

Bioestadística y Epidemiología

**NIVELES DE EVIDENCIA Y GRADOS DE RECOMENDACIÓN:
EL SISTEMA GRADE**DANIELA PÉRSICO T.¹ y DAVID TORRES P.²

- 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 ello, hay varios sistemas de clasificación.
- Gran número de investigadores 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, 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.

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.

Desde fines de los años 90, con la consolidación de la Medicina Basada en la Evidencia o en las Pruebas, cualquier procedimiento realizado en medicina, ya sea preventivo, diagnóstico, terapéutico, pronóstico o rehabilitador, tiene que estar definido por su nivel de evidencia científica. En esencia, la MBE pretende aportar más objetividad al arte de la medicina, siendo su objetivo disponer de la mejor información científica disponible -la evidencia- para aplicarla a la práctica clínica.

La evidencia médica, o las recomendaciones en las que están basadas, pueden ser de diferente

calidad. Las fuentes de evidencia van desde la opinión de expertos, pasando por estudios pequeños en laboratorio o reportes de casos, hasta estudios clínicos grandes y bien diseñados que buscan minimizar la probabilidad de sesgos (ensayos clínicos controlados aleatorios doble ciego, no ciego y metanálisis). Debido a que la evidencia de mala calidad puede llevar a recomendaciones que no benefician a los pacientes, es fundamental conocer si una recomendación es fuerte (si podemos confiar en dicha recomendación) o débil (no podemos confiar en ella).

Los niveles de evidencia provienen no sólo de la solidez del diseño del estudio sino también de la solidez de los resultados finales. Además, todas las recomendaciones deben considerar otras características que no pueden ser sopesadas o resumidas fácilmente, como:

- Efectos adversos y morbimortalidad del factor objeto de estudio.
- Nivel de confianza estadística.
- Intervalos de tiempo y duración de la observación.
- Número de participantes en el ensayo.
- Garantías de la calidad del ensayo.

Emitir juicios sobre la calidad de la evidencia y el grado de recomendaciones mediante un enfoque sistemático puede ayudar a prevenir o disminuir los errores, facilitar la valoración crítica de estos juicios y a mejorar la comunicación de esta información¹.

Desde la década de 1970 un número creciente de organizaciones ha empleado diferentes sistemas para clasificar la calidad de la evidencia y el grado de las recomendaciones. Desafortunadamente no hay un acuerdo sobre el método a utilizar, para clasificar la calidad de la evidencia y el grado de recomendaciones. Según el sistema que se use, la

¹ Pasantía Investigación Clínica en Anestesia, Clínica Santa María.

² Anestesiólogo, Clínica Santa María. Instructor, Universidad de los Andes.

misma evidencia y recomendación puede ser clasificada como II-2, B; C+, 1 o evidencia fuerte, fuertemente recomendada. Esto confunde y evita una comunicación efectiva, además de utilizar el tiempo de los médicos en forma ineficiente y poco realista^{2,3}.

Realizar juicios sobre el nivel o la calidad de la evidencia y el grado de las recomendaciones de una guía clínica es un proceso complejo. Los autores deben decidir qué resultado considerar, qué evidencia incluir para cada resultado, cómo evaluar la calidad de esa evidencia y, frecuentemente, considerar si los costos adicionales justifican los beneficios. Los usuarios de guías clínicas necesitan saber cuánta confianza pueden depositar en la evidencia y las recomendaciones que estas entregan. El grupo GRADE (<http://www.gradeworking-group.org>) lleva más de una década perfeccionando un sistema de clasificación de niveles de evidencia y grado de recomendación, donde se describen los factores en que se debería basar nuestra confianza y nos muestra un enfoque sistemático para la toma de decisiones complejas a partir de guías clínicas y otro tipo de recomendaciones². Numerosas publicaciones han aceptado esta forma de clasificación. El British Medical Journal, entre muchas otras, 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.

El sistema GRADE tiene aspectos diferenciales que se pueden resumir en tres: 1) Inicialmente categoriza los desenlaces de interés (ej. muerte) y su importancia relativa; 2) Luego evalúa la calidad de la evidencia en cuatro categorías para cada uno de los desenlaces para obtener una calidad global de evidencia y 3) Finalmente gradúa la fuerza de las recomendaciones en dos únicas categorías (recomendaciones fuertes o débiles).

Es importante partir de una definición común de los dos conceptos fundamentales en la formulación de las recomendaciones: la calidad de la evidencia (conocida también como nivel de evidencia) y la fuerza de la recomendación. El sistema GRADE propone las siguientes definiciones:

- Calidad de la evidencia: indica hasta qué punto nuestra confianza en la estimación de un efecto permite apoyar una recomendación.
- Fuerza de la recomendación: indica hasta qué punto confiar si el poner en práctica la recomendación conllevará más beneficios que riesgos.

El sistema GRADE, al tener dos etapas claramente diferenciadas (Calidad y Fuerza) puede ser utilizado para evaluar únicamente la calidad de la

información disponible sobre un tema (p. ej. al evaluar la calidad en una revisión sistemática) o además para formular recomendaciones en el marco de la elaboración de una guía.

En el caso concreto de las revisiones sistemáticas el sistema GRADE se ha adoptado por parte de la Colaboración Cochrane para evaluar la calidad de la evidencia de sus revisiones sistemáticas.

Las ventajas de esta aproximación son⁴:

- Está desarrollado por un grupo internacional y representativo de desarrolladores de guías clínicas de varias partes del mundo.
- Muestra conjuntamente la calidad de la evidencia y el grado de recomendación. Por ejemplo: mala calidad, mala recomendación.
- Hay una evaluación explícita de la importancia de los resultados de estrategias de manejo alternativas.
- Existen criterios claros para bajar y subir los puntajes de calidad de la evidencia.
- Existe un proceso transparente del paso de la evidencia a la recomendación.
- Hay un reconocimiento explícito de los valores y preferencias de los autores, lo que también pudiera ser una desventaja.
- Finalmente, se presenta una interpretación clara y pragmática de recomendación fuerte o débil para médicos, pacientes y autoridades.

CALIDAD DE LA EVIDENCIA

Indica hasta qué punto se puede confiar en que la estimación del efecto es correcta. En otras palabras, la utilidad de la estimación de la magnitud del efecto de la intervención depende de la confianza que le demos a dicha estimación. Dos conceptos a clarificar: magnitud del efecto y cuánta confianza en esa estimación de la magnitud.

Muchas recomendaciones se toman en función de evidencia de baja calidad, lo que ha producido guías clínicas inadecuadas y recomendaciones que llevan a los médicos a actuar en perjuicio de sus pacientes. Reconocer la calidad de la evidencia puede prevenir estos errores.

Un sistema formal que categorice la calidad de la evidencia ayuda a que sea mejor comprendido por los clínicos. Pese a que la calidad de la evidencia es un continuo y la categorización tiene algún grado de arbitrariedad, los beneficios de la simpleza y la transparencia superan estas limitaciones².

El sistema GRADE clasifica la calidad de la evidencia en 4 niveles: alta, moderada, baja y muy baja⁵. El significado de calidad alta, moderada, baja o muy baja se presenta en la Tabla 1.

Tabla 1. Significado de los niveles de calidad de la evidencia

Calidad de la evidencia	Definición
Alta	Es muy poco probable que nueva investigación modifique nuestra confianza en la estimación del efecto
Moderada	Es probable que nuevos estudios tengan un impacto importante en nuestra confianza en la estimación del efecto, y quizás pueda modificar esta estimación
Baja	Es muy probable que nuevos estudios tengan un impacto importante en nuestra confianza en la estimación del efecto, y es probable que modifique esta estimación
Muy baja	Cualquier estimación del efecto es muy incierta

La evidencia que proviene de estudios clínicos aleatorizados se presupone como de alta calidad, pero este nivel puede disminuir en función de:

- Limitaciones del estudio.
- Inconsistencia de los resultados.
- Evidencia indirecta.
- Imprecisión.
- Sesgo de reporte.

En la propuesta GRADE los estudios se clasifican de tal manera que los estudios controlados aleatorios corresponden de entrada a “calidad alta” y los observacionales a “calidad baja”. Existen además otros cinco factores que pueden justificar dis-

minuir la calidad y otros tres que pueden justificar aumentar la calidad de la evidencia (Tabla 2).

FUERZA DE LA RECOMENDACIÓN

Indica el grado en que se puede estar seguro de que la adhesión a la recomendación producirá beneficio o daño. Cuando los efectos deseados de la intervención claramente superan o no superan a los indeseados, la recomendación es fuerte. Por otro lado, si los efectos de la intervención son más dudosos, o si la calidad de la evidencia es baja, la recomendación es débil².

Tabla 2. Evaluación de la calidad de la evidencia según el tipo de diseño de los estudios

Diseño de estudio	Calidad de la evidencia inicial	Disminuir sí	Aumentar sí	Calidad del conjunto de la evidencia
Ensayos clínicos aleatorizados	Alta	Limitaciones en el diseño o la ejecución Importante (-1) Muy importante (-2)	Fuerza de asociación Fuerte (+1) Muy fuerte (+2)	Alta ⊕⊕⊕⊕
		Inconsistencia Importante (-1) Muy importante (-2)	Gradiente dosis-respuesta Presente (+1)	Moderada ○⊕⊕⊕
		Incertidumbre en que la evidencia sea directa Importante (-1) Muy importante (-2)	Consideración de los posibles factores de confusión que habrían en el efecto (+1) Sugerirían un efecto espurio si no hay efecto (+1)	Baja ○○⊕⊕
Estudios observacionales	Baja	Impresión Importante (-1) Muy importante (-2)		Muy baja ○○○⊕
		Sesgo de publicación Importante (-1) Muy importante (-2)		

Uno de los problemas en que incurren algunos sistemas de clasificación, es que no separan claramente la calidad de la evidencia de la fuerza de recomendación. La evidencia de alta calidad no necesariamente implica una recomendación fuerte, y las recomendaciones fuertes pueden provenir de evidencia de baja calidad. Por ejemplo, los resultados de estudios de alta calidad de uso de anticoagulantes orales para prevención de enfermedad tromboembólica muestran que efectivamente son beneficiosos; sin embargo, aumentan la probabilidad de hemorragia postoperatoria. A pesar de la alta calidad de la evidencia, esta recomendación es débil si la morbilidad causada por hemorragias es similar a la de eventos trombóticos. Por el contrario, hay varios estudios observacionales que muestran los beneficios del dantrolene en hipertermia maligna. La baja calidad de la evidencia, no impide que dantrolene sea fuertemente recomendado.

El sistema GRADE utiliza sólo 2 grados de recomendación: “fuerte” y “débil”. Cuando los efectos esperados de una intervención claramente superan los eventos adversos, (o los efectos adversos claramente superan los efectos esperados), la recomendación es “fuerte”. Por otro lado, cuando existe un balance entre los efectos beneficiosos y los no deseados, o cuando esta diferenciación es menos clara por evidencia de mala calidad, la recomendación es “débil”⁶.

Los factores que afectan la fuerza de la recomendación son:⁷

- Calidad de la evidencia.
- Incerteza en el balance de efectos deseados versus no deseados.

- Variabilidad en los valores y preferencias de los pacientes.
- Poca claridad en la utilización óptima de recursos.

La fuerza de la recomendación se mueve en un continuo que va de fuertemente recomendado a fuertemente no recomendado, con las recomendaciones débiles alrededor de la ausencia de efecto (Figura 1).

CONCLUSIONES

No considerar la calidad de la evidencia, puede llevar a recomendaciones perjudiciales para los pacientes, por lo que las guías clínicas o recomendaciones debieran informar cuál es la calidad de la evidencia que sustenta la intervención, y si dichas recomendaciones son fuertes o débiles.

El sistema GRADE permite evaluar la calidad de la evidencia y la fuerza de recomendación de manera explícita, sistemática, transparente y pragmática, y rápidamente está siendo adoptada por distintas organizaciones.

La evidencia de alta calidad que muestra que los efectos deseados de una intervención son claramente beneficiosos (o claramente no beneficiosos), merecen una recomendación fuerte. Cuando no hay certeza de las ventajas y desventajas de una intervención (por evidencia de baja calidad o falta de claridad en el balance de efectos adversos y beneficios) la recomendación debe ser débil.

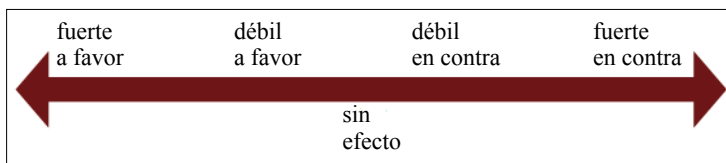


Figura 1. Fuerza de la recomendación.

REFERENCIAS

<ol style="list-style-type: none"> 1. Atkins D, Best D, Briss PA, et al; GRADE Working Group. Grading quality of evidence and strength of recommendations. <i>BMJ</i> 2004; 328: 1490. 2. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. <i>BMJ</i> 2008; 336(7650): 924-926. 3. Atkins D, Eccles M, Flottorp S, et al. Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal 	<ol style="list-style-type: none"> 4. Guyatt GH, Oxman AD, Schünemann HJ, et al. GRADE guidelines: a new series of articles in the <i>Journal of Clinical Epidemiology</i>. <i>J Clin Epidemiol</i> 2011; 64(4): 380-382. 5. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. <i>J Clin Epidemiol</i> 2011; 64(4): 401-406. 6. Andrews J, Guyatt G, Oxman AD, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and 	<ol style="list-style-type: none"> 7. Andrews JC, Schünemann HJ, Oxman AD, et al. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. <i>J Clin Epidemiol</i> 2013; 66(7): 726-735.
--	--	---

Correspondencia a:
Dr. David Torres P.
dtorresp@gmail.com