



Los problemas de calidad de las guías de práctica clínica

resumen ■ **Objetivo:** Presentar el debate en torno a la calidad de las Guías de Práctica Clínica (GPC). **Material y métodos:** se realizó una revisión en *Medline* de la literatura sobre los problemas de calidad de las GPC actualizada a fecha de 1 de enero de 2012 y se evaluó la información publicada por el *Institute of Medicine* sobre el tema. **Resultados:** se describen problemas como el elevado número de guías para un mismo tema, variabilidad de las recomendaciones, los conflictos de interés, el lenguaje de los niveles de evidencia y fuerza de las recomendaciones y, finalmente, su aplicabilidad en la clínica. **Conclusiones:** no todos los documentos que se autodenominan GPC lo son realmente. Hay que evaluar con cuidado la calidad de la evidencia en la que se basan las recomendaciones. **Palabras clave:** guías de práctica clínica, evidencia científica, calidad.



ACCESO ABIERTO

IDOIA GAMINDE INDA
Servicio de Investigación, Innovación y Formación Sanitaria
Departamento de Salud. Navarra

TERESA HERMOSILLA GAGO
C.P.E. Dos Hermanas
Hospital Universitario Virgen de Valme (Servicio Andaluz de Salud)

«Lo primero y más importante es que las guías no pueden ser mejores que la evidencia que las sustenta¹».

«El verdadero problema –insiste David Healy–, es que la industria farmacéutica controla la evidencia científica en la que se basa la práctica médica²».

Introducción

En las últimas décadas el volumen de publicación de Guías de práctica Clínica (GPC) ha aumentado de manera importante, a lo que se ha sumado la publicación de varias guías para un mismo problema o enfermedad. Por ejemplo, el trabajo publicado por Matthys y cols³ identifica 10 GPC de faringitis con recomendaciones diferentes, lo que hace que para el clínico sea realmente complicado decidir cuál es la GPC que debe de tener en cuenta. Decir que siga las GPC basadas en la evidencia no parece suficiente, ¿en qué GPC se señala que está o no está basada en la evidencia? Y, si el objetivo de las GPC era ordenar el tema y reducir la variabilidad, se consigue lo contrario con tantas publicadas sobre un mismo problema.

Además, la propia noción de GPC depende de quién haga referencia a ella. McAlister y cols⁴ afirman que, para el grupo elaborador, las GPC basadas en la evidencia se definen como aquellas que incorporan una búsqueda sistemática de las pruebas empíricas, evaluando de manera explícita la calidad de la evidencia y proponiendo recomendaciones basadas en las mejores pruebas disponibles, incluso cuando estas no sean de calidad elevada. Sin embargo, para los clínicos, el concepto “basado en la evidencia” se interpreta, a menudo, como que las recomendaciones están basadas solamente (y siempre) en evidencia de elevada calidad. De hecho, no se suele citar la fuerza de la recomendación ni la investigación en la que se basa, sino simplemente “lo dice la guía”.

El objetivo de este artículo es revisar el estado actual de las GPC, siguiendo muy de cerca la propuesta recientemente realizada por el IOM (*Institute of Medicine*) sobre GPC, sugerir algunas reflexiones sobre si las guías pueden ser parte del problema que pretenden resolver y cómo pueden identificar los clínicos las guías en las que confiar, o mejor, las recomendaciones en las que confiar.

¿Qué son las GPC?

El objetivo de las GPC –y de la Medicina Basada en la Evidencia (MBE)– es dotar al trabajo clínico de una fuerte base científica para lograr coherencia, eficiencia, efectividad, calidad y seguridad en la atención clínica. Se inician en un contexto de aumento de costes de la atención sanitaria, desigualdad en el acceso al sistema y enormes variaciones en la práctica clínica, debidas tanto a un uso excesivo, infrautilización o una mala utilización de la atención clínica. En este sentido, el *Dartmouth Atlas Project* lleva informando de la enorme variación en los EEUU desde hace más de 20 años⁵. También trabaja en esta línea el Grupo Atlas VPM⁶ (Variación en la Práctica Médica) en España.

La propuesta frente al problema citado es el desarrollo de herramientas estandarizadas –como pueden ser GPC, Vías Clínicas o Protocolos– que recuperen, evalúen y sintetizen la evidencia, resuman los beneficios y riesgos, y determinen la adecuación de la intervención⁷ para disminuir la variabilidad.

Las Guías de Práctica Clínica fueron definidas en 1990 como un conjunto de “recomendaciones desarrolladas de forma sistemática para ayudar a los profesionales y a los pacientes a tomar decisiones sobre la atención sanitaria más apropiada, y a seleccionar las opciones diagnósticas o terapéuticas más adecuadas a la hora de abordar un problema de salud o una situación clínica específica⁸”.

Veinte años más tarde, el *Institute Of Medicine* (IOM) de los EEUU, por motivos que ahora explicaremos, se ha visto obligado a revisar el tema en profundidad dando lugar a dos informes claramente relevantes: “Clinical Practice Guidelines We Can Trust⁹” y “Finding What Works in Health care: Standards for Systematic Reviews¹⁰”.

Para empezar, propone una nueva definición de GPC: “Documentos informativos que incluyen recomendaciones que pretenden optimizar la atención a los pacientes, que están basados en una revisión sistemática de la evidencia y una evaluación de los beneficios y riesgos de las diferentes opciones alternativas de atención⁹”.

Definen las revisiones sistemáticas (RS) como un tipo de investigación científica que se centra en una determinada pregunta –clínica- y que utiliza métodos científicos explícitos y planificados para identificar, seleccionar, evaluar y resumir los resultados de estudios similares pero diferentes. Puede incluir, o no, una síntesis cuantitativa de los resultados de los diferentes estudios (metanálisis)¹⁰. Por su metodología estructurada, explícita, sistemática y multidisciplinar en la recogida de la información, la valoración crítica de los estudios y la síntesis de los mismos, se diferencian metodológicamente de las revisiones clásicas de la literatura científica sobre un tema. En estas últimas, en las que un experto revisa los estudios publicados, decide cuáles son relevantes y resalta sus resultados, sin que se describa habitualmente el proceso seguido hasta llegar a las conclusiones.

¿Por qué el IOM se ha visto obligado a realizar una evaluación crítica de las GPC?

Volumen de GPC y la variabilidad de las recomendaciones

El volumen de elaboración de GPC es sorprendente. La red internacional de guías GIN (*Guidelines International Network*) tiene en su base de datos más de 3.814 guías de 39 países diferentes (acceso 25 de octubre de 2011). La *National Guideline Clearinghouse* de los EEUU contiene 2.654 GPC de 285 organizaciones diferentes (acceso 25 de octubre de 2011).

El problema, sin embargo, no es tanto el elevado volumen de GPC –que lo es-, sino el solapamiento de estas. Es frecuente que exista un elevado número de GPC para un mismo problema y, además, con recomendaciones diferentes. Por ejemplo, como ya se apunta en la introducción, Matthys y cols³ identifican 10 GPC sobre faringitis elaboradas por diferentes grupos, partiendo de la hipótesis de que las recomendaciones sobre el mismo problema deberían de ser similares, pero no es así. Señalan que la evidencia utilizada es diferente entre las GPC de Norte América y Europa, las primeras no citan Ensayos Controlados Aleatorizados (ECA) europeos relevantes, ni siquiera la revisión Cochrane sobre el tema¹¹. A veces, incluso

Existe una gran variabilidad en la calidad de las GPC y en sus recomendaciones para un mismo tema

las recomendaciones son diferentes aún utilizando los mismos ECA⁴. Veamos más ejemplos.

Diabetes Mellitus

Un análisis de 15 GPC para la diabetes tipo 2 elaboradas en 13 países mostraba que la longitud de las guías oscilaba entre 3 y 350 páginas. El solapamiento sobre la evidencia citada por las guías era bajo, un 18% (185/1033) de las citas aparecían en una u otra guía y había 10 estudios (1%) que aparecían en 6 o más guías. Este dato aumenta si se tienen en cuenta las revisiones. Predomina la investigación realizada en los EEUU (40% de las citas) y hay una tendencia significativa a citar evidencia originada en el propio país. A pesar de todo, los autores concluyen que hay un elevado grado de consenso con respecto a las recomendaciones¹².

También nos encontramos con cambios no explicados en los niveles de la evidencia entre las recomendaciones. Por ejemplo, la ADA (*American Diabetes Association*), en sus Estándares de Cuidados Médicos en Adultos, calificó en 2008 como “A” –evidencia procedente de ECA bien diseñados- la siguiente recomendación: “uso de aspirina (ASA, 75-162 mg/día) como una estrategia de prevención primaria en aquellos pacientes diabéticos tipo 1 o 2 con riesgo cardiovascular incrementado, incluyendo aquellos > 40 años o con factores de riesgo adicional (historia familiar de enfermedad coronaria, hipertensión, fumadores, dislipidemia o microalbuminuria)”. En 2009, la calificación inicial “A” se cambió a “C” –evidencia apoyada por estudios mal controlados o evidencia inconsistente-. En 2010, cambiaron el perfil de riesgo aunque mantuvieron el mismo nivel de la evidencia¹³⁻¹⁵.

El interés sobre la calidad de las GPC sigue siendo motivo de preocupación e investigación en este campo, como demuestra una reciente revisión sistemática de las recomendaciones de las GPC

La calidad de las GPC puede verse seriamente comprometida por los conflictos de interés de sus autores con la industria farmacéutica

sobre el uso de antidiabéticos orales para diabetes mellitus tipo 2¹⁶.

Cardiología

McAlister y cols.⁴ estudiaron la calidad de la evidencia en la que se basaban las recomendaciones de la gestión del riesgo cardiovascular de 9 GPC basadas en la evidencia para pacientes con diabetes mellitus, dislipidemia o hipertensión. La longitud de estas oscilaba entre 10 y 284 páginas, dando lugar a un total de 1.005 recomendaciones. De ellas, solo 338 se centraban en la gestión del riesgo cardiovascular y citaban la evidencia en la que se sustentaban las recomendaciones. De estas 338, 231 (68%) citaban un ECA o RS para apoyar la recomendación y, finalmente, solo 105 (45%) tenían un grado A del CHEP (*Canadian Hypertension Education Program*), equivalente a “evidencia de elevada calidad” de GRADE). La no obtención de un grado elevado se debía a que los ECA incluidos se realizaron en poblaciones con características diferentes a las de los pacientes a los que iba dirigida la GPC.

Por ejemplo, la recomendación del uso de espirolactona para reducir la tensión arterial se basaba en un ECA en el que los participantes tenían insuficiencia cardiaca con una tensión arterial normal. Se rebajó la calificación de otras 59 recomendaciones porque su evidencia se sustentaba en ECA que no se habían centrado en variables sólidas de eficacia clínica sino en variables subrogadas.

Con respecto al uso de las estatinas y el punto de corte del colesterol, la ausencia de evidencia y las recomendaciones de las guías también son otro punto de controversia. Hayward y Krumholz¹⁷ señalan que los objetivos terapéuticos y los puntos de corte para iniciar el tratamiento basados en c-LDL divergen de la evidencia clínica, pues lo que se recomienda no ha sido testado en ningún ensayo clínico.

Ginecología y obstetricia

En un reciente estudio sobre la evidencia que sustenta las recomendaciones del “*American College of Obstetricians and Gynecologists*”, se mostró que solo un tercio de las recomendaciones se basan en pruebas de elevada calidad. Atribuyen el problema a la carencia de ECA cuando se compara con otras disciplinas como cardiología, por ejemplo. Señalan que, comparando sus recomendaciones con las del “*Royal College of Obstetricians and Gynaecologists*”, se ve que solo un 28% de las recomendaciones eran las mismas; un 56% no eran comparables y un 16% eran de carácter opuesto.

Por ejemplo, en el caso concreto de “parto vaginal después de parto previo mediante cesárea”, la calidad de los estudios de las 136 referencias –exceptuando un ensayo prospectivo– era solo de nivel II-A (estudios de cohortes bien diseñados o estudios caso-control, preferentemente de más de un centro o grupo de investigación)¹.

Habría que señalar aquí, aunque es un tema en el que nos detendremos más adelante, el significado de niveles de evidencia y fuerza de las recomendaciones. El objetivo de la guía es establecer una serie de recomendaciones a partir de las pruebas científicas disponibles. El nivel o grado de evidencia clínica es un sistema jerarquizado, basado en las pruebas o estudios de investigación, que ayuda a los profesionales de la salud a valorar la fortaleza o solidez de la evidencia asociada a los resultados obtenidos de una estrategia terapéutica. En función del rigor científico del diseño de los estudios, pueden construirse escalas de clasificación jerárquica de la evidencia, a partir de las cuales pueden establecerse recomendaciones respecto a la adopción de un determinado procedimiento o intervención sanitaria. La fuerza de la recomendación significa hasta qué punto podemos confiar en que poner en práctica una recomendación conllevará más posibles beneficios que riesgos.

Acabamos de ver algunos ejemplos de los problemas de las guías: diferencias de volumen; excesivo número de GPC para un mismo problema; diferentes interpretaciones de los mismos estudios; recomendaciones con falta de ECA y/o RS (de estudios de calidad) y el extraño y variado lenguaje de los niveles de evidencia y la fuerza de las recomendaciones.

Sigamos ahora con más detalle otras cuestiones que abundan en los problemas de “credibilidad” de las GPC.

Las recomendaciones basadas en la opinión de los expertos. El método GOBSAT (*Good Old Boys Sat Around the Table / Los colegas sentados alrededor de la mesa*)

Cuando las recomendaciones se sustentan en la opinión de “expertos” en vez de en las pruebas científicas, nos damos de lleno con el problema que las GPC deberían de resolver. El método GOBSAT, utiliza el nivel más bajo de evidencia posible, se fundamenta en la opinión de los expertos, en su experiencia clínica y es evidente que el sesgo puede ser marcado, como veremos después en el problema de los conflictos de interés.

El elevado número de GPC y las diferentes revisiones hace que las recomendaciones “de obligado cumplimiento” –de clara eficacia probada– se pierdan en la ventisca de las menos importantes. Según Tricoci y cols.¹⁸, un ejemplo relevante son las GPC del *American College of Cardiology (ACC) / American Heart Association (AHA)* de 2002 y 2007 sobre la evaluación previa a la operación de los pacientes que van a ser sometidos a cirugía programada no cardíaca. El análisis de las guías muestra importantes incoherencias entre la evidencia que se cita y las recomendaciones asociadas. Por ejemplo, los autores de las guías señalan discordancias en los argumentos que apoyan las pruebas de estrés y la terapia con betabloqueantes y, sin embargo, sus algoritmos proponen esta intervención. Esto da lugar a un conjunto de recomendaciones poco claras, no sustentadas en la evidencia, que pueden llevar al clínico a iniciar tratamientos “para curarse en salud”¹⁹.

Los conflictos de interés. ¿Es posible la objetividad en la elaboración de las recomendaciones?²⁰

Muchas de las denominadas GPC son informes de consenso de expertos –método GOBSAT–, por lo que no debería sorprendernos que las GPC producidas por el *American College of Cardiology (ACC) / American Heart Association (AHA)* hayan cambiado a más recomendaciones de tipo II (pruebas conflictivas y/o disparidad de opiniones sobre la utilidad/eficacia de un determinado procedimiento o tratamiento) y que, en un 48% de las ocasiones, estas recomendaciones se basen en el nivel más bajo de la evidencia (nivel C: opinión de expertos, estudios de caso, o estándares de cuidados). Solo 245 de las 1305 recomendaciones de nivel I tenían una evidencia A¹⁸.

SE Nissen²¹ en *Archives of Internal Medicine* se pregunta “¿Podemos confiar en las GPC cardio-

“Lo pone en la GPC”
no es sinónimo de
recomendación sólida.
Hay que fijarse en el
nivel de evidencia
científica para cada
recomendación

vasculares?”, comentando un artículo de TB Mendelson²² en el que el autor se pregunta sobre la independencia y fiabilidad de las GPC en medicina cardiovascular. Informan de la presencia de relaciones financieras con entidades comerciales en más de la mitad los 498 autores de las 17 GPC cardiovasculares más importantes publicadas entre 2004 y 2008. Más de la mitad de los elaboradores de guías habían trabajado como ponentes promocionales de la industria y un número importante de estos tenía acciones en las compañías afectadas por las GPC. Además, los líderes de las guías tienen un conflicto potencial de interés con respecto a los miembros del grupo elaborador (81% frente a 55%, $p=0,03$), lo que sugiere que la selección de líderes elaboradores de guías está claramente sesgada.

Abundando en el tema, ni si quiera se informa de los potenciales conflictos. Una evaluación de 431 GPC elaboradas por sociedades científicas mostró que un 67% no informó de la composición del grupo elaborador, impidiendo la evaluación de los potenciales conflictos de interés²³.

Según un editorial del “*Canadian Medical Association Journal*”, en 2003 un comité responsable de la GPC de la *Canadian Diabetes Association (CDA)* recomendó que se utilizara insulina glargina como alternativa a una insulina NPH de larga duración en el tratamiento de la diabetes mellitus tipo I y II en pacientes con problemas de control de la glucosa en ayunas o con hipoglucemia. En el otoño de 2005, los miembros del grupo *Common Drug Review (CDR)*, recomendaron no incluir ese medicamento. Lo sorprendente es que los dos grupos de expertos evaluaron prácticamente la misma información procedente de unos 20 ECA. Es interesante resaltar que las GPC de la CDA no informan acerca de si los miembros de su panel de expertos tienen algún tipo de relación financie-

ra con la empresa elaboradora de la insulina glargina²⁴. El editorial también hace referencia a los potenciales conflictos de interés en las recomendaciones sobre el tratamiento de la hipertensión, pues las GPC nacionales recomiendan el uso de caros medicamentos con patente frente a los compuestos genéricos.

Una reciente investigación sobre la composición de los grupos elaboradores de guías –que define como documentos que orientan el diagnóstico y tratamiento de los pacientes–, muestra que más de un tercio de los autores declararon relaciones financieras con la industria farmacéutica, afectando a un 70% de los grupos elaboradores. En un caso, todos los miembros del grupo habían cobrado dinero de la compañía responsable del medicamento que se acabó recomendando en la GPC²⁵. En este mismo estudio, sobre más de 200 guías internacionales incluidas en la *National Guideline Clearinghouse* en 2004, solo 90 (45%) incorporaban detalles individuales de los conflictos de interés de los autores. En la mitad de las GPC que dieron detalles individuales había un autor asesor en la industria, más de un tercio de los grupos tenían al menos una persona que daba seminarios en nombre de la compañía y, una de cada 10, contaba con un miembro que tenía acciones en la compañía cuyo producto estaba incluido en la guía.

Parece, por tanto, que la relación entre los grupos elaboradores de las guías y la industria farmacéutica es un tema clave por las implicaciones financieras que la recomendación de un determinado medicamento puede tener. Por ejemplo, tras la publicación de una recomendación no favorable al uso de rosiglitazona en una declaración de consenso de la ADA/EASD, el precio de las acciones de GlaxoSmithKline cayó un 20%²⁶.

Parece existir un elevado grado de interacción entre los autores de GPC y la industria farmacéutica. Esta interacción puede influir en la práctica de un elevado número de clínicos²⁷. Pero, y esto es clave, el problema puede ser aún más grave cuando ni siquiera se informa de los potenciales conflictos de interés.

¿Cómo se gestionan los valores de los pacientes?²⁸

De Kort et al²⁹ estudiaron los valores (implícitos) de los clínicos en la elaboración de recomendaciones en oncología paliativa. Estos juicios de valor pueden ser motivo de variación en las recomendaciones de diferentes GPC. Encontraron que en las

GPC había una tendencia a preferir la supervivencia sin tener en cuenta la calidad de vida que puede esperarse cuando no hay opciones curativas. Sin embargo, estos valores de los miembros del grupo elaborador, que no de los pacientes, no se explicitan en la GPC. Si tenemos en cuenta que los valores de los pacientes y los médicos pueden diferir, si estos no se explicitan en la GPC, pueden considerarse como evidencia –ojo, como pruebas científicas–, cuando deberían ser una herramienta que permitiera discutir los pros y los contras de una determinada opción con los pacientes.

Hoy en día, la mayor parte de las agencias elaboradoras de GPC proponen la incorporación de pacientes en la elaboración de guías³⁰. Se postula que la participación de estos incrementará la calidad de las GPC y, además, reflejará sus valores y preferencias. Sin embargo, no hay evidencia, o esta es muy escasa, de que por el hecho de que participen pacientes en la elaboración la GPC los resultados sean diferentes. No hemos de olvidar que las GPC se sustentan en la existencia de investigación de calidad. Es más, Groh³¹ manifiesta que, aún siendo un tema interesante, todavía no es una práctica habitual y además no está claro cuál es la mejor manera de incorporar sus valores y preferencias. ¿Cómo debería concretarse la participación de los pacientes? ¿el paciente que participa en la GPC representa a todos los pacientes?

La evaluación de la calidad de las GPC

Aunque definir la calidad de una GPC no es algo sencillo, en principio, una guía de calidad debería mejorar la salud de los pacientes. Además, debería de ser válida científicamente, utilizable y fiable. Una de las herramientas propuestas para evaluar la calidad de las GPC es AGREE (*Appraisal of Guidelines Research and Evaluation*)^{32,33}. AGREE pretende evaluar el proceso de desarrollo de la guía pero no valora el contenido clínico ni la calidad de la evidencia que sustenta las recomendaciones. Consta de 23 ítems, organizados en 6 dominios: alcance y objetivos, participación de los implicados, rigor en la elaboración, claridad de la presentación, aplicabilidad e independencia editorial. Los criterios se evalúan con una escala Likert de 4 puntos que mide la intensidad con la que se cumple el criterio. Permite realizar una evaluación global –subjettiva– con respecto a si es “recomendable o no”.

En un estudio realizado en 2005³⁴ se identificaron 278 GPC producidas en España entre enero de 1990 y diciembre del 2002. Se recuperaron 61 guías a las que se pasó el Instrumento AGREE de

forma independiente por cuatro evaluadores. Se analizó la calidad de las guías, la factibilidad y reproducibilidad del AGREE. Para todos los atributos del AGREE, excepto en la independencia editorial, más del 60% de las guías evaluadas obtienen una puntuación de mala calidad, no habiendo ningún dominio en el que más del 10% de las guías sea excelente. Las áreas de peor calificación son las de aplicabilidad, participación y rigor en el desarrollo de las guías. De las 61 guías evaluadas, los revisores califican a 25 de ellas de ‘muy poco recomendables’, a 26 de ‘no recomendables’ y tan solo a 6 y 4 guías como ‘recomendables’ o ‘muy recomendables’, respectivamente.

Una evaluación reciente de 7 guías de depresión de diferentes países (cinco de las cuales fueron elaboradas después de que se publicara el instrumento AGREE en 2003) mostraron puntuaciones bajas en términos generales. Respecto al valor de referencia del 100%, que supone un acuerdo máximo, se obtuvieron puntuaciones entre 25 y 65% en la **participación de los implicados** (se centra en el grado en el que se representan las percepciones de los usuarios interesados), 1-64% en **rigor en la elaboración** (proceso de síntesis de la evidencia), 0-56% en **aplicabilidad** (implicaciones organizativas, comportamentales, de costes), y 8-75% en **independencia editorial** (independencia con respecto a las recomendaciones y conflictos de interés). Para los autores, esto muestra importantes limitaciones con respecto a la relevancia y aplicabilidad de las guías estudiadas en atención primaria³⁵.

¿Es tu evidencia 1++, 1+ 1a o nivel 1? ¿Nos entendemos cuando decimos 2c o 2+D? Evidencia y fuerza de la recomendación: sistemas de gradación de la evidencia

En el desarrollo de una GPC, los conceptos de calidad, nivel de evidencia y grado de recomendación forman el eje central de la definición de GPC basada en la evidencia. Son los instrumentos que intentan estandarizar y proporcionar a los clínicos reglas sólidas para valorar la investigación publicada, determinar su validez y resumir su utilidad en la práctica clínica³⁶. Este es un proceso que debería ser transparente y realizarse de forma sistemática, pero ya hemos visto antes que no es infrecuente que haya diferentes recomendaciones ante la misma evidencia.

La estrategia de separar los niveles de evidencia y los grados de recomendación tuvo su origen en la *Canadian Task Force on Preventive Health Care* (CTFPHC, 2005). La CTFPHC fue creada en el año

Las GPC hacen recomendaciones generales pero siempre hay que valorar si son adecuadas en cada paciente concreto

1976 a partir de una Conferencia de los Ministros de Sanidad de las diez provincias de Canadá. La CTFPHC, en el desarrollo de su metodología, contempló el peso de la evidencia científica para hacer recomendaciones, ya fueran a favor o en contra de incluir intervenciones de prevención en las personas asintomáticas (CTFPHC, 2005). En esta metodología se introduce la noción de que la evidencia científica se presenta de manera jerárquica y que algunos diseños de estudio están más sujetos a sesgos que otros y, por lo tanto, justifican en menor medida las decisiones clínicas³⁶.

Hay múltiples clasificaciones para evaluar y estructurar la evidencia y las diferentes interpretaciones de los grados de las recomendaciones –posiblemente más de 100–. Los códigos utilizados son de tres tipos: letras (A, B, C, etc.), números (I, II, III, etc.) y una mezcla de letras y números (Ia, Ib, IIa, etc.). Estas diferentes clasificaciones parece que confunden más que ayudan. Por ejemplo, la administración de anticoagulación oral a pacientes con fibrilación auricular y problemas reumáticos de la válvula mitral reciben diferentes grados de recomendación según las diferentes organizaciones: Clase ‘I’ basado en un nivel de evidencia ‘B’ por parte de la *American Heart Association*, una recomendación ‘C’ basada en un nivel de evidencia ‘IV’ por parte de SIGN y un grado ‘1C+’ por parte del *American College of Chest Physicians*³⁷, donde ‘1’ indica el equilibrio riesgo/beneficio y ‘C+’ la calidad metodológica de la evidencia que la sustenta.

Esto es un problema serio porque un enfoque estandarizado en la evidencia debería de reducir la variabilidad y, sin embargo, el hecho de que la nomenclatura no sea coherente incorpora problemas³⁶. Para empezar, las GPC dan un peso diferente al consenso pues algunas de ellas permiten la incorporación del consenso en sus declaraciones. Esto incorpora una dimensión opaca en la

Queda por resolver en qué manera se puede incorporar la visión de los pacientes a las GPC

forma en la que se realizan las recomendaciones y compromete la objetividad que los clínicos demandan de la evidencia.

Curiosamente, la creación de toda esta cantidad de clasificaciones no está informada por la investigación sino por la opinión de expertos –¡una vez más!-. Schünemann et al.³⁷ no pudieron encontrar ningún trabajo de investigación que sustentara las diferentes formas de presentación de grados de evidencia y recomendaciones.

Problemas de aplicabilidad en la clínica

Las declaraciones sobre la eficacia clínica dominan las GPC y asumen a “pacientes ideales sin comorbilidades”. A menudo, no reflejan correctamente las cuestiones relevantes para la atención sanitaria cotidiana como la seguridad y la gestión de riesgos, la colaboración multidisciplinar, el efecto en los costes o la adherencia y la autogestión de los pacientes. El análisis de 7 guías sobre depresión mostró que estas estaban sobre todo centradas en el tratamiento farmacológico y que prestaban muy poca atención a las terapias psicológicas, al riesgo de suicidio y otras cuestiones sociales³⁸.

En realidad, la mitad de los pacientes con algún problema crónico tiene al menos otros dos problemas con interacciones potenciales que requieren cuidados complejos y apoyo³¹. Sin embargo, las GPC a menudo están formuladas por especialistas en un tema concreto y con un determinado modelo de paciente en la cabeza.

En una reciente revisión³⁹ se constata que las GPC no orientan correctamente las comorbilidades. Tras revisar las GPC de 4 problemas crónicos, se concluyó que posiblemente sean de poca aplicabilidad para ese tipo de pacientes con problemas de comorbilidad. Muchas de estas guías no orientan de manera explícita sobre el tratamiento cuando se dan determinadas combinaciones de estas

enfermedades. Cuando en las GPC se hace referencia a la comorbilidad, se hace de manera genérica y se presentan muy pocas recomendaciones de tratamiento para los pacientes con comorbilidades. Es más, la evidencia que sustenta las recomendaciones relacionadas con la comorbilidad es limitada, de calidad moderada a baja y, a menudo, no se expresa correctamente en las GPC. Lo más llamativo es que estas limitaciones no se describen adecuadamente, lo que puede dar lugar a una confianza en las recomendaciones inadecuada.

A pesar de los excelentes ejemplos de adopción de las recomendaciones, los estudios muestran que las GPC se utilizan en solo un 50-70% de las decisiones diarias y que la variabilidad es enorme. Parece ser que una posible explicación de la poca relevancia para clínicos y pacientes puede ser que se escriben como si de un manual de medicina se tratara y no como un breve conjunto de recomendaciones para la práctica (un máximo de 5-10 páginas), o no son compatibles con las normas y los valores de los usuarios a los que va dirigida³¹.

¿Y ahora que?

“No es oro todo lo que reluce”, no todos los documentos que se autodenominan guías son realmente GPC y no todas cumplen los requisitos necesarios como los que propone el IOM (ver caja 1).

Además, en las GPC puede haber recomendaciones basadas en la evidencia y otras recomendaciones basadas en el consenso, por lo que es clave tener en cuenta la calidad de la evidencia en la que están basadas las recomendaciones.

Queremos señalar que en el documento sobre “elaboración de Guías” del *Scottish Intercollegiate Guidelines Network (SIGN)*⁴⁰ se afirma que: “Hay que recordar que las Guías de SIGN pretenden ayudar a tomar decisiones en la clínica no a reemplazarlas. Las guías no tienen la respuesta para cada una de las posibles preguntas clínicas, ni garantizan un resultado correcto en cada una de las situaciones. La decisión última sobre un determinado procedimiento clínico o tratamiento dependerá siempre de la situación individual de cada paciente, sus circunstancias y deseos, y el juicio clínico del equipo sanitario”

Lo ideal sería seleccionar las GPC de calidad –que cumplan los requisitos como los mencionados en la caja 1–, incluso, aún mejor, sería seleccionar recomendaciones sólidas basadas en probada eficacia.

Existen organizaciones elaboradoras de Guías como NICE –*National Institute for Clinical Excellence*–, SIGN en el Reino Unido y el Proyecto Guía-Salud en España que han desarrollado procesos estructurados de elaboración de GPC que garantizan la calidad en el proceso de elaboración. De hecho, hay que señalar que entre los objetivos de GuíaSalud se encuentran los siguientes:

- Promover el desarrollo –elaboración, adaptación, actualización– de Guías de Práctica Clínica (GPC) y de otros productos basados en la evidencia científica (OPBE) siguiendo una metodología común.
- Aportar metodología para desarrollo de GPC y de otros productos basados en la evidencia científica de calidad
- Facilitar el acceso de las GPC y OPBE al conjunto del Sistema Nacional de Salud (SNS). GuíaSalud cuenta con un catálogo de GPC que cumplen una serie de criterios de inclusión con diferentes versiones.

Finalizaremos con una reflexión de Bonfill, X y Marzo, M⁴¹: *“Lógicamente, la existencia de una buena GPC que aborde un determinado problema de salud es un elemento necesario para cumplir los requisitos de explicación, fundamentación y transparencia de los procesos clínicos a los que aludíamos al principio, aunque no es suficiente para modificar la práctica clínica existente ni tan siquiera para asegurar la propia implantación de la guía. También es necesario que se den las condiciones de dirección, gestión, formación, motivación, etc., imprescindibles para que un documento útil y bien elaborado, como puede ser una GPC, no se quede únicamente en el armario. Hemos de crear las condiciones para que las GPC, una vez elaboradas y validadas, salgan de los armarios y normalicen su existencia como un componente esencial, dinámico y definitorio de la política clínica de cada servicio, hospital o centro de atención primaria”*.

Caja 1. Criterios de credibilidad de las GPC según el Informe “Clinical Practice Guidelines We can Trust”

Para tener credibilidad las GPC deberían:

- Basarse en una revisión sistemática de la evidencia disponible.
- Estar desarrolladas por un grupo multidisciplinario de expertos informados y representantes de los grupos clave afectados.
- Tener en cuenta a los subgrupos importantes de pacientes y las preferencias de estos, según corresponda.
- Basarse en un proceso explícito y transparente que minimiza la distorsión, los sesgos y los conflictos de interés.
- Proporcionar una explicación clara de las relaciones lógicas entre las diferentes opciones de tratamiento y los resultados en salud, y proporcionar índices tanto de la calidad de la evidencia como de la fuerza de las recomendaciones.
- Reconsiderarse y revisarse según corresponda cuando las nuevas pruebas disponibles justifiquen la modificación de las recomendaciones.

Conclusiones

No todos los documentos que se autodenominan GPC realmente lo son.

Es clave tener en cuenta la calidad de la evidencia en la que están basadas las recomendaciones.

Las GPC pretenden ayudar a tomar decisiones en la clínica, no a reemplazarlas.

Existen organizaciones que han desarrollado procesos estructurados de elaboración de GPC que garantizan la calidad del proceso.

Publicaciones de interés

1. Castiñeira C, Medrano F. ¿Cómo leer una GPC? En: Guías de Práctica Clínica Conceptos básicos sobre su elaboración y utilización. A Coruña: Casiterides SL; 2007.

2. Gisbert JP, Alonso-Coello P, Pique JM. ¿Cómo localizar, elaborar, evaluar y utilizar guías de práctica clínica? *Gastroenterol Hepatol* 2008;31:239-57.

3. Martín Muñoz P, Ruiz-Canela Cáceres J, Guerra de Hoyos JA, Rivas Aguayo L. Guías de práctica clínica en Internet: cómo separar el grano de la paja. *Rev Pediatr Aten Primaria*. 2003;5:73-88

4. www.guiasalud.org cuenta con el Catálogo de GPC del Sistema Nacional de Salud, producidas en España, en cualquiera de las lenguas oficiales en el Estado, por grupos de profesionales y que han de cumplir una serie de criterios de calidad para ser incluidas.

5. La sección “herramientas y recursos” de guía salud cuenta con un interesante listado sobre elaboración, actualización e implementación de guías en castellano [<http://portal.guiasalud.es/web/guest/herramientas-gpc>]

Bibliografía

1. Wright JD, Pawar N, Gonzalez JSR, Lewin SN, Burke WM, Simpson LL, et al. Scientific Evidence Underlying the American College of Obstetricians and Gynecologists' Practice Bulletins. *Obstet Gynecol* 2011; 118(3):505-512

2. Elliot C. *White Coat, Black Hat: Adventures on the Dark Side of Medicine* Boston:Beacon Press, 2010.

3. Matthys J, De Meyere M, van Driel ML, De Sutter A. Differences Among International Pharyngitis Guidelines: Not Just Academic. *Ann Fam Med* 2007;5(5):436-443.

4. McAlister FA, van Diepen S, Padwal RS, Johnson JA, Majumdar SR. How Evidence-Based Are the Recommendations in Evidence-Based Guidelines? *PLoS Med* 2007;4(8):e250.

5. The Dartmouth Atlas of Health Care. [<http://www.dartmouthatlas.org/>]

6. Atlas VPM. Atlas de Variaciones en la Práctica Médica en el Sistema Nacional De Salud [<http://www.atlasvpm.org/avpm/>]

7. Timmermans S, Mauck A. The Promises And Pitfalls Of Evidence-Based Medicine. *Health Aff* 2005; 24(1):18-28.

8. IOM (Institute of Medicine). *Clinical Practice Guidelines: Directions for a New Program* Washington, DC: National Academia Press, 1990.

9. IOM (Institute of Medicine). *Clinical Practice Guidelines We Can Trust*. Washington, DC: The National Academies Press, 2011.

10. IOM (Institute of Medicine). *Finding what works in Health Care: Standards for Systematic Reviews*. Washington, DC: The National Academies Press, 2011.

11. Matthys J, De Meyere M, van Driel M, De Sutter A. Differences among international pharyngitis guidelines: not just academic. *Ann Fam Med* 2007;5(436-43).

12. Burgers JS, Bailey JV, Klazinga NS, Van der Bij AK, Grol R, Feder G. Inside Guidelines: Comparative analysis of recommendations and evidence in diabetes guidelines from 13 countries. *Diabetes Care* 2002; 25(11):1933-1939.

13. Association AD. Standards of Medical Care in Diabetes - 2008. *Diabetes Care* 2008;31(Supplement 1): S12-S54.

14. Association AD. Standards of Medical Care in Diabetes - 2009. *Diabetes Care* 2009;32 (Supplement 1): S13-S61.

15. Association AD. Standards of Medical Care in Diabetes - 2010. *Diabetes Care* 2010;33 (Supplement 1): S11-S61.

16. Bennett WL, Odelola OA, Wilson LM, Bolen S, Selvaraj S, Robinson KA, et al. Evaluation of guideline recommendations on oral medications for type 2 diabetes mellitus: a systematic review. *Ann Intern Med* 2012;156(1 Pt 1):27-36.

17. Hayward RA, Krumholz HM. Three Reasons to Abandon Low-Density Lipoprotein Targets. *Circ Cardiovasc Qual Outcomes*;5(1):2-5.

18. Tricoci P, Allen JM, Kramer JM, Califf RM, Smith SC, Jr. Scientific Evidence Underlying the ACC/AHA Clinical Practice Guidelines. *JAMA* 2009;301(8):831-841.

19. Schwenk TL. Critique of Clinical Practice Guidelines. *Journal Watch (General)* 2009;March 10.

20. Shaneyfelt T, Centor R. Reassessment of clinical practice guidelines: Go gently into that good night. *JAMA* 2009;301:868.

21. Nissen SE. Can We Trust Cardiovascular Practice Guidelines?: Comment on “Conflicts of Interest in Cardiovascular Clinical Practice Guidelines”. *Arch Intern Med*;171(6):584-585.

22. Mendelson TB, Meltzer M, Campbell EG, Caplan AL, Kirkpatrick JN. Conflicts of Interest in Cardiovascular Clinical Practice Guidelines. *Arch Intern Med*; 171(6): 577-584.

23. Grilli R, Magrini N, Penna A, Mura G, Liberati A. Practice guidelines developed by specialty societies: the need for a critical appraisal. *Lancet* 2000; 355(9198): 103-106.

24. Clinical practice guidelines and conflict of interest. *Can Med Assoc J* 2005;173(11):1297.

25. Taylor R, Giles J. Cash interests taint drug advice. *Nature* 2005;437(7062):1065-6.
26. Gale E. Collateral damage: the conundrum of drug safety. *Diabetologia* 2009;52(10):1975-1982.
27. Choudhry NK, Stelfox HT, Detsky AS. Relationships between authors of clinical practice guidelines and the pharmaceutical industry. *JAMA* 2002; 287(5): 612-7.
28. Wollersheim H. Beyond the evidence of guidelines. *Neth J M* 2009;67:39-40.
29. de Kort S, Burgers J, Willems D. Value judgements that matter to patients remain implicit in oncology guidelines: an observational study. *Neth J M* 2009;67(2): 62-8.
30. Gaminde I, Hermosilla T, Orrego C, Díaz del Campo P. Implicación de los pacientes/cuidadores en la elaboración de las GPC. In: GPC Gdts, editor. *Elaboración de Guías de Práctica Clínica en el Sistema Nacional de Salud. Manual Metodológico*. Madrid: Plan Nacional para el SNS del MSC, 2007.
31. Grol R. Has guideline development gone astray? Yes. *BMJ* 2010;340.
32. 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(1):18-23.
33. <http://www.agreetrust.org/>.
34. Navarro Puerto MA, Ruiz Romero F, Reyes Domínguez A, Gutiérrez Ibarlucea I, Hermosilla Gago T, Ortiz del Río CA, et al. ¿Las guías que nos guían son fiables? Evaluación de las guías de práctica clínica española. *Rev Clin Esp* 2005;205(11):533-40.
35. Hegarty K, Gunn J, Blashki G, Griffith T, Dowell T, Kendrick T. How could depression guidelines be made more relevant and applicable to primary care? *Br J Gen Pract* 2009;59(322-8).
36. Upshur REG. Are all evidence-based practices alike? Problems in the ranking of evidence. *Can Med Assoc J* 2003;169(7):672-673.
37. Schünemann HJ, Best D, Vist G, Oxman AD, Group fTGW. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. *Can Med Assoc J* 2003;169(7):677-680.
38. Hegarty K, Gunn J, Blashki G, Griffiths F, Dowell T, Kendrick T. How could depression guidelines be made more relevant and applicable to primary care?: A quantitative and qualitative review of national guidelines. *The Br J Gen Pract* 2009;59(562):e149.
39. Lugtenberg M, Burgers JS, Clancy C, Westert GP, Schneider EC. Current Guidelines Have Limited Applicability to Patients with Comorbid Conditions: A Systematic Analysis of Evidence-Based Guidelines. *PLoS ONE* 2011;6(10):e25987.
40. Network SIG. SIGN 50: A guideline developers' handbook. Edinburgh: Scottish Intercollegiate Guidelines Network, 2011.
41. Bonfill X, Marzo M. Guías de práctica clínica: tenerlas, que sean de calidad y que salgan del armario. *Med Clin (Barc)* 2003;120(13):496-7.



Servicio Navarro de Salud
Osasunbidea



ISSN

1138-1043

DEPÓSITO LEGAL

NA-1263/1997

INFORMACIÓN Y SUSCRIPCIONES

Servicio Navarro de Salud / Osasunbidea
Plaza de la Paz, s/n
31002 Pamplona
T 848429047
F 848429010

E-mail

farmacia.atprimaria@cfnavarra.es

WEB

www.bit.navarra.es

COMITÉ DE REDACCIÓN

PRESIDENTA

Cristina Ibarrola Guillén

VICEPRESIDENTE

Ignacio Yurss Arruga

VOCALES

Cristina Agudo Pascual

M^a José Ariz Arnedo

Miguel Ángel Imízcoz Zubicaray

Jesús Arteaga Coloma

Idoia Gaminde Inda

M^a Mar Malón Musgo

Rodolfo Montoya Barquet

Javier Gorricho Mendivil

Javier Elizondo Armendáriz

Javier Lafita Tejedor

COORDINADOR

Juan Erviti López