargo, Europa ha ido adquiriendo importancia en el empleo de este método. Así, el Reino Unido<sup>26,27</sup>, Suiza<sup>28,29</sup>, España<sup>30,31</sup>, los Países Bajos<sup>32,33</sup> e Italia<sup>34,35</sup> engloban casi el 50% de las referencias incluidas en la presente revisión, lo que indica que la línea de investigación sobre el uso apropiado de este método se va consolidando en Europa. Por tanto, en el período comprendido entre los años 1999 y 2004 no hay diferencias entre Estados

**Tabla 1**  
Procedimientos estudiados en los artículos

Tipo de procedimiento	n	%
Quirúrgicos	71	34,6
Médicos	44	21,5
Diagnósticos	34	16,6
Elaboración de indicadores de calidad	29	14,1
Evaluación de procesos de cuidados	5	2,4
Otros	22	10,7

**Tabla 3**  
Resultados del análisis del tipo de procedimiento según el tipo de estudio

Tipo de estudio	Tipo de procedimiento						Total
	Quirúrgico	Diagnóstico	Médico	Calidad	Proceso de cuidados	Otros	
Metodológico			9 (20,5%)				
Aplicar criterios	42 (59,2)	31 (91,2)	22 (50)	1 (3,4)	3 (60)	1 (4,5)	100 (48,8)
Comparar paneles	10 (14,1)	3 (8,8)	2 (4,5)	1 (3,4)	0 (0)	0 (0)	16 (7,8)
Nuevas aplicaciones	1 (1,4)	0 (0)	11 (25)	23 (79,3)	1 (20)	6 (27,3)	42 (20,5)
Total	71 (34,6)	34 (16,6)	44 (21,5)	29 (14,1)	5 (2,4)	22 (10,7)	205 (100)

**Tabla 4**  
Diferencias en el tipo de procedimiento según el tipo de revista

Tipo de procedimiento <sup>a</sup>	Tipo de revista								Total
	Salud pública y servicios de salud	Medicina general	Psiquiatría y neurología	Otras especialidades médicas	Cirugía y otras especialidades quirúrgicas	Gastroenterología y hepatología	Sistemas cardíaco y cardiovascular	Otros	
Quirúrgico	30 (53,6)	14 (37,8)	7 (53,8)	2 (9,1)	9 (81,8)	0 (0)	8 (66,7)	1 (3,4)	71 (34,6)
Diagnóstico	3 (5,4)	3 (8,1)	1 (7,7)	1 (4,5)	0 (0)	25 (100)	1 (8,3)	0 (0)	34 (16,6)
Médico	8 (14,3)	10 (27)	4 (30,8)	7 (31,8)	1 (9,1)	0 (0)	3 (25)	11 (37,9)	44 (21,5)
Calidad	8 (14,3)	8 (21,6)	1 (7,7)	7 (31,8)	0 (0)	0 (0)	0 (0)	10 (34,5)	34 (16,6)
Otros	7 (12,5)	2 (5,4)	0 (0)	5 (22,7)	1 (9,1)	0 (0)	0 (0)	7 (24,1)	22 (10,7)
Total	56 (27,3)	37 (18)	13 (6,3)	22 (10,7)	11 (5,4)	25 (12,2)	12 (5,9)	29 (14,1)	205 (100)

<sup>a</sup> Se han agrupado las categorías «calidad» y «proceso de cuidados», dado que tratan temas similares y tienen pocas referencias.

**Tabla 5**  
Datos generales de los artículos

Variables	n	%
País de origen de los autores		
Estados Unidos	92	44,9
Suiza	36	17,6
Reino Unido	24	11,7
España	18	8,8
Países Bajos	11	5,4
Canadá	10	4,9
Italia	8	3,9
Otros	6	2,9
Año de publicación		
1999	52	25,4
2000	37	18
2001	36	17,6
2002	24	11,7
2003	28	13,7
2004	28	13,7

Unidos y Europa en cuanto al volumen de estudios publicados en los cuales se haya empleado el método RAND.

En el caso concreto de España, 18 (8,8%) de los artículos incluidos en la revisión habían sido redactados por autores españoles, y la mayoría (77,8%) trataba sobre el uso apropiado de procedimientos quirúrgicos. El 78% de los estudios de autores españoles se publicó entre los años 2000 y 2002. Más del 60% apareció en revistas del ámbito de la salud pública y los servicios de salud, y el 77% se publicó en revistas internacionales. Por último, el factor de impacto de las revistas en que aparecieron estos artículos oscilaba entre 0 y 4,102, con una media de 1,322 y una mediana de 1,05.

Respecto al número de artículos por año de publicación, 52 de las referencias aparecieron en el año 1999, mientras que las demás se distribuyeron de forma homogénea en el resto de los años (tabla 5). No se hallaron diferencias estadísticamente significativas en cuanto al tipo de procedimiento estudiado en función del

**Tabla 6**  
Datos de las revistas en que se han publicado los artículos

Variables	n	%
Tipo de revista		
Salud pública y servicios de salud	56	27,3
Medicina general	37	18,0
Gastroenterología y hepatología	25	12,2
Sistemas cardíaco y cardiovascular	12	5,9
Psiquiatría y neurología	13	6,3
Otras especialidades médicas	22	10,7
Cirugía y otras especialidades quirúrgicas	11	5,4
Otros	29	14,1
Factor de impacto		
0	45	22
≤2	48	23,4
≤4	49	23,9
≤7	38	18,5
>7	25	12,2

año de publicación. Estos datos reflejan la actualidad del método, puesto que se sigue empleando en la evaluación de distintos tipos de procedimientos.

Más de la mitad de los artículos se publicaron en revistas del ámbito de la salud pública y los servicios de salud (27,3%), medicina general (18%), y gastroenterología y hepatología (12,2%). El resto de las referencias aparecieron en revistas de otro tipo de especialidades médicas, de cirugía o en formatos distintos al artículo científico, como informes o monografías<sup>36-38</sup>. Las revistas en que más artículos se publicaron fueron *Endoscopy*, *Medical Care* y *Quality & Safety in Health Care*. Su factor de impacto oscilaba entre 0 y 38,75, con una media de 4,07 y una mediana de 2,21. Una vez categorizada esta variable, en un extremo se encontraron 45 artículos de revistas que no tenían factor de impacto, mientras que en el extremo opuesto se identificaron 76 (12,2%) cuyo factor de impacto se situó por encima de 7 (tabla 6).

Se revisaron también las veces que cada artículo había sido citado por otros autores. El número de citas oscilaba entre 0 y 377, con una media (DE) de 14,25 (34,71).

La revisión bibliográfica aquí descrita presenta algunas limitaciones que hay que tener en cuenta al interpretar sus resultados. En primer lugar, es posible que haya un sesgo en la identificación de las referencias, pues aunque se han consultado las bases de datos más empleadas en medicina, algunos estudios no se publican en revistas indexadas en estas bases. Sin embargo, se ha tratado de completar la búsqueda comparando los resultados de nuestra revisión con la realizada en los últimos años por el grupo RAND Europe.

Algunos autores estiman que la carga de trabajo que supone para los panelistas es excesiva, no está clara la unión entre la evidencia científica y las puntuaciones de los panelistas, la composición de los paneles puede condicionar los resultados de las puntuaciones, y se trata de un método que presenta dificultades cuando se consideran múltiples tratamientos posibles para una misma enfermedad<sup>1</sup>. Kahan y Van het Loo<sup>1</sup> llevaron a cabo una acción concertada a escala europea para diseminar el método, con la cual trataron de minimizar estas dificultades.

Además, hay dudas acerca de la validez y la fiabilidad del método<sup>39-42</sup>. Algunos estudios han analizado estos problemas<sup>43,44</sup>, tratando de buscar soluciones. Así, Shekelle et al<sup>45</sup> afirman que, si bien es cierto que el método podría mejorarse en varios aspectos, el resto de las técnicas de consenso no son mejores<sup>46</sup>. Los mismos investigadores que desarrollaron el método han publicado una serie de recomendaciones para mejorarlo y lograr que sea más aplicable en la práctica clínica y la gestión de los servicios sanitarios<sup>47</sup>.

## Conclusiones y perspectivas

El método RAND, como técnica que combina la evidencia científica disponible y la opinión de expertos, se sigue utilizando ampliamente, a juzgar por el número de estudios publicados en los últimos 5 años. Tradicionalmente se ha empleado en la evaluación del uso apropiado de procedimientos médicos<sup>10,15,48</sup>, quirúrgicos<sup>6,30,34,49,50</sup> y diagnósticos<sup>14,51-53</sup>. Sin embargo, en los últimos años se han realizado estudios que utilizan el RAND con otros objetivos, como la elaboración de indicadores de calidad<sup>38,39</sup> o la evaluación de los procesos de cuidados<sup>40,41</sup>.

El desarrollo de criterios explícitos para la toma de decisiones aplicadas a los procedimientos medicoquirúrgicos es el área más estudiada con este método, aunque las nuevas aplicaciones mencionadas van cobrando mayor importancia en los estudios incluidos en la revisión. En cuanto a los procedimientos diagnósticos, los artículos identificados tratan los que tienen relación con la gastroenterología, analizando el uso apropiado de ciertas técnicas, como la endoscopia y la gastroscopia, pero se han encontrado estudios sobre otros tipos de procedimientos diagnósticos, lo que señala una probable área de desarrollo para los estudios de uso adecuado. La mayor frecuencia de artículos sobre procedimientos médicos, quirúrgicos y diagnósticos está relacionada con el tipo de revistas en que se han publicado, sobre todo en las de medicina general, gastroenterología y hepatología, y de salud pública y servicios de salud.

Las perspectivas de uso del método RAND vendrán condicionadas en parte por la confianza en él de los profesionales, y por las mejoras que se vayan haciendo para superar sus limitaciones. Los criterios explícitos de uso apropiado creados con el método RAND tienen en su aplicabilidad uno de sus grandes potenciales, dado que pueden emplearse para las revisiones sobre la utilización de un determinado procedimiento, como base para la elaboración de guías de práctica clínica y apoyo a la toma de decisiones por parte de los clínicos, los gestores y los proveedores sanitarios. En cualquier caso, se trata de herramientas que han de someterse a continuas revisiones con el fin de mejorarlas, e incluso se deben

investigar nuevas aplicaciones, de manera que sea posible obtener resultados cada vez más válidos y fiables.

## Agradecimientos

Los autores desean agradecer a Hermann Stoevelaar, del Grupo RAND Europe, su colaboración en la presente revisión al aportar referencias a las cuales nos habría sido imposible acceder. Asimismo, es de agradecer el trabajo realizado por las personas participantes en la Línea 2.2 de Uso Apropriado de la Red IRYSS, tanto en las revisiones bibliográficas como en el resto de las actividades llevadas a cabo, que además de los autores del artículo han sido las siguientes: Eduardo Aguayo, Elena Andradás, Marisa Baré, José M. Beguiristain, Juan Antonio Blasco, Belén Elizalde, Antonio Escobar, Susana García, Gemma Navarro, Emilio Perea y Jesús Martínez Tapias.

## Financiación

La revisión bibliográfica cuyos resultados se presentan en este artículo se realizó dentro de un estudio más amplio, que fue financiado por la Subdirección General de Redes del Fondo de Investigación Sanitaria (G03/202).

## Bibliografía

- Kahan JP, Van het Loo M. Defining appropriate health care. *Eurohealth*. 1999;5:16-8.
- Martínez-Sahuquillo ME, Echevarría Ruiz de Vargas MC. Métodos de consenso. Uso adecuado de la evidencia en la toma de decisiones. *Método RAND/UCLA. Rehabilitación*. 2001;35:388-92.
- Wennberg J. The paradox of appropriate care. *JAMA*. 1987;258:2568-9.
- Chassin MR, Koseoff J, Park RE, et al. Does inappropriate use explain geographic variation in the use of health care services? *JAMA*. 1987;258:2533-7.
- Feinstein AR, Horwitz RJ. Problems in the «evidence» of «evidence-based medicine». *Am J Med*. 1997;103:529-35.
- Quintana JM, Arostegui I, Azkarate J, et al. Evaluation of explicit criteria for total hip joint replacement. *J Clin Epidemiol*. 2000;53:1200-8.
- Campbell SM, Cantrill JA. Consensus methods in prescribing research. *J Clin Pharm Ther*. 2001;26:5-14.
- Brook RH, Chassin MR, Fink A, et al. A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care*. 1986;2:53-63.
- Aguilar MD, Fitch K, Lazaro P, et al. The appropriateness of use of percutaneous transluminal coronary angioplasty in Spain. *Int J Cardiol*. 2001;78:213-21.
- Alexopoulos GS, Streim J, Carpenter D, et al. Expert Consensus Panel for Using Antipsychotic Drugs in Older Patients. Using antipsychotic agents in older patients. *J Clin Psychiatry*. 2004;65(Suppl 2):5-99.
- Escobar A, Quintana JM, Arostegui I, et al. Development of explicit criteria for total knee replacement. *Int J Technol Assess Health Care*. 2003;19:57-70.
- Oishi SM, Morton SC, Moore AA, et al. Using data to enhance the expert panel process. Rating indications of alcohol-related problems in older adults. *Int J Technol Assess Health Care*. 2001;17:125-36.
- Stoevelaar HJ, McDonnell J, Van de Beek C, et al. Appropriate treatment of benign prostatic hyperplasia; refining the indications for treatment by systematic analysis of expert opinion [Passende behandeling bij benigne prostaathyperplasie; verfijning van de indicatiestelling door systematische analyse van klinische expertise]. [Dutch] *Ned Tijdschr Geneesk*. 1999;143:2425-9.
- Vader JP, Burnand B, Froehlich F, et al. The European Panel on Appropriateness of Gastrointestinal Endoscopy (EPAGE): project and methods. *Endoscopy*. 1999;31:572-8.
- Vader JP, Hammig R, Besson J, et al. Appropriateness of methadone maintenance treatment for opiate addiction: evaluation by an expert panel. *Sozial und Präventivmedizin*. 2003;48:1S-14S.
- Van het Loo M, Kahan JP. The RAND appropriateness method: an annotated bibliography through 1999. RAND Europe. Santa Monica: RAND Distribution Services; 1999. Disponible en: [http://www.rand.org/pubs/rand\\_europe/RE99-010](http://www.rand.org/pubs/rand_europe/RE99-010).
- Reference Manager v.11 for Windows. Thomson ISI Research Soft, 2004.
- SPSS versión 12.0 para Windows. Chicago: SPSS Inc.; 2003.
- Shekelle PG, MacLean CH, Morton SC, et al. Assessing care of vulnerable elders: methods for developing quality indicators. *Ann Intern Med*. 2001;135:647-52.
- Muijrs PE, Janknegt R, Sijbrandij J, et al. Prescribing indicators. Development and validation of guidelinebased prescribing indicators as an instrument to

- measure the variation in the Oprescribing behaviour of general practitioners. *Eur J Clin Pharmacol.* 2004;60:739–46.
21. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. *N Engl J Med.* 2003;348:2635–45.
  22. Campbell SM, Roland MO, Shekelle PG, et al. Development of review criteria for assessing the quality of management of stable angina, adult asthma, and noninsulin dependent diabetes mellitus in general practice. *Qual Health Care.* 1999;8:6–15.
  23. Tobacman JK, Zimmerman B, Lee P, et al. Visual acuity following cataract surgeries in relation to preoperative appropriateness ratings. *Med Decis Making.* 1999;23:122–30.
  24. Halm EA, Chassin MR, Tuhim S, et al. Revisiting the appropriateness of carotid endarterectomy. *Stroke.* 2003;34:1464–71.
  25. Garg PP, Landrum MB, Normand SL, et al. Understanding individual and small area variation in the underuse of coronary angiography following acute myocardial infarction. *Med Care.* 2002;40:614–26.
  26. Hemingway H, Crook AM, Dawson JR, et al. Rating the appropriateness of coronary angiography, coronary angioplasty and coronary artery bypass grafting: the ACRE study. Appropriateness of Coronary Revascularisation Study. *J Public Health Med.* 1999;21:421–9.
  27. Feder G, Crook AM, Magee P, et al. Ethnic differences in invasive management of coronary disease: prospective cohort study of patients undergoing angiography. *BMJ.* 2002;324:511–6.
  28. Froehlich F, Gonvers JJ, Vader JP, et al. Appropriateness of gastrointestinal endoscopy: risk of complications. *Endoscopy.* 1999;31:684–6.
  29. SeematterBagnoud L, Vader JP, Wietlisbach V, et al. Overuse and underuse of diagnostic upper gastrointestinal endoscopy in various clinical settings. *Int J Qual Health Care.* 1999;11:301–8.
  30. Quintana JM, Cabriada J, De Tejada IL, et al. Development of explicit criteria for cholecystectomy. *Qual Saf Health Care.* 2002;11:320–6.
  31. Aguilar MD, Lazaro P, Fitch K, et al. Gender differences in clinical status at time of coronary revascularisation in Spain. *J Epidemiol Commun Health.* 2002;56:555–9.
  32. Stoevelaar HJ, McDonnell J, Stals H, et al. Gastroprotective treatment in patients using NSAIDs. Development of appropriateness criteria by a multidisciplinary expert panel. *Scand J Rheumatol.* 2003;32:162–7.
  33. McDonnell J, Stoevelaar HJ, Bosch JL, et al. The appropriateness of treatment of benign prostatic hyperplasia: a comparison of Dutch and multinational criteria. *Health Policy.* 2001;57:45–56.
  34. Fantini MP, Negro A, Accorsi S, et al. Development and assessment of a priority score for cataract surgery. *Can J Ophthalmol.* 2004;39:48–55.
  35. Filardo G, Maggioni AP, Mura G, et al. The consequences of underuse of coronary revascularization; results of a cohort study in Northern Italy. *Eur Heart J.* 2001;22:654–62.
  36. Goodman JC. «Necessary» and «unnecessary» medical care. En: Goodman JC, editor. *The American experience with managed care.* Dallas: National Center for Policy Analysis; 2002. p. 11–7.
  37. Frinking E, Kahan JP, Van het Loo M, et al. Risk profiles and appropriate treatment therapies for whiplash associated disorders. Synthesis report conducted for the Swiss Insurance Association (SIA). Lausanne: SIA; 2003.
  38. Kerr EA, Asch SM, Hamilton EG, et al. Quality of care for general medical conditions: a review of the literature and quality indicators. Santa Monica: RAND Health; 2000. MR1280AHRQ. Disponible en: <http://www.rand.org/cgi-bin/Abstracts/abdb.pl>.
  39. Phelps CE. The methodological foundations of studies of the appropriateness of medical care. *N Engl J Med.* 1993;329:1241–5.
  40. Kravitz RL, Laouri M, Kahan JP, et al. Validity of criteria used for detecting underuse of coronary revascularization. *JAMA.* 1995;274:632–8.
  41. McKee M, Priest P, Ginzler M, et al. How representative are members of expert panels. *Qual Assur Health Care.* 1991;3:89–94.
  42. Scott EA, Black N. When does consensus exist in expert panels. *J Public Health Med.* 1991;13:35–9.
  43. Shekelle P, Kahan JP, Park RE, et al. Assessing appropriateness by expert panels: how reliable? *J Gen Intern Med.* 1996;10:81.
  44. Naylor CD. What is appropriate care? *N Engl J Med.* 1998;338:1918–20.
  45. Shekelle P, Kahan JP, Bernstein SJ, et al. The reproducibility of a method to identify the overuse and underuse of procedures. *N Engl J Med.* 1998;338:1888–95.
  46. Shekelle PG. The appropriateness method. *Med Decis Making.* 2004;24:228–31.
  47. Fitch K, Berstein SJ, Aguilar MD, et al. The RAND/UCLA appropriateness method user's manual. MR1269DGXII/RE. Santa Monica: RAND; 2001.
  48. Altshuler LL, Cohen LS, Moline ML, et al. The Expert Consensus Guideline Series. Treatment of depression in women. *Postgrad Med.* 2001. p. 1–107.
  49. Hannan EL, Van Ryn M, Burke J, et al. Access to coronary artery bypass surgery by race/ethnicity and gender among patients who are appropriate for surgery. *Med Care.* 1999;37:68–77.
  50. Porchet F, Vader JP, Larequi-Lauber T, et al. The assessment of appropriate indications for laminectomy. *J Bone Joint Surg Br.* 1999;81:234–9.
  51. Kohut M, Romanczyk T, Nowak A, et al. Estimation of appropriateness of colonoscopy using RAND/UCLA method [Ocena zasadności przeprowadzania kolonoskopii z wykorzystaniem metody RAND/UCLA]. [Polish]. *Wiad Lek.* 2004;57:321–6.
  52. NicollierFahrni A, Vader JP, Froehlich F, et al. Development of appropriateness criteria for colonoscopy: comparison between a standardized expert panel and an evidencebased medicine approach. *Int J Qual Health Care.* 2003;15:15–22.
  53. Bochud M, Gonvers JJ, Vader JP, et al. Appropriateness of gastroscopy: gastroesophageal reflux disease. *Endoscopy.* 1999;31:596–603.