

Sistema GRADE para clasificar nivel de evidencia y grado de las recomendaciones para la elaboración de guías de buena práctica clínica

Luis F. Oñate-Ocaña,* Francisco Javier Ochoa-Carrillo**

Resumen

Actualmente existe gran inconsistencia en cómo los “desarrolladores de guías clínicas” en todo el mundo clasifican la calidad de la evidencia disponible y la fuerza de las recomendaciones, debido a que hay varios sistemas de clasificación. Por esa razón, los usuarios de las guías clínicas se enfrentan a retos profundos en la comprensión de guías clínicas. Importantes investigadores y gran número de ellos están adoptando el sistema GRADE para clasificar la calidad de la evidencia disponible y la fuerza de las recomendaciones cuando preparan manuscritos sobre guías clínicas. El sistema GRADE fue desarrollado por un grupo representativo de expertos en la elaboración de guías clínicas, muestra una clara separación entre el grado de evidencia y la fuerza de las recomendaciones, presenta una evaluación explícita sobre la importancia de los desenlaces finales de las diferentes estrategias de manejo, entre otras características. En esta breve revisión describimos las razones que explican la adopción de esta clasificación.

Palabras clave: Sistema GRADE, nivel de evidencia, grado de recomendación, guía de práctica clínica.

Summary

There is great inconsistency on how “Clinical Guideline Developers” worldwide rate the quality of evidence and strength of recommendations because there are several classification systems in use. Therefore, medical guideline users face profound challenges in understanding articles on Clinical Guidelines. Many researchers are adopting the GRADE system to classify the quality of the available evidence and the strength of recommendations when preparing a manuscript on Clinical Guidelines. The GRADE system has advantages over previous rating systems, including being developed by a representative group of guideline developers, patent separation between quality of evidence and strength of recommendations, explicit evaluation of the importance of outcomes of alternative management strategies, and others. Consequently, in this brief review, we describe the reasons underlying the adoption of GRADE.

Key words: GRADE system, level of evidence, grade of recommendations, clinical practice guidelines.

Introducción

Actualmente es común que se reúna un comité de expertos con la intención de desarrollar guías clínicas para el estudio o tratamiento de una enfermedad, para que estas guías sean publicadas y utilizadas por la comunidad médica como una ayuda para unificar y estandarizar la atención médica, con la intención de lo-

grar mejorar la calidad de los servicios médicos. Sin embargo, cuando se integra un comité de expertos, la heterogeneidad y diversidad de los individuos que lo conforman, sus ideas, puntos de vista y métodos, así como la variabilidad de la información científica disponible sobre el tema, hacen que la elaboración de guías útiles y de alta calidad sea una labor sumamente difícil.^{1,2}

El desarrollo de guías clínicas adecuadas se ha vuelto un motor fundamental que determina la calidad de la práctica clínica. La abundante cantidad de información científica generada en todo el mundo hace imposible que el clínico se mantenga actualizado en todos los campos de la medicina. Por ejemplo, si hoy hiciéramos una consulta electrónica por internet en el portal Pubmed (www.ncbi.nlm.nih.gov/sites/entrez) buscando la palabra “cáncer”, encontraríamos 2 244 584 publicaciones. Es obvia la imposibilidad de revisar la totalidad de la información, y se hace imperativo conocer un método que nos permita seleccionar solamente las publicaciones que tengan mayor relevancia para el problema sobre el que queremos generar recomendaciones. Por esa razón, numerosas sociedades médicas nacionales e internacionales han tomado la responsabilidad de

* Departamento de Gastroenterología.

** División de Cirugía.

Instituto Nacional de Cancerología, México, D. F.

Solicitud de sobretiros:

Luis F. Oñate-Ocaña.

Departamento de Gastroenterología, Instituto Nacional de Cancerología, San Fernando 22, Col. Sección XVI, Del. Tlalpan, 14080 México, D. F.
Fax: (55) 5666 5569.

E-mail: fonate@gmail.com

Recibido para publicación: 14-10-2008

Aceptado para publicación: 27-03-2009

identificar áreas clínicas relevantes y reunir comités de expertos, con el objeto de formular recomendaciones prácticas sólidamente fundamentadas.

Mucho se ha avanzado en cuanto a la metodología para poner de acuerdo a un grupo grande de expertos en disciplinas diversas, aun cuando hace no pocos años se seguía una metodología informal para integrar acuerdos, resolver desacuerdos o establecer cuando hay diferencias importantes.

Por otro lado, existe una gran variedad de sistemas de clasificación del nivel de evidencia que aporta la información científica disponible. Frecuentemente estos sistemas se basan en metodologías poco claras y poco prácticas.

Recientemente, el grupo de estudio GRADE (*Grading of Recommendations, Assessment, Development and Evaluation*, constituido por un grupo internacional y multidisciplinario de metodólogos, desarrolladores de guías clínicas y médicos involucrados en la práctica clínica) ha desarrollado un sistema para la organización de paneles de discusión, jerarquizando los aspectos a considerar e integrando finalmente las recomendaciones más relevantes.³

Lo novedoso de este sistema es que es un método riguroso y transparente para la clasificación de la calidad de la evidencia, para balancear los beneficios y los riesgos de una intervención médica, permitiendo un reconocimiento explícito de los valores y preferencias que representan las recomendaciones (para los pacientes y para los médicos), y si la intervención constituye un uso racional de los recursos desde el punto de vista de la autoridad administrativa (cuadro I). Estos cuatro elementos determinan si una recomendación debe ser catalogada como fuerte o débil.⁴

La calidad de la evidencia debe juzgarse en función de una pregunta de investigación claramente planteada, con una varia-

ble de desenlace también claramente formulada (por ejemplo, frecuencia de respuestas objetivas o supervivencia global). Por otro lado, la calidad de la evidencia debe juzgarse con base en evidencia experimental diseñada para responder la pregunta de investigación específica que se ha planteado *a priori*.

La pregunta de investigación puede corresponder a una de cuatro posibilidades: a un estudio de prueba diagnóstica, a un estudio de riesgo, a un estudio de pronóstico, a un estudio de utilidad terapéutica. La metodología de investigación requerida es diferente según el tipo de pregunta de investigación planteada.⁵

Sobre estas bases, la evidencia puede ser alta, moderada, baja o muy baja (cuadro II).⁶ En las investigaciones que evalúan el efecto del tratamiento, los estudios controlados aleatoriamente brindan mayor evidencia que los estudios observacionales, y probablemente la menor evidencia procede de casos anecdóticos. Además, la calidad inherente al diseño del estudio puede reducirse como resultado de limitaciones metodológicas, conducción inapropiada del estudio, resultados poco precisos o muy variables. Pero la calidad también puede incrementarse cuando se encuentra un gran efecto terapéutico o por el efecto “dosis-respuesta”.⁶

Desde el punto de vista del paciente, las recomendaciones son fuertes cuando virtualmente todos los pacientes adecuadamente informados escogen la misma estrategia de tratamiento. Por otro lado, una recomendación débil implica que las elecciones varían ampliamente dentro de cierta gama de valores y preferencias.^{3,4}

Respecto a los médicos, una recomendación fuerte implica que éstos sugieren que se utilice la estrategia de tratamiento en la mayoría de los pacientes. Una recomendación débil implica que debe discutirse la evidencia con el paciente según sus preferencias y valores.⁷

En cuanto a los administradores, una recomendación fuerte implica que la adherencia a esta estrategia de tratamiento puede

Cuadro I. Factores que deben considerarse para definir la clasificación del nivel de evidencia

Factor	
Calidad de la evidencia	Mientras mayor la calidad de la evidencia, mayor debe ser la fuerza de la recomendación
Ventajas y desventajas	Mientras mayor la diferencia entre las ventajas y desventajas, mayor debe ser la fuerza de la recomendación
Valores y preferencias	Mientras más variación exista entre los valores y preferencias, menor debe ser la fuerza de la recomendación
Costos	Mientras mayores los costos de la intervención, menor debe ser la fuerza de la recomendación

Cuadro II. Definición de la calidad de la evidencia

Calidad de la evidencia	
Alta	Más investigación al respecto no va a modificar nuestra confianza en la estimación del efecto
Moderada	Más investigación puede modificar nuestra confianza en la estimación del efecto
Baja	Muy probablemente más investigación va a tener un efecto importante en nuestra confianza en la estimación del efecto
Muy baja	Cualquier estimación del efecto es sumamente incierta

Cuadro III. Clasificación de la calidad de la evidencia y la fuerza de las recomendaciones

	Código
Calidad de la evidencia	
Alta	A
Moderada	B
Baja	C
Muy baja	D
Fuerza de la Recomendación	
Fuerte a favor de la intervención	1
Débil a favor de la intervención	2
Débil en contra de la intervención	2
Fuerte en contra de la intervención	1

usarse como una medición de calidad de la atención. Si los médicos no usan este tratamiento, deben justificar su omisión. Por otro lado, una recomendación débil implica que los clínicos deben discutir y llegar a un consenso respecto a la estrategia terapéutica. El documento de un consenso es lo que puede ser usado como medida de calidad de la atención.⁸

De tal forma, si la evidencia que sustenta la intervención terapéutica es de alta calidad y se concluye en el consenso que la recomendación debe ser fuerte, hablamos de evidencia y recomendación A1; si la calidad de la evidencia es baja y se concluye que la recomendación debe ser débil, hablamos de evidencia y recomendación C2 (cuadro III). Pero existen otras posibilidades: evidencia de alta calidad que sustenta débilmente el uso de la intervención terapéutica (A2), o incluso evidencia de moderada calidad que sustenta una fuerte recomendación en contra del uso de la intervención terapéutica (B1).

El sistema GRADE ha sido probado en múltiples circunstancias de la práctica clínica cotidiana y ha cubierto un amplio espectro de preguntas de investigación. Es sumamente práctico: ha logrado mantener la simplicidad y además capturar los complejos aspectos que involucran la evaluación de la evidencia científica y la clasificación de las recomendaciones terapéuticas.

Numerosas organizaciones y sociedades médicas han adoptado el sistema GRADE para el proceso de desarrollo de guías clínicas.^{9,10} La revista *British Medical Journal* ha incluido en sus “recomendaciones para los autores” el requerimiento de utilizar el sistema GRADE en cualquier artículo sobre guías clínicas que se pretenda publicar en esa revista. Por estas razones, sugerimos que los grupos de expertos que se reúnen para elaborar guías de tratamiento consideren el uso del sistema GRADE con la intención de integrarnos a este esfuerzo internacional.

Referencias

1. Jaeschke R, Guyatt GH, Dellinger P, Schünemann H, Levy MM, Kunz R, et al. Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. *BMJ* 2008;337:a744.
2. Tricoci P, Allen JM, Kramer JM, Califf RM, Smith SC Jr. Scientific evidence underlying the ACC/AHA clinical practice guidelines. *JAMA* 2009;301:831-841.
3. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-926.
4. Kunz R, Djulbegovic B, Schunemann HJ, Stanulla M, Muti P, Guyatt G. Misconceptions, challenges, uncertainty, and progress in guideline recommendations. *Semin Hematol* 2008;45:167-175.
5. Guyatt G, Rennie D, Meade M, Cook D. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. Second Edition. *JAMA & Archives Journals*: 2008.
6. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ. What is “quality of evidence” and why is it important to clinicians? *BMJ* 2008;336:995-998.
7. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, et al. Going from evidence to recommendations. *BMJ* 2008;336:1049-1051.
8. Guyatt GH, Oxman AD, Kunz R, Jaeschke R, Helfand M, Liberati A, et al. Incorporating considerations of resources use into grading recommendations. *BMJ* 2008;336:1170-1173.
9. van Spronsen E, Ingels KJ, Jansen AH, Graamans K, Fokkens WJ. Evidence-based recommendations regarding the differential diagnosis and assessment of nasal congestion: using the new GRADE system. *Allergy* 2008;63:820-833.
10. Akl EA, van Doormaal FF, Barba M, Kamath G, Kim SY, Kuipers S, et al. Parenteral anticoagulation may prolong the survival of patients with limited small cell lung cancer: a Cochrane systematic review. *J Exp Clin Cancer Res* 2008;27:4.