

Metodología

de investigación y lectura crítica de estudios

Las guías de práctica clínica y el instrumento AGREE II

Iván Darío Flórez Gómez¹
Diana Carolina Montoya²

Resumen

Introducción: La medicina basada en la evidencia surge por la necesidad sustentar la práctica clínica en algo de mayor fortaleza que la experiencia, y de ella se derivan las guías de práctica clínica (GPC), las cuales son documentos que contienen recomendaciones fuertes y válidas desde diferentes enfoques y apoyan la toma de decisiones clínicas. El instrumento AGREE II fue desarrollado para realizar una evaluación crítica de las GPC disponibles en la literatura.

Método: Análisis de la literatura concerniente a las GPC y breve descripción del método de lectura crítica de guías mediante el instrumento AGREE II. *Resultado:* Existe gran variabilidad entre las GPC dirigidas a un tema en particular y pululan las guías de baja calidad, repletas de sesgos y, por lo tanto, con recomendaciones erróneas. Se crea entonces el AGREE y su versión mejorada, AGREE II, el sistema de mayor aceptación para evaluar la calidad de las GPC. Este sistema consiste en un documento de 23 ítems agrupados en 6 dominios diferentes, en los cuales se evalúan el alcance y objetivos, la participación de los implicados en la elaboración de la guía, el rigor en la elaboración de esta, la claridad de la presentación, su aplicabilidad en la práctica clínica y la independencia editorial. *Conclusiones:* El AGREE II puede ser utilizado por clínicos, gestores de políticas en salud, editores de revistas médicas y estudiantes de medicina, entre otros; además, les permite a quienes no sean expertos en metodología de la investigación ni en epidemiología clínica realizar de forma fácil y detallada la evaluación crítica de una GPC antes de decidir si le será de utilidad si conviene aplicarla.

Palabras clave: Guía de práctica clínica, medicina basada en evidencia, salud mental

Title: Clinical Practice Guidelines and the AGREE II Instrument.

Abstract

Introduction: Evidence-based medicine emerges from the need to find enough evidence to support medical practices. It makes these practices stronger and more reliable than solely relying on the experience of the experts. Clinical practice guidelines (CPG) are created from evidence based medicine. These are documents intended to develop strong and valid recommendations from different approaches in order to assist in the making of medical decisions. The AGREE II instrument was created to evaluate the quality of these guidelines. *Methods:*

¹ Pediatra epidemiólogo clínico, jefe del Departamento de Pediatría y Puericultura, Universidad de Antioquia, Medellín, Colombia.

² Médica general, Universidad de Antioquia, Medellín, Colombia.

Analysis of the literature related to the CPG and brief description of the method of critical appraisal of guidelines by the AGREE II instrument. *Results:* There are countless varieties of CPG describing one particular condition. There are also many guidelines of poor quality, containing misleading information with many biases, and with erroneous recommendations. In order to improve CPG quality the AGREE instrument and its new version AGREE II, are the most accepted ways to evaluate medical guidelines. They consist of 23 items grouped into 6 different domains. The AGREE instrument evaluates assessments and objectives, participation of different groups related to the guideline, the clarity and applicability of the CPG, the medical staff participation, and the editorial independence. *Conclusions:* AGREE II can be used by doctors, health policy makers, publishers of medical journals, and medical students among others. This allows people who are not experts in any scientific methodology to more easily detail critical evaluations of any CPG before deciding on their usefulness and applicability.

Key words: Practice guideline, evidence-based medicine, mental health

Introducción

La medicina basada en la evidencia (MBE) surge en las últimas décadas como una necesidad de encontrar en los conocimientos y en los resultados obtenidos de la investigación científica, la evidencia que sustente nuestra práctica clínica en algo de mayor fortaleza y aceptación que la experiencia de los expertos, y se caracteriza por ser una utilización consciente, explícita y juiciosa de la mejor evidencia clínica disponible para tomar decisiones sobre el estado

de salud de un paciente o de una población en particular (1).

Todos los tipos de estudios en la investigación clínica aportan información relevante y todos tienen algún nivel de evidencia: estudios de corte transversal, series de casos, estudios de casos y controles, de cohortes, ensayos clínicos, metaanálisis y guías de práctica clínica (GPC), pero son estas últimas las que tienen mayor aceptación entre los clínicos, porque resumen la información de los otros tipos de estudios y, además, brindan recomendaciones para la toma de decisiones clínicas, basadas en la evidencia.

Las guías de práctica clínica (GPC)

Las GPC son documentos que, por medio de un análisis exhaustivo de la literatura acerca de un problema clínico con relevancia para una población o institución específica, buscan la elaboración de recomendaciones desde diferentes enfoques (preventivo, diagnóstico, tratamiento y/o rehabilitación) para ayudar al clínico en la toma de decisiones y, así, establecer la manera como será abordado el paciente que presenta dicha condición (2).

Su principal característica es la selección exquisita de la literatura médica por medio de un detallado y concienzudo análisis de esta, para emitir recomendaciones fuertes y válidas que puedan ser puestas en práctica en diferentes contex-

tos asistenciales. De esta manera, las GPC promueven intervenciones con claro beneficio y desalientan aquellas ineficaces para reducir la morbilidad, mortalidad y/o mejorar la calidad de vida; además, son la mejor herramienta disponible para homogeneizar la práctica clínica, es decir, disminuir la variabilidad clínica indeseada y actualizarla con el más reciente estado del arte en un tópico específico (2,3). Todo esto, indudablemente, permite al clínico ahorrar una importante cantidad de tiempo que se hubiese invertido en las búsquedas y lectura de la literatura, y se traduce en mejoría en los costos de la atención de los pacientes.

El impacto de las GPC va más allá de solo resumir y analizar la literatura, y aportar recomendaciones para el clínico, ya que adicionalmente pueden apoyar en la toma de decisiones en torno a un sistema de salud en una región o país (4). A pesar de que todo lo anterior se encuentra relativamente sedimentado y aceptado, aún no se tienen a disposición GPC adecuadas para todas las enfermedades relevantes y prioritarias en los sistemas de salud de nuestros países no desarrollados (5).

La construcción de una GPC es un proceso estructurado y riguroso que demanda tiempo, dedicación y recursos importantes; es considerado una investigación original integrativa, por ello, intentar desarrollar una de ellas debe ser una decisión muy cautelosa (6,7). Se debe considerar

la construcción de una GPC en los casos de enfermedades con elevada prevalencia o con elevada mortalidad, cuando se observe un alto grado de variabilidad clínica en su manejo y en aquellos casos en los que se haga necesario definir las recomendaciones considerando las estrategias más costo-efectivas (7,8).

Además de estos criterios para considerar desarrollar una GPC, existen algunas condiciones adicionales que no deben ser negociables en los procesos de desarrollo. La metodología debe ser clara (los métodos de búsqueda de la literatura y evaluación de la evidencia deben quedar explícitos y estar al alcance de los usuarios), los objetivos, el alcance, los usuarios y la población objeto de la guía deben ser generales, pero claros y concisos (nunca ambiguos), sin dejar de reconocer y exponer explícitamente que existen diferencias individuales entre los pacientes; por ende, deben permitir la variación racional y sustentada de la práctica clínica. Por último, los temas o enfermedades escogidos para desarrollar una GPC deben contar con la suficiente evidencia en la literatura médica para su análisis (2,8).

En los últimos años, muchas instituciones, entes académicos, sociedades científicas, entre otros, se han adjudicado la tarea de desarrollar guías de numerosos temas, publicarlas e, incluso, implementarlas. Pero, a pesar del “boom” de las GPC en el mundo, con frecuencia vemos una enorme variabilidad en-

tre las guías de un tema específico. Esto ocurre porque muchas de ellas carecen de un rigor metodológico adecuado que garantice la calidad de la información analizada y de las recomendaciones elaboradas (9-11). Además, con frecuencia se elaboran guías con adecuada metodología, son bien desarrolladas, pero no se consideran los aspectos necesarios para que se puedan implementar en el quehacer diario, y permanecen en el papel, sin poder llegar a cumplir su objetivo: brindar la mejor atención a los pacientes (12).

Una verdadera GPC basada en la evidencia es una herramienta fuerte y confiable para la toma de decisiones clínicas, pero, ¿cómo identificar estas guías entre un mar de ellas que carecen de rigor metodológico y, por lo tanto, son inadecuadas?

Para responder esta pregunta, inicialmente se debe evaluar si la guía que tiene en sus manos es una verdadera GPC basada en la evidencia. Para ello, esta debe poseer tres características fundamentales: constar de una estrategia estructurada, sistematizada y reproducible de búsqueda de la literatura y selección de la evidencia; tener bien especificado el proceso de evaluación y síntesis de la evidencia; por último, los métodos de formulación de las recomendaciones deben ser claros, estructurados y explícitos. Estos tres puntos deben estar claramente descritos en la metodología expuesta por la guía, bien sea en la versión original o, en ocasiones, en la pági-

na web de la institución o sociedad científica que produce la guía, y deben considerarse fundamentales para considerar una guía como una GPC (6).

La estrategia de búsqueda y la selección de la evidencia son pasos esenciales para responder adecuadamente las preguntas que se construyan dentro de la elaboración de una GPC, y aunque no es claro cuántas y cuáles deben ser las fuentes de información que deben utilizarse, si existen recomendaciones bastante aceptadas sobre cómo debe ser dicha búsqueda. Se ha recomendado consultar las bases de datos más importantes, como MedLine, Embase, Central y PsycINFO, que abarcan la gran mayoría de las publicaciones importantes en lengua inglesa (13). Pero, para los casos como el de nuestra región, deben considerarse indudablemente de forma adicional las bases Lilacs y SciELO, para abarcar la literatura de lengua hispana y portuguesa en nuestro continente (14).

Luego de la búsqueda de la literatura, los procesos de calificación de la evidencia y formulación de recomendaciones son pasos fundamentales, ya que a partir de ellas se obtiene el producto definitivo. De la calidad de la selección y calificación de la evidencia dependerá la fortaleza de las recomendaciones y, por ende, gran parte de la calidad de la guía (15). Ante la necesidad de facilitar el proceso de calificación de la evidencia se han creado numerosas escalas que la categorizan según el

tipo de estudio y la probabilidad de contener sesgos; de esta manera, estas escalas ayudan a determinar la calidad de los estudios y aportan las pautas para que a partir de allí se formulen las recomendaciones y la fuerza de estas (15).

De dichas escalas, el sistema SIGN (Scottish Intercollegiate Guidelines Network) (16) es el que ha tenido mayor tradición, disponibilidad y diseminación en los últimos años. Pero ante la abundancia de sistemas y escalas que buscaban el mismo objetivo, en el 2000 se creó una colaboración informal de investigadores interesados en abordar las deficiencias de las anteriores formas de clasificación, con el fin de unificarlas en un solo sistema, y por ello se creó el grupo de trabajo GRADE (Grading of Recommendations Assessment, Development and Evaluation) (17-19). El sistema GRADE, si bien con menor diseminación dada su menor antigüedad, por ser una iniciativa global de agregación que busca superar las limitaciones de todos los sistemas anteriores se constituye, a la fecha, en el sistema de mayor aceptación y recomendación en el mundo.

¿Cuáles GPC pueden ser usadas en el escenario clínico diario?

Una revisión de la calidad de las GPC producida por organizaciones importantes entre 1988 y 1998 mostró que de 431 guías analizadas, el 67% no describen a los actores

implicados (del inglés *stakeholders*, sin traducción exacta en el idioma español), 88% no reportan cómo se identificó la evidencia, 82% no indican la fuerza de las recomendaciones y solamente 5% de estas describen estos tres importantes criterios (10,20). Es decir, de las innumerables GPC disponibles en la literatura, solo unas pocas cuentan con un rigor metodológico mínimo y describen los procesos de su construcción adecuadamente, como la búsqueda, la síntesis de la evidencia y la generación de recomendaciones. Queda claro, entonces, que en la literatura médica en general pululan las guías de baja calidad, con gran cantidad de sesgos y, por lo tanto, constituidas por recomendaciones erróneas. Por todo lo anterior, aparece el instrumento AGREE, que permitió poner las GPC ante el microscopio del escrutinio público y aumentar el nivel de calidad en su construcción (21).

¿Qué es el instrumento AGREE y en qué consiste?

El grupo internacional de investigadores y elaboradores de guías, The AGREE collaboration (*Appraisal of Guidelines, Research and Evaluation*), fue creado en 1998 y su objetivo primordial fue la creación de un instrumento con el cual se evaluara el desarrollo y el reporte de las GPC. El instrumento AGREE se construyó bajo una rigurosa metodología, fue publicado por primera vez en 2003 y rápidamente se con-

solidó como una herramienta útil, de fácil aplicación, y como la mejor estrategia para evaluar la calidad y la posibilidad de aplicación, entre otros aspectos, de las GPC (22).

Posteriormente, se creó '*Next steps consortium*', proyecto que tenía como objetivo mejorar las propiedades métricas del instrumento, incluyendo fiabilidad y validez; refinar los ítems para servir mejor a las necesidades de los usuarios potenciales y mejorar la documentación de apoyo; para este propósito se hicieron dos estudios (23,24). El resultado de estos esfuerzos fue el rediseño del instrumento, es decir, la creación de una nueva y mejorada versión a la que se denominó AGREE II, que se encuentra publicada y disponible en el sitio web de la colaboración AGREE³ (25,26).

Esta nueva versión logró mejorar la inicial en varios aspectos, como ampliar el alcance del instrumento más allá del ámbito clínico (abrió la posibilidad de que se utilizara por personas encargadas de tomar decisiones en salud pública y en los sistemas de salud, no solo por clínicos) e incluye un nuevo ítem relacionado con la evaluación de la validez interna y de las fortalezas y debilidades de la evidencia analizada en la guía (25). Este instrumento ha tenido excelente acogida y reconocimiento, ha sido citado en más de 200 publicaciones y es utilizado por

muchas organizaciones, entre ellas The National Institute for Health and Clinical Excellence (NICE), del Reino Unido (21). Adicionalmente, existe una reciente versión del AGREE II, traducida al español en el mismo sitio web de la colaboración (27), y se ha demostrado la facilidad de aplicación y su superioridad frente a las otras escalas disponibles para el mismo objetivo (28).

El AGREE II, sus dominios e ítems

El AGREE II consiste en un documento de 23 ítems agrupados en 6 dominios diferentes, seguidos de 2 ítems de puntuación global (evaluación global de la calidad de la guía) (27). Está diseñado para evaluar la calidad de las GPC por medio de un espectro de salud, proporciona una estrategia metodológica para el desarrollo de las guías y establece qué información debe estar presentada en las guías y de qué forma debe hacerse (24,25). Este instrumento está diseñado para ser utilizado por proveedores de atención en salud (médicos generales, especialistas, estudiantes de medicina), elaboradores de guías, gestores de políticas en salud y educadores en medicina. Por su fácil y rápida utilización, se puede aplicar a guías sobre promoción de la salud, tratamiento, intervenciones médicas, salud pública, entre otros.

³ Véase en <http://www.agreetrust.org>.

Una enorme ventaja con que cuenta el instrumento es que no hay necesidad de ser un experto en epidemiología clínica para aplicar el AGREE II, aunque, claro está, los conocimientos en estas áreas facilitarán su utilización. Dependiendo de la estructura y del tamaño de la guía, se calcula que la aplicación del instrumento debe tardar en promedio 1,5 horas por evaluador (25).

El AGREE II se puede usar por grupos constructores de guías, por clínicos, por gestores de políticas en salud (para decidir la implementación en un sistema de salud), por editores de revistas médicas (ayuda a definir los requerimientos mínimos de reporte de las guías en las publicaciones); por último, el AGREE II será un marco de referencia para lograr consensos en la metodología y en los reportes en los procesos de colaboración internacional (25).

A continuación se detallan los seis dominios de la herramienta y las preguntas que contiene cada uno de ellos. Cada dominio aborda un aspecto o área distinta de las características de una guía, como se verá a continuación. En la Tabla 1, en la columna de la izquierda se detallan las preguntas originales de la versión anterior del AGREE y en la columna derecha, los ítems del AGREE II, si los hubiese.

*Dominio 1: alcance y objetivos
(ítems 1-3)*

Esta área se encarga de evaluar la claridad con la que se especifica

el alcance y los objetivos, y, por ello, este aspecto se relaciona directamente con el potencial impacto que puede llegar a tener en la población, así como las fronteras de la revisión y los criterios de inclusión y exclusión de pacientes. Se analiza mediante tres preguntas (Tabla 1).

*Dominio 2: participación de los
implicados (ítems 4-7)*

Este dominio incluye la evaluación de los aspectos relacionados con los participantes en su construcción. Es importante que para cada enfermedad se incluyan todos los grupos de profesionales (especialidades médicas, médicos generales y profesiones afines a medicina relacionadas con el tema, por ejemplo profesionales en nutrición, enfermería y fisioterapia, etc.), que tenga en cuenta un representante de los pacientes o de los padres de niños con la enfermedad en cuestión (para el caso de guías sobre temas pediátricos). Además, deben quedar específicamente descritos los usuarios de la guía (personas o profesionales que usarán la guía para la toma de decisiones). Se evalúa este dominio mediante tres ítems (Tabla 1). En la primera versión del AGREE se contemplaba un cuarto ítem que fue suprimido (Tabla 1).

*Dominio 3: rigor en la elaboración
(ítems 8-14)*

Es uno de los dominios más importantes y el que marca la pauta de

Tabla 1. Ítems y dominios de los instrumentos AGREE y AGREE II

Dominios	Ítem del AGREE original	Ítem del AGREE II
Dominio 1. Alcance y objetivo	1. El(los) objetivo(s) general(es) de la guía está(n) específicamente descrito(s).	Sin cambios.
	2. El (los) aspecto(s) clínico(s) cubierto(s) por la guía está(n) específicamente descrito(s).	El (los) aspecto(s) de salud cubierto(s) por la guía está(n) específicamente descrito(s).
	3. Los pacientes a quienes se pretende aplicar la guía están específicamente descritos.	La población (pacientes, público etc.) a la cual se pretende aplicar la guía esta específicamente descrita.
Dominio 2. Participación de los implicados	4. El grupo que desarrolla la guía incluye individuos de todos los grupos profesionales relevantes.	Sin cambios.
	5. Se han tenido en cuenta los puntos de vista del paciente y sus preferencias.	Se han tenido en cuenta los puntos de vista de la población diana (pacientes, publico, etc.).
	6. Los usuarios diana de la guía están claramente definidos.	Sin cambios.
	7. La guía ha sido probada entre los usuarios diana.	Ítem suprimido. Incorporando en la descripción del ítem 19 de la guía del usuario.
Dominio 3. Rigor en la elaboración	8. Se han utilizado métodos sistemáticos para la búsqueda de la evidencia.	Sin cambios en el ítem. Renumerado como 7.
	9. Los criterios para seleccionar la evidencia se describen con claridad.	Sin cambios en el ítem. Renumerado como 8. Nuevo ítem 9: Las fortalezas y limitaciones
	10. Los métodos utilizados para formular las recomendaciones están claramente descritos.	Sin cambios.
	11. Al formular las recomendaciones han sido considerados los beneficios en salud, los efectos secundarios y los riesgos.	Sin cambios.
	12. Hay una relación explícita entre cada una de las recomendaciones y las evidencias en las que se basan.	Sin cambios.
	13. La guía ha sido revisada por expertos externos antes de su publicación.	Sin cambios.
	14. Se incluye un procesamiento para actualizar la guía.	Sin cambios.

(Continúa)

(Continuación)

Dominios	Ítem del AGREE original	Ítem del AGREE II
Dominio 4. Claridad de la presentación	15. Las recomendaciones son específicas y no son ambiguas.	Sin cambios.
	16. Las distintas opciones para el manejo de la enfermedad o condición se presentan claramente.	Las distintas opciones para el manejo de la enfermedad o condición de salud se presentan claramente.
	17. Las recomendaciones clave son fácilmente identificables.	Sin cambios.
Dominio 5. Aplicabilidad	18. La guía se apoya con herramientas para su aplicación.	La guía proporciona consejo y/o herramientas sobre como las recomendaciones pueden ser llevadas a la práctica. Cambio de dominio (de Claridad de la presentación) y renumerando como 19.
	19. Se han discutido las barreras organizativas potenciales a la hora de aplicar las recomendaciones.	La guía describe factores facilitadores y barreras para su aplicación. Renumerado como 18.
	20. Han sido considerados los costos potenciales de la aplicación de las recomendaciones.	Se han considerado las posibles implicaciones de la aplicación de las recomendaciones, sobre los recursos.
	21. La guía ofrece una relación de criterios clave con el fin de realizar monitorización y/o auditoría.	La guía se ofrece para monitorización y/o auditoría.
Dominio 6. Independencia editorial	22. La guía es editorialmente independiente de la entidad financiadora.	Los puntos de vista de la entidad financiadora No han influido en el contenido de la guía.
	23. Se han registrado los conflictos de intereses de los miembros del grupo de desarrollo.	Se han registrado y abordado los conflictos de intereses de los miembros del grupo elaborador de la guía.

la calidad de la información. Se basa en siete ítems dirigidos a determinar la calidad en la búsqueda de la información (que sea sistematizada, reproducible, amplia, fundamentada en bases de datos adecuadas y completas, entre otros aspectos), los

criterios de inclusión y exclusión de los trabajos, la calidad de las recomendaciones, la consideración de los riesgos y beneficios, la relación entre la evidencia y las recomendaciones, la existencia de una revisión externa de la guía antes de su publicación y,

por último, la consideración por los autores de la actualización de esta.

En relación con la versión anterior, fue el dominio que sufrió las modificaciones más importantes, como la generación de un nuevo ítem (ítem 9), que se refiere a la necesidad de que la guía exprese detalladamente “las fortalezas y limitaciones del conjunto de la evidencia”. Es el dominio más grande y, con frecuencia, el de mayor requerimiento de tiempo para ser evaluado. Debe ponerse mucha atención a los resultados de estos ítems, ya que, en pocas palabras, este es “el dominio que determina la calidad de la guía”.

Dominio 4: claridad de la presentación (ítems 15-17)

Este dominio, relacionado con el lenguaje en que se expresan las recomendaciones, la estructura de la presentación y su formato, evalúa la calidad de la presentación de las recomendaciones (que sean específicas, no ambiguas; de fácil identificación, y que al expresarlas contemplen todas las opciones de manejo posibles reconocidas). En este punto desempeña un papel importante la apariencia, la facilidad de visualizar de forma rápida las recomendaciones. Este dominio consta de cuatro ítems (Tabla 1).

Dominio 5: aplicabilidad (ítems 18-21)

Este es uno de los dominios que con mayor frecuencia incumplen las

guías, ya que en la literatura abundan las GPC que se esmeran por llevar a cabo una adecuada metodología de búsqueda, selección y recomendaciones, pero bien sea porque los autores no se percatan muchas veces de la importancia de generar estrategias de aplicabilidad, o bien sea porque el objetivo de la GPC es meramente académico y no es competencia del grupo desarrollador proponer la forma de aplicar las recomendaciones, no se considera la implementación de la guía. Aquí se analizan las posibles barreras para implementar las GPC, así como los costos potenciales y la necesidad de establecer las estrategias para monitorizar dicha aplicación. Las preguntas van dirigidas a evaluar que tan fácil puede ser que esa guía se lleve a la práctica que se desea (Tabla 1).

Dominio 6: independencia editorial (ítems 22 y 23)

Este dominio evalúa la independencia del grupo que desarrolla la guía y, por lo tanto, la independencia de los resultados de la GPC, de la entidad que la financió. En este ítem se buscan activamente los registros de los potenciales conflictos de intereses de cada uno de los integrantes del grupo, y se basa en los ítems 22 y 23. No se debe olvidar que los conflictos no son solamente económicos, sino que también pueden ser de interés académico, gubernamental o político. Una GPC que no exprese la existencia o la

inexistencia de conflicto de intereses debe ser analizada con cautela, y la que exprese conflictos importantes entre sus autores, debe tenerse en cuenta dicha situación a la hora de analizar las recomendaciones particulares que pueden relacionarse con dicho conflicto (Tabla 1).

La evaluación y la recomendación de la guía

La calificación de cada ítem, en la primera versión del AGREE, se basaba en una escala de 4 puntos (4: fuertemente de acuerdo; 3: de acuerdo; 2: en desacuerdo; 1: fuertemente en desacuerdo), pero la nueva versión se basa en una escala de 7 puntos, desde el valor 1 (fuertemente en desacuerdo) hasta el 7 (fuertemente de acuerdo) (12) (Tabla 1). Los valores entre estos extremos permiten un distinto nivel de acuerdo o desacuerdo, según considere el evaluador.

De esta manera, cada ítem puede recibir una calificación entre 1 y 7. Al final de la evaluación se determina la calificación para cada dominio, al sumar las calificaciones de cada ítem y calcular un porcentaje, siendo 100% la máxima calificación posible por obtener (7 en todos los ítems del dominio) y 0%, la mínima (calificación de 1 en todos los ítems del dominio). Esta ponderación a porcentaje permite una mejor interpretación de la calidad de cada dominio. Se recomienda que cada GPC sea evaluada, idealmente, por

dos a cuatro personas, para aumentar la confiabilidad del instrumento.

Además, al final de la herramienta aparecen dos criterios finales y globales de la guía, que son: la calificación global de toda la guía con la misma escala, es decir, entre 1 (la calidad más baja posible) y 7 (la más alta calidad posible), y la respuesta a la pregunta que establece la herramienta: ¿recomendaría esta guía para su uso? En este punto, queda a discreción del (de los) evaluador (es) la decisión, con base en lo evaluado hasta el momento, de la calificación numérica descrita y la decisión de recomendar o no la guía.

Aunque las puntuaciones de los dominios pueden ser útiles para comparar guías y para recomendar su empleo, no han establecido puntuaciones mínimas o patrones de puntuación entre dominios que diferencien entre guías de alta y baja calidad; por lo tanto, no se recomienda crear puntos de corte para tal efecto, aunque algunos grupos han sugerido que una GPC se debe considerar de alta calidad y puede ser recomendada si el puntaje de todos los dominios es superior al 60% (29). Estas decisiones deben hacerse por el usuario y guiadas por el contexto en el cual se utilice el AGREE II (27).

Ventajas e inconvenientes de la herramienta

Es una herramienta de fácil utilización, que puede ser usada por clínicos, pacientes y encarga-

dos de la toma de decisiones en los sistemas de salud. Es un instrumento muy completo, ya que valora todas las dimensiones importantes de la GPC (desde la metodología, la presentación de la información, la aplicabilidad, hasta la independencia editorial). Las dos grandes desventajas de la herramienta son el requerimiento de un tiempo de aplicación relativamente largo, y que se recomienda que haya, mínimo, dos evaluadores. Ambas situaciones pueden hacer ver el instrumento como poco práctico para el clínico que requiere toma de decisiones rápidas, pero teniendo en cuenta que disponer de una herramienta adecuada para la evaluación de guías no es una nimiedad —dado que con base en ellas se toman desde decisiones clínicas individuales, locales e institucionales, hasta grandes decisiones en un sistema de salud de un país o región—, vale la pena invertir el tiempo en una juiciosa evaluación de las guías antes de considerar su utilización.

Conclusiones y consideraciones para los clínicos

Las GPC son recomendaciones basadas en la evidencia científica disponible para asistir a profesionales de la salud, a los encargados de la creación de políticas en salud, así como a los propios pacientes, en la toma de decisiones respecto a una condición clínica en particular o algún procedimiento clínico, con

el fin de que este sea el más apropiado, aceptado y seguro. Las GPC abundan en la literatura, pero infortunadamente la mayoría de ellas no cuentan con un rigor metodológico aceptable, lo que puede traer consigo la generación de recomendaciones con profundos sesgos y, por lo tanto, con errores importantes.

La herramienta AGREE II es una interesante propuesta de evaluación de la calidad de las guías, que permiten al clínico que no es experto en metodología de la investigación realizar, de una forma fácil y detallada, la evaluación crítica de una GPC antes de decidir si se aplicará. Es importante que todos los clínicos conozcan y dominen esta herramienta, y que aprendan a realizar una evaluación crítica de todas las guías que consideren puedan ser de utilidad en su práctica clínica, lo cual les permitirá cometer menos errores, conocer la manera correcta de filtrar la información masiva que se encuentra en la literatura, discernir entre la información que realmente tiene sustentación científica y la que no, y, de esta manera, brindar el mejor manejo o el mejor procedimiento ante una condición clínica determinada.

Referencias

1. Sackett DL, Rosenberg WMC, Gary JAM, et al. Evidence based medicine: What is it and what it isn't. *BMJ*. 1996;312:71-2.
2. Pardo R. Manual metodológico para la elaboración y adaptación de guías de práctica clínica basadas en la evidencia

- 1ra ed. Bogotá: Universidad Nacional de Colombia; 2009.
3. Ortiz Z, Esandi ME, De Luca M. Desafíos en la elaboración e implementación de Guías de prácticas clínicas en Argentina. *Bol Acad Nacional de Medicina*. 2005;83:232-9.
4. Oxman AD, Freitheim A, Schünemann HJ. Improving the use of research evidence in guideline development: introduction. *Health Res Policy Syst*. 2006;4:4-12.
5. Alemán A, Cafferatta ML, Tomasso G, et al. Guías clínicas basadas en la evidencia. En: Sanchez-Mendiola M. *Medicina Pediátrica basada en evidencia* 1ra ed. México: Ed. McGraw Hill Interamericana; 2004. p. 197-210.
6. Lozano JM, Cuervo LG. Desarrollo de guías de práctica clínica. En: Ruiz A, Morillo LE. *Epidemiología clínica*. 1ra ed. Bogotá: Editorial Panamericana; 2004. p. 373-92.
7. Bergan BA. Evidence based guidelines and critical pathways for quality improvement. *Pediatrics*. 1999;103:225-32.
8. American Academy of Pediatrics. Task force on ethical management guidelines. Guiding principles, attributes and process to review medical management guidelines. *Pediatrics*. 2001;108:1378-82.
9. Shaneyfelt TM, Mayo-Smith MF, Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. *JAMA*. 1999;281:1900-5.
10. Grilli R, Magrini N, Penna A, et al. Practice guidelines developed by specialty societies: the need for critical appraisal. *Lancet*. 2000;355:103-6.
11. Burgers JS, Fervers B, Haugh M, et al. International assessment of the quality of clinical practice guidelines in oncology using the Appraisal of Guidelines and Research and Evaluation Instrument. *J Clin Oncol*. 2004;22:2000-7.
12. Graham RP, James PA, Cowan TM. Are clinical guidelines valid for primary care? *J Clin Epidemiol*. 2000;53:949-54.
13. Crumley ET, Wiebe N, Cramer K, et al. Which resources should be used to identify RCT/CCTs for systematic reviews: a systematic review. *BMC Medical Research Methodology*. 2005;5:24.
14. Clark OA, Castro AA. Searching the literatura latino Americana e do caribe em ciencias da Saude (LILACS) database improves systematics reviews. *Int J Epidemiol*. 2002;31:112-4.
15. Atkins D, Briss PA, Eccles M, et al. Systems for grading the quality of evidence and recommendations II. Pilot study of a new system. *BMC Health Serv Res*. 2005;5:25.
16. Scottish Intercollegiate Guidelines Network SIGN 50. A guidelines developer's handbook. Edinburgh: SIGN [internet]. 2008 [citado: 10 de junio del 2011]. Disponible en: www.sign.ac.uk.
17. Marzo-Castillejo M, Alonso-Coello P, Rotaecche del Campo R. Clasificación de la calidad de la evidencia y fuerza de las recomendaciones. *Aten Primaria*. 2006;37:5-8.
18. Guyatt G, Oxman AD, Akl EA, et al. GRADE Guidelines: 1. Introduction – GRADE evidence profiles and summary of finding tables. *J Clin Epidemiol*. 2011;64:383-94.
19. GRADE Working Group. Grading the the quality of evidence and the strength of recommendations [internet]. 2000 [citado: 10 de junio del 2011]. Disponible en: <http://www.gradeworkinggroup.com/>.
20. Burls A. AGREE II – improving the quality of clinical care. *Lancet*. 2010;376:1128-9.
21. Dans AL, Dans LF. Appraising a tool of guideline appraisal (the AGREE II instrument). *J Clin Epidemiol*. 2010;63:1281-22.
22. 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.
23. Brouwers M, Kho ME, Browman GP, et al. Development of the AGREE II. part 1: performance, usefulness and areas for improvement. *CMAJ*. 2010;182:1045-52.
24. Brouwers M, Kho ME, Browman GP, et al. Development of the AGREE II, part

- 2: assesment of validity of items and tools to support application. CMAJ. 2010;182:E472-8.
25. Brouwers M, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. J Clin Epidemiol. 2010;63:1308-11.
26. AGREE Next Steps Consortium. The AGREE II Instrument [Electronic version] [internet]. 2009 [citado: 5 de Julio del 2011]. Disponible en: <http://www.agreetrust.org>.
27. AGREE Next Steps Consortium. El Instrumento AGREE II Versión electrónica [internet]. 2009 [citado: 5 de julio del 2011]. Disponible en: <http://www.guiasalud.es>.
28. Iturrioz RR, Gutiérrez-Ibarluzea I, Batarrita JA, et al. Valoración de escalas y criterios para la evaluación de guías de práctica clínica. Rev Esp Salud Pública. 2004;78:457-67.
29. German Instrument for Methodological Guideline Appraisal. Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI) [internet]. 2005-2006 [citado: 5 de junio del 2011]. Disponible en: www.english.delbi.de.

Conflictos de interés: Los autores manifiestan que no tienen conflictos de interés en este artículo.

Recibido para evaluación: 10 de mayo del 2011

Aceptado para publicación: 3 de agosto del 2011

Correspondencia
Iván Darío Flórez Gómez
Facultad de Medicina
Universidad de Antioquia
Calle 67 No. 50-27
Medellín, Colombia
ivoflorez@gmail.com
ivoflorez@hotmail.com