

COLABORACIÓN ESPECIAL**VALORACIÓN DE ESCALAS Y CRITERIOS PARA LA EVALUACIÓN DE GUÍAS DE PRÁCTICA CLÍNICA (*)**

Rosa Rico Iturrioz (1), Iñaki Gutiérrez-Ibarluzea (1), José Asua Batarrita (1), M^a Asunción Navarro Puerto (2), Antonio Reyes Domínguez (2), Ignacio Marín León (2), Eduardo Briones Pérez de la Blanca (3)

- (1) Osteba. Servicio de Evaluación de Tecnologías Sanitarias. Gobierno Vasco.
(2) Servicio de Medicina Interna. Hospital Universitario de Valme. Sevilla
(3) Agencia de Evaluación de Tecnologías Sanitarias de Andalucía. AETSA. Consejería de Salud.

RESUMEN

La gran proliferación de guías, protocolos y otros instrumentos de apoyo a la decisión clínica en el Sistema Nacional de Salud español está acompañada de una amplia variabilidad que refleja inconsistencias y baja calidad de estos documentos. Este estudio pretende realizar un inventario de escalas de valoración e instrumentos para el análisis crítico de las Guías de Práctica Clínica y proponer una escala o conjunto de criterios que sirva para evaluar la calidad de las Guías de Práctica Clínica producidas en España. Se realizó una búsqueda sistemática de escalas de valoración crítica. Se aplicaron los criterios de inclusión y el análisis de concordancia de los ítems por tres evaluadores de forma independiente. Las discordancias se resolvieron por consenso explícito. Se identificaron 10 propuestas de escalas de valoración crítica y de grupos de criterios procedentes de 11 instituciones. Ocho son escalas e instrumentos que se proponían evaluar la calidad de las Guías de Práctica Clínica, y dos propuestas para evaluar la implantación y la inclusión de las Guías de Práctica Clínica en un registro. En el análisis comparativo los criterios más repetidos en las escalas analizadas se refieren a las áreas incluidas en el instrumento AGREE. Las áreas consideradas en la mayor parte de las escalas de valoración crítica se corresponden con el Instrumento AGREE. Esta herramienta, aunque no aborda criterios para evaluar la implementación de guías, se considera como una escala adecuada para utilizar en la evaluación previa a la inclusión en un registro nacional de GPC.

Palabras clave: Control de calidad. Investigación sobre servicios de salud. Evaluación de Procesos y Resultados (Atención de Salud). Garantía de la calidad de atención de salud. Guías de práctica clínica. Estándares

Correspondencia:
Rosa Rico Iturrioz.

Osteba. Servicio de Evaluación de Tecnologías Sanitarias.
Departamento de Sanidad. Gobierno Vasco.
Donostia-San Sebastián, 1. 01010 - Vitoria-Gasteiz.
Correo electrónico: rosarico-osteba@ej-gv.es.

(*) El presente trabajo ha recibido financiación del Fondo de Investigaciones Sanitarias en la modalidad Evaluación de Tecnologías Sanitarias. Proyecto FIS: 01/1057 y forma parte de las actividades de la Línea 5 de la Red MBE, FIS: G03/090.

ABSTRACT**Assessment of Clinical Practice Guidelines Evaluation. Scales and Criteria**

Not only are there large number of guides, protocols and other support tools available for the clinical decision-making process in the Spanish National Health System, but there is also a major degree of variability among them, reflecting inconsistencies and low quality of those documents. This study is aimed at conducting an inventory of the Clinical Practice Guideline assessment scales and clinical analysis tools and to propose a scale or set of criteria for assessing the quality of the Clinical Practice Guidelines put out in Spain. A systematic search of critical evaluation scales was conducted. The inclusion criteria and the concordance analysis of the items by three evaluators were independently applied. The discordances were resolved by explicit consensus. Ten suggested critical assessment scales and sets of criteria from eleven institutions were identified, eight of which consist of scales and tools proposed for assessing the quality of the Clinical Practice Guidelines, the other two being proposals for assessing the implementation and inclusion of the Clinical Practice Guidelines in a register. In the comparative analysis, the criteria most often repeated on the scales analysed were related to the areas included in the AGREE Instrument. The areas considered in most of the critical assessment scales were the same as those of the AGREE Instrument. Although this tool does not take in criteria for guide implementation assessment purposes, it is considered suitable for use in the assessment prior to inclusion a national CPG register.

Key words: Process Assessment (Health Care). Practice Guidelines. Quality Assurance, Health Care. Quality control. Health services research. Standards.

INTRODUCCIÓN

En nuestro país existe una gran proliferación de guías, protocolos, vías y normas de actuación clínica. La terminología utilizada para definir los diferentes instrumentos puede resultar confusa y refleja la existencia de productos con diferentes utilidades y de calidad variable. Consideramos que las Guías de Práctica Clínica (GPC) se distinguen por su carácter indicativo y su flexibilidad clínica *versus* la rigidez y carácter normativo de los protocolos, vías y normas¹.

Recientemente se han desarrollado dos estudios dirigidos a conocer el grado de elaboración, uso y evaluación de GPC en España^{2, 3}. Ambos trabajos dibujan un pobre panorama en cuanto a la calidad, adopción y uso de las GPC en el Sistema Nacional de Salud (SNS). Se aprecia una gran variabilidad entre los diferentes y numerosos guías y protocolos que refleja la falta de rigor, de consistencia, de actualización y de estrategias de difusión o integración en programas de mejora de la calidad o de compra de servicios sanitarios. Tampoco existen evaluaciones sobre los efectos en salud derivados de su uso. Asimismo, la pobre aceptación de las guías entre los profesionales es un indicador de su escasa percepción como un valor añadido para la asistencia.

La evaluación de las GPC en España se hace además difícil debido a que frecuentemente hay ausencia de elementos clave para su identificación. Así, la ausencia de fecha, autoría u organismo financiador es una barrera para tenerlas en cuenta como instrumentos válidos para la ayuda en las decisiones clínicas².

De forma ideal, al usuario de una GPC antes de utilizarla le interesa asegurarse de que se puede calificar como una guía de calidad, válida, útil y fiable. Es decir, que los autores han evaluado la evidencia científica antes de trasladarla a las recomendaciones, de modo que se consigan los resultados en

salud que se pretenden, que partiendo de la misma información otra guía pueda alcanzar las mismas recomendaciones y que los diferentes usuarios las interpreten de la misma manera⁴.

Sin embargo, la evaluación de las GPC se suele realizar a partir de elementos subjetivos, tanto para la interpretación de la calidad como para la relevancia clínica de sus conclusiones.

Por todo ello, en los últimos años se está reconociendo la necesidad de evaluar y comparar el proceso y la metodología empleada en la elaboración de GPC. Se trataría de aportar un marco sistemático para la evaluación de la calidad de las GPC con el fin de ayudar: a) a quienes elaboran las GPC para que sigan una metodología estructurada y sistemática, b) a los proveedores de cuidados de salud para que evalúen las GPC antes de adoptar sus recomendaciones en la práctica clínica y c) a los gestores de diferentes niveles y organismos para recomendar GPC para su uso en la práctica.

A partir de trabajos realizados en el marco de programas de elaboración de GPC^{5, 6} en Europa se pusieron en marcha iniciativas de colaboración internacional que tenían como objetivo armonizar los instrumentos y estrategias para la elaboración, evaluación de la calidad (*Appraisal of Guidelines Research and Evaluation in Europe*⁷, AGREE) e implementación (*Changing Professional Practice*)⁸ de las GPC. Proyectos en los que han participado Agencias de Evaluación del Sistema Nacional de Salud (Osteba y AETSA), el Hospital Universitario Valme de Sevilla y la Biblioteca Fundación Josep Laporte. Este tipo de iniciativas se han desarrollado en el marco conceptual de la Medicina Basada en la Evidencia, sus propuestas abordan diferentes aspectos cruciales para elaborar GPC de calidad y hacen hincapié en que las recomendaciones que se emitan estén basadas en pruebas científicas que demuestren su efectividad.

Este estudio se ha desarrollado, en el marco de un proyecto de investigación más amplio para conocer el grado de desarrollo y la calidad de las GPC elaboradas en España, con el objetivo de realizar un inventario de escalas de valoración crítica e instrumentos para el análisis crítico de GPC y proponer una escala o un conjunto de criterios que sirva para evaluar su calidad.

METODOLOGÍA

Se realizó una búsqueda de escalas de valoración crítica de GPC en las siguientes bases de datos: Medline (ver estrategia de búsqueda en tabla 1) INAHTA y National Guidelines Clearinghouse. Dada la naturaleza de los instrumentos y de los documentos a recuperar, en su mayor parte en formato de literatura gris, se contactó con una serie de productores, evaluadores y compiladores de GPC de reconocido prestigio, cuyo listado se recoge en la tabla 2.

A las escalas identificadas y recuperadas se les aplicó los siguientes criterios de inclusión: que fueran instrumentos que abordaran la valoración crítica de guías, la validación o la elaboración de éstas o que fueran criterios propuestos para la inclusión de GPC en registros. La selección de los instrumentos se realizó por tres evaluadores de manera independiente.

Los instrumentos incluidos se agruparon para su clasificación y análisis posterior en tres grupos según su enfoque fuera dirigido a la calidad, implementación o inclusión en un registro (tabla 3). Para realizar el análisis de contenido de los instrumentos, el grupo de trabajo consideró de común acuerdo que las áreas del instrumento AGREE eran las más exhaustivas y establecían la taxonomía más adecuada para evaluar los aspectos de interés de las GPC. La Colaboración AGREE⁷ propuso en 2002 un conjunto de 23 criterios agrupados en 6 áreas relativas al: 1) alcance y objetivo, 2) participación de los implicados,

Tabla 1

Estrategia de búsqueda en Medline

Búsqueda en Medline	
1-	practice guideline [pt]
2-	guideline [pt]
3-	#1 OR #2
4-	guidelines [mh]
5-	practice guidelines [mh]
6-	protocols [mh]
7-	consensus development conferences [mh]
8-	#4 OR #5 OR #6 OR #7
9-	practice guideline*
10-	guideline*
11-	protocol* [ti]
12-	consensus development conference*
13-	recomend* [ti]
14-	consensus*
15-	#9 OR #10 OR #11 OR #12 OR #13 OR #14
16-	clinical [ti]
17-	protocol [ti]
18-	#15 AND #16
19-	medical [ti]
20-	protocol [ti]
21-	#19 AND #20
22-	#3 OR #8 OR #15 OR #18 OR #21
23-	appraisal [ti]
24-	criteria [ti]
25-	tool* [ti]
26-	instrument* [ti]
27-	#23 AND (#24 OR #25 OR #26)
28-	evaluat* [ti]
29-	assessm* [ti]
30-	quality [ti]
31-	#27 OR #28 OR (#29 AND #30)
32-	#22 AND #31

3) rigor en la elaboración, 4) claridad y presentación, 5) aplicabilidad y 6) independencia editorial. Este instrumento ha sido validado internacionalmente y adoptado por un amplio grupo de organizaciones profesionales de investigación, gestión sanitaria y agencias de evaluación de tecnologías sanitarias.

A las áreas del instrumento AGREE se añadieron dos más, no incluidas en éste y

Tabla 2

Listado de direcciones de organizaciones encargadas en la elaboración, evaluación y recopilación de GPC

Organización	Dirección web
AGREE	http://www.agreecollaboration.org
NZGC	http://www.nzgg.org.nz
SIGN	http://www.sign.ac.uk
Health Service Technology Assessment	http://text.nlm.nih.gov
Australian National Health and Medical Research Council	http://www.health.gov.au/nhmrc/publications/cphome.htm
Monash Institute of Health Services and Research	http://www.med.monash.edu.au/healthservices/
Wessex Institute for Health Research and Development	http://www.soton.ac.uk/~wi/
US National Clearinghouse	http://www.guideline.gov/index.asp
UK Clearinghouse on Health Outcomes	http://www.leeds.ac.uk/nuffield/pubs/index.htm
Scottish Health Purchasing Information Centre	http://www.nhsconfed.org/Scotland/shpic/
RAND Health Program	http://www.rand.org/health/
St George's Hospital Medical School	http://www.sghms.ac.uk/depts/phs/hceu/
Guideline Appraisal Project	http://www.cche.net/usersguides/main.asp
Group Health Northwest	http://www.ghc.org/web/
e-guidelines	http://www.ghc.org/web/
Clinical Practice Guidelines. University of California	http://medicine.ucsf.edu/resources/guidelines/
CMA e-practice tools	http://www.cma.ca/cma/common/linkNavigate.do?skin=129
AHRQ	http://www.ahrq.gov/
Alberta Medical Association Clinical Practice Guidelines	http://www.albertadoctors.org/resources/guidelines.html
Cancer Care Ontario	http://www.cancercare.on.ca/
ICSI	http://www.icsi.org/

Tabla 3

Listado de instrumentos, asignación de enfoque y número de criterios que contienen

Instrumento de evaluación de GPC	Grupo	Nº de criterios
Cluzeau et al., 1997	Calidad	*
AGREE, 2002	Calidad	23
SIGN, 2002	Calidad	*
NZGG, 2002	Calidad	*
ACPG, 1999	Calidad	16
NHMRC, 2002	Calidad	4
AHRQ, 2002	Calidad	*
Hayward, 1995	Calidad	10
Nice, 2001	Calidad	8
Shaneyfelt et al., 1999	Calidad	25
AZQ, 2002	Calidad	44
Pilote y Tager, 2002	Implementación	9
National Guidelines Clearinghouse, 2000	Inclusión en registro	4
TOTALES	3	143

* Aunque contaban con criterios propios de evaluación han asumido los 23 criterios del Instrumento AGREE en el 2002 (23).

Tabla 4

Agrupación de criterios en áreas y concordancia con los criterios AGREE

Ámbitos a evaluar	CRITERIOS	número concordancias con áreas y criterios*
Alcance y objetivo	Los objetivos generales de la guía están específicamente descritos.	5
	Los aspectos clínicos cubiertos por la guía están específicamente descritos	6
	Los pacientes a quienes se pretende aplicar la guía están específicamente descritos	4
Participación de los implicados	El grupo que desarrolla la guía incluye individuos de todos los grupos profesionales relevantes	6
	Se han tenido en cuenta los puntos de vista del paciente y sus preferencias	6
	Los usuarios diana de la guía están claramente definidos	3
	La guía ha sido probada entre los usuarios diana	6
Rigor en la elaboración	Se han utilizado métodos sistemáticos para la búsqueda de la evidencia	9
	Los criterios para seleccionar la evidencia se describen con claridad	8
	Los métodos utilizados para formular las recomendaciones están claramente descritos	6
	Al formular las recomendaciones han sido considerados los beneficios en salud, los efectos secundarios y los riesgos	3
	Hay una relación explícita entre cada una de las recomendaciones y las evidencias en las que se basan	7
	La guía ha sido revisada por expertos externos antes de su publicación	7
	Se incluye un procedimiento para actualizar la guía	5
Claridad y presentación	Las recomendaciones son específicas y no son ambiguas	7
	Las distintas opciones para el manejo de la enfermedad o condición se presentan	6
	Las recomendaciones clave son claramente identificables	2
	La guía se apoya con herramientas para su aplicación	3
Aplicabilidad	Se han discutido las barreras organizativas potenciales a la hora de aplicar las recomendaciones	8
	Han sido considerados los costes potenciales de la aplicación de las recomendaciones	7
	La guía ofrece una relación de criterios clave con el fin de realizar monitorización y/o auditoría	5
Independencia editorial	La guía es editorialmente independiente de la entidad financiadora	2
	Se han registrado los conflictos de intereses de los miembros del grupo de desarrollo	2
Implementación	¿Se puede identificar una gran base de datos que contenga información sobre actuaciones médicas para el tratamiento de la condición para la que se ha elaborado una GPC?	1
	A partir de los datos ¿se puede realizar un diagnóstico preciso?	1
	¿Se pueden establecer criterios que permitan la creación de grupos de comparación con diferentes actuaciones médicas?	1
	¿Existen datos que permitan asegurar la comparabilidad de los grupos?	1
	¿Se pueden medir estas actuaciones médicas?	1
	¿Se pueden identificar actuaciones médicas de acuerdo con las recomendadas por las GPC's?	1
	¿Existen datos sobre pacientes, médicos o factores ambientales que podrían explicar desviaciones de lo recomendado en las GPC's y ayudar a validar cualquier inferencia realizada sobre la asociación "actuación médica-resultados"?	1
	Los resultados de interés relacionados con el objetivo de la GPC de mejorar la calidad, uso apropiado y efectividad de la atención médica, ¿están disponibles y se pueden medir con precisión?	1
Las tasas de incidencia y prevalencia de los resultados de interés ¿Son suficientemente grandes para permitir asociaciones útiles sobre "actuación médica-resultados"?	1	
Inclusión en registro	La guía debe de contener recomendaciones desarrolladas sistemáticamente, estrategias u otra información que ayude a la toma de decisiones en situaciones clínicas específicas	1
	La guía ha sido elaborada bajo los auspicios de una organización profesional de reconocido prestigio (p. Ej. Sociedad médica especializada, agencia gubernamental, organización que se ocupa de los cuidados en salud o en el marco de un Plan de Salud)	1
	En el proceso de desarrollo de la guía se ha incluido una búsqueda de la literatura y una revisión de la evidencia existente y publicada en revistas con revisores por pares y este hecho se puede contrastar.	1
	La guía debe ser actual y la versión más reciente tiene que haberse elaborado, validado o revisado en los últimos 5 años.	1
Sin concordancia		7
TOTALES		143

recuperadas de otros instrumentos, son las áreas de implementación e inclusión en registro (tabla 4).

El análisis comparativo de las diferentes escalas seleccionadas se realizó mediante la comparación uno a uno de los criterios respecto a la escala estándar de criterios y ámbitos del instrumento AGREE. Dicha selección de estándar se hizo tras comprobar que cuatro organizaciones analizadas y que previamente poseían criterios propios renunciaron a éstos para asumir el instrumento AGREE (tabla 3). Posteriormente se asignó un número a cada criterio de las escalas recuperadas. Se realizó un análisis de concordancia en base a la similitud de definiciones de cada uno de los criterios frente al estándar por tres evaluadores de forma independiente (tabla 4). En aquellas asignaciones en las que existía discordancia entre los evaluadores ésta se resolvió por consenso explícito.

RECUPERACIÓN Y ANÁLISIS DE ESCALAS DE VALORACIÓN CRÍTICA

El grado de respuesta de los 21 organismos consultados fue del 100%. Se han identificado las propuestas de escalas de valoración crítica y de grupos de criterios de 13 organismos (tabla 3). Ocho de trece se encontraron en una búsqueda de literatura gris. Este hecho refleja una baja inclusión de documentos sobre escalas y grupos de criterios en las bases de datos más reconocidas. Las únicas escalas de valoración crítica de GPC validadas han sido el Instrumento AGREE⁷, y el utilizado por Shaneyfelt et al.⁹.

ESCALAS DE VALORACIÓN DE LA CALIDAD

Se recuperaron 11 escalas para la valoración de la calidad de las GPC, apreciándose solapamiento y cruces entre los instrumentos. Entre éstos es de destacar la propuesta de 1997 *Appraisal Instrument for Clinical guidelines*

(AICG)¹⁰ del St George Hospital Medical School, ya que a partir de ella y con la participación de sus autores, se desarrolló más adelante el instrumento AGREE dentro de un proyecto multicéntrico europeo. Otros organismos como SIGN, AHRQ y NZGG que disponían de su propia escala han adoptado recientemente el instrumento AGREE. Por su parte, el *National Health and Medical Research Council* (NHMRC) de Australia¹¹ al proponer el cumplimiento de sus principios de elaboración recomienda los criterios desarrollados por *Field & Lohr* en 1992¹² y SIGN en 1995⁵.

El *Evidence Medicine Working Group* en 1995, utiliza los criterios basados en *Users guides to Evidence-Based Medicine* publicados en JAMA^{13, 14}.

Hay que destacar que en 1999 el grupo de Shaneyfelt et al. publicó el «*Guidelines Quality Assessment Questionnaire*» (GQAQ)⁹. Este instrumento consta de 25 ítems de los cuales 10 evalúan el desarrollo y formato, otros 10 la identificación de la evidencia y su síntesis y finalmente 5 la formulación de sus recomendaciones. Fue utilizado para evaluar y valorar la calidad de más de 200 guías.

En cuanto al grupo de *Alberta Clinical Practice Guidelines*¹⁵ adopta las dimensiones publicadas por Battista y Hodge¹⁶, y Davis y Taylor-Vaisey¹⁷ y recomienda, asimismo, la propuesta elaborada por Cluzeau en 1999¹⁸.

La Agency for Quality in Medicine Joint Institution of the German Medical Association and the National Association of the Statutory Health Insurance Physicians (ÄZQ)¹⁹ proponen un cuestionario con 21 ítems relacionados con el desarrollo, 16 ítems con el contenido y el formato y 4 ítems con la aplicabilidad de la guía. Para su realización utilizan instrumentos anteriormente elaborados por organizaciones alemanas e internacionales y de versiones previas del Instrumento AGREE.

El National Institute for Clinical Excellence (NICE) del NHS publica en 2001 *The Guideline Development Process – Information for National Collaborating Centres and Guideline Development Groups*²⁰, con los atributos que se exigirán a las GPC. Utiliza los criterios identificados por el programa estadounidense del AHRQ²¹.

Tanto Scottish Intercollegiate Guidelines Network (SIGN)⁵ como New Zealand Guidelines Group²² adoptan las áreas y criterios del Instrumento AGREE en 2002.

Finalmente, aunque fue excluido del análisis ya que no se consideró en nuestro estudio la comparación entre guías, la «National Guideline Clearinghouse» de la Agency for Health Research and Quality, AHRQ²³ utiliza 27 campos para que los usuarios del registro las comparen entre sí.

ESCALAS DE IMPLEMENTACIÓN

Se ha identificado una propuesta para evaluar las GPC en relación con la implementación, Pilote L y Tager IB²⁴. En ella se propone utilizar la investigación de resultados para evaluar las GPC.

ESCALAS PARA LA INCLUSIÓN EN REGISTROS

La *National Guideline Clearinghouse* de la AHRQ²³, propone cuatro criterios básicos para considerar la inclusión de una GPC en su registro (tabla 3). Dichos criterios son diferentes y complementarios a los 27 criterios que utilizan para comparar la calidad de las guías entre sí.

CRITERIOS RECUPERADOS

Se identificaron 143 criterios pertenecientes a 9 escalas de valoración crítica (tabla 4). Del análisis de los criterios se desprende que

el Instrumento AGREE cubre la mayor parte de criterios expresados en otras escalas de análisis y valoración crítica de Guías de Práctica Clínica, resultando la taxonomía más exhaustiva para clasificar los criterios por su capacidad para abordar los diversos aspectos de interés de las GPC.

Este hallazgo valida la decisión del grupo de trabajo que consideró como estándar la clasificación de dominios o áreas propuesta en el Instrumento AGREE para la comparación 1 a 1 de escalas de valoración crítica.

ANÁLISIS DESCRIPTIVO DE LAS ESCALAS DE VALORACIÓN CRÍTICA DE GPC ENCONTRADAS RESPECTO A LA TAXONOMÍA

En la comparación de las diferentes escalas o instrumentos evaluados, los criterios más repetidos se refieren a las siguientes 6 áreas del AGREE (tabla 4): 45 criterios para evaluar el Rigor en la elaboración, 21 para la Participación de los implicados, 20 criterios en relación con la Aplicabilidad y 18 relativos a la Claridad en la presentación. Por el contrario, las áreas Alcance y objetivos e Independencia editorial han sido las menos representadas en los diferentes instrumentos recuperados con 15 y 4 criterios respectivamente.

Respecto a los dos ámbitos no reflejados en el Instrumento AGREE, uno es el de la Implementación de las GPC que está representado por los 9 criterios de la propuesta de Pilote et al²⁴. Otro es el relativo a la inclusión en un registro tal como propone el *National Guidelines Clearinghouse* que contiene 4 criterios. Esta organización es responsable de la base de datos de GPC más amplia del mundo.

COMENTARIOS

La definición de calidad de las GPC es la confianza en que se han evitado los sesgos

potenciales producidos en su desarrollo, que las recomendaciones tienen validez interna y externa y que son útiles para la práctica clínica¹². Por tanto a los instrumentos de evaluación de la calidad de las GPC se les exige valorar estos cuatro aspectos. El conjunto de criterios que mejor lo cumple es el Instrumento AGREE.

Aunque las otras 9 diferentes escalas encontradas y revisadas (tabla 3) sólo cumplen estos criterios parcialmente, consideramos que este hecho es un indicador del interés y esfuerzo realizado en los últimos años por diversas organizaciones para conseguir herramientas que les permitan decidir con rigor qué GPC van a implementar y cuáles no.

Sólo se han encontrado dos escalas que realizan un proceso de validación, la de Shaneyfelt et al, 1999⁹ y el AGREE⁷. El primero considera 25 ítems procedentes de importantes sociedades norteamericanas (*American Medical Association, Institute of Medicine y Canadian Medical Association*), que utiliza para evaluar 279 guías desde 1985 a 1997. En el caso del Instrumento AGREE su validez fue evaluada por 195 personas en 100 guías de 11 países, obteniendo una alta fiabilidad en la mayoría de sus 6 áreas⁷. De esas dos escalas el Instrumento AGREE es el que hasta el momento cuenta con la máxima aceptación por diversos organismos elaboradores y compiladores de GPC, pues además está estructurado en un formato que facilita la cuantificación y la comparación de la evaluación.

Tras el análisis comparativo realizado consideramos que AGREE es un instrumento exhaustivo en la agrupación de criterios, excepto en lo relativo a la evaluación de la implementación de las GPC. Entendemos que AGREE ha descartado ese ámbito ya que se trata de un instrumento genérico y diseñado para ser utilizado en cualquier contexto de aplicación.

De alguna manera AGREE tiene en cuenta los aspectos relativos a la implementa-

ción, ya que exige tanto la realización de una prueba piloto antes de la difusión de la guía para probar su aceptación como la elaboración, por parte de los autores, de indicadores que faciliten la evaluación posterior de los efectos de su aplicación.

La implementación es un ámbito necesario en el caso de auditar a posteriori la utilización de las propias guías evaluadas o las diferencias en la aplicación de las guías en diferentes contextos²⁵ y exige criterios específicos que se fundamenten en un conocimiento apriorístico de la propia GPC y del contexto local. En nuestro análisis hemos mostrado los criterios de implementación e inclusión en un registro por la información que aportan en el proceso continuo de evaluación de GPC.

Por otra parte es interesante observar que las diferentes escalas de valoración crítica analizadas dan menos importancia a los criterios que se agrupan bajo 3 de los epígrafes: «Alcance y objetivos», «Participación de los implicados» e «Independencia editorial». Estos ámbitos que podrían considerarse de tipo formal o de menor trascendencia que los criterios relacionados con la metodología de elaboración son, sin embargo, fundamentales a la hora de describir para qué sirve la guía, a quién se dirige (usuarios-profesionales), a quién se aplica (tipo de pacientes) y la posibilidad de que haya intereses espúreos tras los contenidos aparentes de la guía²⁶.

Identificar la autoría de las guías es fundamental para establecer las responsabilidades, solicitar información adicional y adjudicar méritos. La ausencia de autores puede reflejar varias deficiencias, tales como la existencia de grupos de consenso con metodología no explícita, las traducciones encubiertas de guías o, relacionándolo con el área de la Independencia editorial, podría enmascarar la existencia de conflictos de intereses²⁷.

La búsqueda bibliográfica de evidencia sobre escalas de valoración crítica de GPC ha sido un proceso complejo por las caracte-

rísticas del objeto de búsqueda y su baja indexación. De hecho, en el presente trabajo y aunque se siguieron los criterios sistemáticos de búsqueda en bases de datos primarias (MEDLINE) y secundarias (INAHTA y *National Guidelines Clearinghouse*) se obtuvo un menor rendimiento de recuperación que en la búsqueda de literatura gris. Esto es debido a que las instituciones y programas de producción de GPC son las que tienen elaboradas herramientas de evaluación, y éstas pueden estar publicadas o no según el formato de artículo original. La exhaustividad en la búsqueda de literatura gris ha permitido encontrar más escalas que las que hubieran sido identificadas en la bibliografía, pero por supuesto no nos garantiza que haya otras escalas en uso, de las que desconocemos su existencia. En cualquier caso queda con nuestra búsqueda garantizada que las escalas más difundidas y aceptadas han sido evaluadas.

Una posible limitación de nuestro estudio es la ausencia de un *gold estándar* establecido de comparación de escalas de evaluación de GPC. Ante la inexistencia de un patrón aceptado el grupo de investigación ha optado por una doble aproximación: primeramente, extraer todos los criterios de las diferentes escalas y compararlos entre sí y como segundo paso seleccionar una escala sobre la cual establecer la comparación entre las publicadas, recuperadas y que han realizado un proceso de validación. Así, se encontraron dos: el Instrumento AGREE⁷ y el elaborado por Shaneyfelt et al, 1999⁹.

Graham et al, 2000²⁶ en un estudio similar sobre comparación de escalas de valoración de calidad de GPC agrupó 44 criterios en 10 atributos procedentes de trece instrumentos y realizó un análisis de contenido. En su estudio sólo el instrumento de Cluzeau et al¹⁸, que fue la base del proyecto de la Colaboración AGREE, cubría los 10 atributos propuestos y tenía en cuenta 28 de los 44 criterios generados. Sin embargo no se realizó un estudio de comparación.

En nuestra investigación, se decidió seleccionar el Instrumento AGREE como estándar de comparación por dos razones: su riguroso proceso de mejora y validación finalizado en 2001 y su creciente grado de aceptación por parte de los organismos elaboradores, compiladores y evaluadores de GPC.

Disponer de una herramienta de valoración de guías no es una cuestión menor. La evaluación de guías adquiere una especial relevancia para la recomendación o selección de la GPC que se propone adoptar, adaptar e implementar en un contexto determinado. Hay que tener en cuenta asimismo el valor que tiene el propio proceso de elaboración como garante inicial de calidad²⁸. Así, en algunos países están invirtiéndose energías y fondos para la elaboración de GPC, sin embargo, dichas iniciativas no se corresponden con una homogeneidad en la calidad de las GPC publicadas²⁹. De hecho existe una amplia variabilidad dependiente de las entidades elaboradoras de las mismas y muchas veces derivada de la inexperiencia en la elaboración de GPC por parte de los autores³⁰. Con el fin de garantizar la calidad mínima están surgiendo en estos momentos movimientos e iniciativas similares a las creadas en su día para regularizar la publicación de ensayos controlados y aleatorizados (CONSORT)^{31,32} y sobre los metanálisis^{33,34} que abordan aspectos para la publicación común de GPC. Así, *The Conference on Guideline Standardization (COGS)*²⁵ en 2002 estableció por consenso un instrumento para estandarizar la publicación de GPC, promover su calidad y favorecer así su implementación.

Del mismo modo, la colaboración entre clínicos y metodólogos con experiencia en la elaboración de GPC, así como la existencia de programas locales de elaboración y evaluación de guías que nutran registros estatales es esperable que aseguren la calidad de dichos registros y los productos que en ellos se compilen.

Para concluir, nuestro estudio identifica una serie de escalas de valoración crítica de la calidad de las GPC, entre las que destaca el Instrumento AGREE por su exhaustividad y sistematización de los atributos exigibles a una guía. Además, por su fiabilidad demostrada y formato facilita la comparabilidad entre los evaluadores, individuales o institucionales de las guías.

BIBLIOGRAFÍA

1. Jovell AJ y Navarro-Rubio MD. Guías de Práctica Clínica. Formación Médica Continuada 1995; 2:152-6.
2. REDEGUIAS. Informe de Evaluación de Tecnologías FIS 01/1057. ISCIII 2003. Disponible en: www.redeguias.org.
3. Fundación Salud Innovación y Sociedad. Documento de trabajo nº17: Análisis del cuestionario sobre «Implantación y utilización de guías de práctica clínica: instrumentos, recursos y redes». Proyecto Guía Salud [citado feb 2003]. Disponible en: www.fundsis.org/Actividades/publicaciones/Documento%20de%20trabajo%2017.pdf.
4. Littlejohns P y Cluzeau F. Guidelines for evaluation. *Fam Pract* 2000; 17:s3-s6.
5. Scottish Intercollegiate Guidelines Network (SIGN). Clinical Guidelines: criteria for appraisal for national use. Edinburgh: SIGN; 1995.
6. Cluzeau F y Littlejohns P. Appraising clinical guidelines in England and Wales: The development of a methodological framework and its application to policy. *Jt Comm J on Qual Improv* 1999; 25:514-521.
7. The AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: The AGREE project. *Qual Saf Health Care* 2003; 12:18-23.
8. Thorsen T, Mäkelä M, editores. *Cambiar la Práctica de los Profesionales*. Sevilla: Agencia de Evaluación de Tecnologías Sanitarias de Andalucía, AETSA; 2002.
9. Shaneyfelt TM, Mayo-Smith MF y Rothwangl J. Are Guidelines Following Guidelines? The methodological Quality of Clinical Practice Guidelines in the Peer Reviewed Medical Literature. *JAMA* 1999; 281(20):1900-1905.
10. Cluzeau F, Littlejohns P, Grimshaw J y Feder G. Appraisal instrument for clinical guidelines. St George's Hospital Medical School. London:1997.
11. National Health and Medical Research Council (NHMRC). A Guide to the development, implementation and evaluation of clinical practice guidelines. Camberra;1998.
12. Field MJ, Lohr KN, editors. *Guidelines for Clinical Practice: from development to use*. Washington (DC): Institute of Medicine, National Academy Press; 1992.
13. Wilson MC, Hayward RS, Tunis SR, Bass EB y Guyatt G: for the Evidence –Based Medicine Working Group. Users guides to the medical literature. VIII. How to Use Clinical Practice Guidelines. A. Are the recommendations valid? *JAMA* 1995; 274(7):570-574.
14. Wilson MC, Hayward RS, Tunis SR, Bass EB y Guyatt G: for the Evidence –Based Medicine Working Group. Users guides to the medical literature. VIII. What are the recommendations and will they help you in caring for your patients? *JAMA* 1995; 274(20):1630-1632.
15. Alberta Clinical Practice Guidelines. Evaluation framework for the clinical practice guidelines program. Prepared by Howard Research and Instructional systems. Inc. January 2002.
16. Battista R y Hodge MJ. Setting priorities and selecting topics for clinical practice guidelines. *Can Med Assoc J* 1995; 159:1233-7.
17. Davis DA y Taylor-Vaisey A. Translating guidelines into practice: A systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines. *Can Med Assoc J* 1997; 157:408-15.
18. Cluzeau F, Littlejohns P, Grimshaw J, Feder G y Moran S. Development and application of a generic methodology to assess the quality of clinical guidelines. *Int J Qual Health Care* 1999; 11:21-8.
19. Agency for Quality in Medicine Joint Institution of the German Medical Association and the National Association of the Statutory Health Insurance Physicians (ÄZQ). Checklist «Methodological Quality of Clinical Practice Guidelines». 2nd version (8/1999) Valid until 02/28/2003 [citado sept. 2002]. Disponible en: www.azq.de www.leitlinien.de.

20. National Institute for Clinical Excellence. NICE. The Guideline Development Process – Information for National Collaborating Centres and Guideline Development Groups. N0040; 2001.
21. Field MJ, Lohr KN, editores. Committee to Advise the Public Health Service on Clinical Practice Guidelines. Clinical Practice Guidelines: directions for a new Program. Washington (DC): National Academy Press; 1990.
22. New Zealand Guidelines Group. Tools for Guideline Development and Evaluation 2002 [citado sept 2002]. Disponible en: www.nzgg.org.nz/tools/adaptation.cfm.
23. National Guideline Clearinghouse. AHRQ [citado sept 2002]. Disponible en: www.ahrp.gov/clinic/cpgsix.htm.
24. Pilote L y Tager IB. Outcomes research in the development and evaluation of practice guidelines. BMC Health Serv Res 2002; 2:7 [citado sept 2002]. Disponible en: www.biomedcentral.com
25. Shiffman RN, Shekelle P, Overhage M, Slutsky P, Grinshaw J y Deshpandre AM. Standardized reporting of clinical practice guidelines: a proposal from the conference on guidelines standardization. Ann Intern Med. 2003; 139:493-8.
26. Graham ID, Calder LA, Hebert PC, Carter AO, Tetroe JM. A comparison of clinical practice guideline appraisal instruments. Int J Technol Assess Health Care. 2000; 16: 1024-38.
27. Calderón Sandubete E, Cotos Cancas R, Marín León I, Briones Pérez de la Blanca E, Pérez Cano B, Grilo Reina A y Grupo CAMBIE. Mapa de evidencias para el tratamiento de la angina inestable en la cabecera del paciente. Etapa preliminar para el desarrollo de una guía de práctica clínica. Med Clin (Barc) 2000; 114 (Supl)2:19-23.
28. Choudhry NK, Stelfox HT y Detsky AS. Relationships between authors of clinical practice guidelines and the pharmaceutical industry. JAMA 2002; 287:612-17.
29. Burgers JS, Cluzeau FA, Hanna SE, Hunt C y Grol R. Characteristics of high-quality guidelines: evaluation of 86 clinical guidelines developed in ten European countries and Canada. Int J Technol Assess Health Care 2003; 19:148-56.
30. Marzó M, Alonso P y Bonfill X. Guías de práctica clínica en España. Med Clin (Barc) 2002; 118(Supl 3):30-5.
31. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I et al. Mejora de la calidad de los informes de los ensayos clínicos aleatorios controlados. Recomendaciones del grupo de trabajo CONSORT. Rev Esp Salud Pública 1998; 72 (1): 5-11.
32. Moher D, Schulz KF y Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized-trials. Ann Intern Med 2001; 134:657-62.
33. Rodríguez Artalejo F y Guallar Castellón P. Sobre los meta-análisis y el QUOROM. Rev Esp Salud Pública 2000; 74 (2): 103-5.
34. Moher D J Cook, Eastwood S, Olkin I, Rennie D, Stroup DF. Mejora de la calidad de los informes de los metaanálisis de los ensayos clínicos controlados: el acuerdo QUOROMD. Rev Esp Salud Pública 2000; 74 (2): 107-18.