



ARTÍCULO ESPECIAL

Sistema GRADE: metodología para la realización de recomendaciones para la práctica clínica



Andrea Juliana Sanabria^a, David Rigau^a, Rafael Rotaeche^b, Anna Selva^a,
Mercè Marzo-Castillejo^c y Pablo Alonso-Coello^{a,*}

^a Centro Cochrane Iberoamericano, Instituto de Investigación Biomédica (IIB-Sant Pau), Barcelona, España

^b Grupo Kronikgune sobre gestión del conocimiento, Centro de Salud de Alza, Donostia-San Sebastián, España

^c Unitat de Suport a la Recerca de Costa de Ponent-Institut d'Investigació en Atenció Primària (IDIAP) Jordi Gol, Direcció d'Atenció Primària Costa de Ponent, Institut Català de la Salut, Barcelona, España

Recibido el 19 de diciembre de 2013; aceptado el 20 de diciembre de 2013

Disponible en Internet el 29 de marzo de 2014

PALABRAS CLAVE

Guías de práctica clínica;
Medicina basada en la evidencia;
Adhesión a las guías

Resumen Las guías de práctica clínica proporcionan recomendaciones sobre los beneficios y desventajas de diferentes intervenciones disponibles en la asistencia sanitaria. Su adecuado desarrollo e implementación permitirían reducir la variabilidad en la práctica clínica, así como mejorar su calidad y su seguridad. El sistema GRADE es una herramienta que permite evaluar la calidad de la evidencia y graduar la fuerza de las recomendaciones en el contexto de desarrollo de guías de práctica clínica, revisiones sistemáticas o evaluación de tecnologías sanitarias. El objetivo de este artículo es describir las principales características del sistema GRADE a través de ejemplos relevantes en el contexto de la atención primaria.

© 2013 Elsevier España, S.L.U. Todos los derechos reservados.

KEYWORDS

Practice guidelines;
Evidence-based medicine;
Guideline adherence

GRADE: Methodology for formulating and grading recommendations in clinical practice

Abstract Clinical practice guidelines (CPG) provide recommendations on the benefits and harms of different healthcare interventions. Proper CPG development and implementation can potentially reduce variability in clinical practice while improving its quality and safety. The GRADE system is used to assess the quality of evidence and to grade the strength of recommendations in the context of the development of CPGs, systematic reviews or health technology assessments. The aim of this article is to describe the main characteristics of the GRADE system through relevant examples in the context of primary care.

© 2013 Elsevier España, S.L.U. All rights reserved.

Introducción

Las guías de práctica clínica (GPC) se definen como el conjunto de recomendaciones basadas en una revisión

* Autor para correspondencia.

Correo electrónico: palonso@santpau.cat (P. Alonso-Coello).

sistemática de la evidencia y en la evaluación de los riesgos y beneficios de las diferentes alternativas, con el objetivo de optimizar la atención sanitaria a los pacientes¹. Su éxito depende del rigor en su elaboración, así como de las estrategias de diseminación, implantación y actualización.

En los últimos años, con el objetivo de obtener GPC de mayor calidad, diversas instituciones e iniciativas han aportado progresivamente la metodología de su elaboración²⁻⁴. GRADE surgió como una iniciativa internacional con el propósito de optimizar la evaluación de la calidad de la evidencia y la graduación de la fuerza de las recomendaciones, superando las limitaciones de los sistemas de clasificación previos y proponiendo un sistema nuevo que mejora la trazabilidad y la transparencia del proceso (<http://www.gradeworkinggroup.org/>). Actualmente, más de 70 instituciones como la Organización Mundial de Salud, la Colaboración Cochrane o el *National Institute of Clinical Excellence* (NICE) se adhieren o utilizan GRADE en la realización de sus recomendaciones. En nuestro entorno, el Programa de GPC del Sistema Nacional de Salud, coordinado por GuíaSalud (<http://portal.guiasalud.es/web/guest/gpc-sns>), o la Sociedad Española de Medicina de Familia y Comunitaria también comienzan a introducirlo en la elaboración de recomendaciones. El objetivo del presente artículo es describir las características básicas del sistema GRADE utilizando ejemplos relevantes para atención primaria.

Planteamiento de la pregunta clínica y clasificación de los desenlaces de interés

Las GPC están estructuradas para responder a preguntas clínicas. Una pregunta clínica debe incorporar una población, una intervención a evaluar, un comparador, así como unos desenlaces (resultados) de interés (p. ej., mortalidad o efectos indeseables)⁵. Esta estrategia de formulación de preguntas se conoce con el acrónimo PICO (paciente/intervención/comparación/desenlace [*outcome*, en inglés]), y a pesar de no ser exclusiva del sistema GRADE, es una parte fundamental para aplicarlo.

Si tomamos como ejemplo la GPC para el tratamiento de pacientes con enfermedad pulmonar obstructiva crónica (EPOC) del Programa de GPC del Sistema Nacional de Salud, una de las preguntas fue: la utilización del mucolítico carbocisteína ¿está justificada en los pacientes con EPOC? Esta pregunta puede estructurarse de la siguiente forma: a) paciente/problema: pacientes con EPOC en fase estable; b) intervención: mucolítico carbocisteína; c) comparación: placebo, y d) desenlaces: número de exacerbaciones y hospitalizaciones, calidad de vida, síntomas respiratorios, tolerancia al ejercicio, mortalidad por cualquier causa y eventos adversos⁶.

En la selección de los desenlaces de interés a evaluar, el grupo elaborador debe considerar aquellos que sean más importantes para los pacientes, así como para otros actores relacionados con la intervención evaluada (p. ej., cuidadores)⁵. Serán estos desenlaces los que deban tenerse en cuenta a la hora de evaluar la calidad de la evidencia y de graduar la fuerza de las recomendaciones. La [figura 1](#)

muestra cómo se definieron estos desenlaces en la pregunta sobre el uso de carbocisteína en la GPC de la EPOC.

Identificación de la literatura científica disponible

Cada pregunta de una GPC debe basarse en una revisión sistemática (RS) de la literatura⁷. Las RS permiten presentar de forma resumida la evidencia disponible sobre las intervenciones que se están evaluando. Los resultados de la RS sirven para informar al grupo elaborador, facilitando el juicio entre los beneficios y riesgos de las intervenciones y apoyando la elaboración de recomendaciones en la GPC⁷. La RS puede realizarse por el grupo elaborador, usando los resultados de los estudios individuales disponibles (ensayos clínicos aleatorizados [ECA], estudios observacionales, etc.), o pueden identificarse a través de una búsqueda de la literatura⁷. Independientemente del proceso seguido, es importante que la búsqueda de estudios relevantes cuente con una descripción clara de los criterios para su identificación y selección. Estos criterios deben ser acordes con los componentes de la pregunta planteada.

Evaluación de la calidad de la literatura científica

La evaluación de la calidad de la evidencia nos permite conocer el grado de confianza que tenemos en los resultados de la literatura. En otras palabras, hasta qué punto confiamos o no en los resultados obtenidos en los estudios localizados. Algo muy característico de GRADE es que la evaluación de la calidad de la evidencia se realiza inicialmente para cada uno de los desenlaces de interés. Esto quiere decir que, una vez encontremos los estudios relevantes (p. ej., RS de ECA sobre mucolíticos en EPOC), identificaremos el efecto que tiene la intervención en los desenlaces de interés que hemos elegido (p. ej., mortalidad, frecuencia de exacerbaciones, etc.). Evaluaremos entonces la calidad de la evidencia para cada desenlace de forma individual para, posteriormente, emitir un juicio acerca de la calidad global del conjunto de desenlaces. Esta calidad global se determina tomando la calidad de la evidencia más baja de los desenlaces de interés que se han considerado como claves para la elaboración de las recomendaciones, como explicaremos más adelante. GRADE propone clasificar la calidad de la evidencia en 4 categorías: calidad alta, moderada, baja y muy baja⁸. Estas categorías reflejan un gradiente de confianza en los resultados, así como en que estudios posteriores puedan modificar los resultados disponibles.

De entrada los ECA se consideran como calidad alta y los estudios observacionales como calidad baja. En esta valoración se tienen en cuenta los potenciales limitaciones de los estudios: limitaciones en el diseño y ejecución (riesgo de sesgo), resultados inconsistentes, resultados imprecisos, ausencia de evidencia directa y sesgo de publicación. Tanto para los ECA como para los estudios observacionales se han determinado una serie de factores que pueden aumentar o disminuir la confianza en la estimación del efecto observado ([tabla 1](#)).

Puntuación	Tipo de desenlace	Desenlace
9	Desenlaces claves para la toma de decisiones	<ul style="list-style-type: none"> • Eventos adversos • Calidad de vida • Frecuencia hospitalizaciones • Frecuencia exacerbaciones
8		
7		
6	Desenlaces importantes , pero no claves para la toma de decisiones	<ul style="list-style-type: none"> • Tolerancia al ejercicio • Mortalidad por cualquier causa • Síntomas respiratorios
5		
4		
3	Desenlaces no importantes	<ul style="list-style-type: none"> • Función pulmonar (ej. VEMS1, CVF)
2		
1		

Figura 1 Clasificación de los desenlaces de interés. Uso de carbocisteína en los pacientes con EPOC. El sistema GRADE clasifica los desenlaces mediante una escala de 9 puntos en: claves (7-9), importantes pero no claves (4-6), y poco importantes (1-3)⁵. El grupo de autores de la GPC de EPOC valoraron de forma individual la importancia de los 8 desenlaces y por consenso consideraron un total de 7, entre importantes y claves, a incluir en el proceso de realización de las recomendaciones en esta pregunta⁶. CVF: capacidad vital forzada; VEMS1: volumen espirado máximo en el primer segundo de espiración forzada.

Factores que pueden disminuir la calidad de la evidencia

Limitaciones en el diseño y ejecución (riesgo de sesgo)

En los ECA se consideran fundamentalmente los siguientes factores: método de generación de la secuencia de aleatorización inapropiado, ausencia de ocultamiento de la secuencia de aleatorización, enmascaramiento inadecuado de las intervenciones, pérdidas de seguimiento importantes, ausencia de análisis por intención de tratar y/o descripción selectiva de desenlaces de interés⁹. En el caso de los estudios observacionales, los factores que limitan la confianza (calidad) en los resultados son: criterios inapropiados de selección de la población, medidas de exposición o desenlace inapropiadas, control inadecuado de los factores de confusión o seguimiento incompleto de los pacientes⁹.

Inconsistencia en los resultados

Si existen diferencias entre los resultados de los distintos estudios que han evaluado un desenlace (resultados heterogéneos), y estas diferencias persisten tras haber explorado las razones que podrían explicarlas, se considera que se debe disminuir la calidad de la evidencia. Habitualmente esta variabilidad es debida a que los estudios presentan diferencias en las poblaciones o intervenciones, en la evaluación de los desenlaces o a que presentan diferentes riesgos de

sesgo. En estas circunstancias pueden existir diferencias reales entre las estimaciones del efecto y, por tanto, nuestra confianza disminuye¹⁰.

La GPC para el tratamiento de pacientes con EPOC evaluó la eficacia de la rehabilitación respiratoria en el contexto de una exacerbación. Uno de los desenlaces fue la tolerancia al ejercicio evaluada con la prueba de andar durante 6 min. Una RS que analizó 6 estudios para esta prueba mostró unos resultados muy variables entre los estudios, con intervalos de confianza que no se solapaban, y se objetivó heterogeneidad estadística (fig. 2)¹¹. En situaciones como esta, como en el caso de esta guía, nuestra confianza en los resultados disminuye⁶.

Ausencia de evidencia directa

Es frecuente que ciertas decisiones se basen en estudios con poblaciones, intervenciones o desenlaces distintos a los de nuestra pregunta de interés. También es habitual que no existan estudios que hayan comparado de forma directa 2 o más alternativas terapéuticas¹². Por ejemplo, en la evaluación del efecto de diferentes bisfosfonatos en la prevención de fracturas vertebrales en mujeres posmenopáusicas con osteoporosis, una RS comparó la eficacia del ácido zolendronico respecto a otros fármacos (alendronato, ibandronato y risedronato)¹³. Al no encontrarse estudios que los evaluaran de forma directa, los autores realizaron una comparación

Tabla 1 Evaluación de la calidad de la evidencia según el tipo de diseño de los estudios

Diseño de estudio	Calidad de la evidencia inicial	En ensayos clínicos, disminuir si	En estudios observacionales, aumentar solo si	Calidad del conjunto de la evidencia
Ensayos clínicos aleatorizados	Alta	Limitaciones en el diseño o la ejecución – Importantes – Muy importantes Inconsistencia – Importante – Muy importante Incertidumbre en que la evidencia sea directa – Importante – Muy importante Imprecisión –Importante –Muy importante	Fuerza de asociación – Fuerte – Muy fuerte Gradiente dosis-respuesta – Presente Consideración de los posibles factores de confusión que: – Habrían reducido el efecto – Sugerirían un efecto espurio si no hay efecto	Alta Moderada Baja Muy baja
Estudios observacionales	Baja			

Adaptada de: Balshem et al.⁸.

indirecta a través de ensayos clínicos que comparaban los diferentes fármacos respecto a placebo. En estas comparaciones indirectas el ácido zolendrónico se mostró más eficaz que el resto de bisfosfonatos. Si utilizáramos los resultados de esta RS, al no tratarse de una comparación directa, disminuiría la confianza sobre la superioridad observada del ácido zolendrónico en la prevención de fracturas.

Otro ejemplo lo encontramos en una reciente GPC sobre el tratamiento antitrombótico y prevención de la trombosis, la cual evaluó el uso de las medias de compresión graduada como profilaxis inicial para trombosis pulmonar respecto a su no uso, en pacientes sometidos a cirugía mayor ortopédica¹⁴. Sin embargo, a pesar de utilizarse con frecuencia en combinación con otros tratamientos trombotoprolácticos, los autores identificaron solo una RS en la que la mayoría de los estudios incluidos fueron realizados en pacientes sometidos a cirugía no ortopédica. El hecho de que la información provenga de una población diferente a la pregunta de interés hace que disminuya, la confianza en sus resultados¹⁴.

Imprecisión

Para considerar que el efecto de una intervención es impreciso se debe valorar, según las circunstancias, el intervalo de confianza del estimador y el número de eventos o el número de sujetos incluidos. Por ejemplo, en la GPC de la EPOC, una RS mostró que los mucolíticos reducían las hospitalizaciones respecto a placebo (fig. 3). Si nos fijamos en el intervalo de confianza, esta reducción, en términos relativos, puede ser tan importante como un 51% (sin duda recomendaríamos los mucolíticos) o llegar a aumentar 1% las hospitalizaciones (probablemente no recomendaríamos los mucolíticos. Si lo valoramos en términos absolutos, por cada 1.000 pacientes el tratamiento con mucolíticos podría reducir 131 hospitalizaciones o aumentar 3 hospitalizaciones. Si una decisión clínica o recomendación puede cambiar dependiendo de si se tiene en cuenta un extremo u otro del intervalo de confianza, nuestra confianza en la estimación del efecto debería disminuir^{6,15}.

Asimismo, incluso con un intervalo de confianza preciso, si el número de episodios o el número de sujetos evaluados son escasos, debería considerarse la necesidad de disminuir la calidad¹⁶. En la RS del ejemplo anterior se evaluó la mortalidad (fig. 4). En los 6 ECA incluidos hubo 26 muertes: 11 en pacientes que recibieron mucolítico y 15 en pacientes que recibieron placebo. Cualquier resultado con este escaso número de episodios es muy incierto, y nuestra confianza en los resultados disminuye¹⁵.

Sesgo de publicación

Cada vez se conoce mejor que hay situaciones en las que podrían existir estudios, principalmente con resultados negativos, que no se han publicado y que por tanto el efecto observado con los estudios disponibles podría estar sobreestimándose. Este potencial sesgo debe sospecharse ante situaciones en las que se dispone de un conjunto de ensayos de pequeño tamaño, positivos y financiados por la industria. En estos casos se podría llegar a reducir la confianza en la estimación de un efecto¹⁷. Para detectar el sesgo de publicación también existen pruebas estadísticas (test de Egger) o gráficas (*funnel plot*) que podrían proporcionar información complementaria de utilidad.

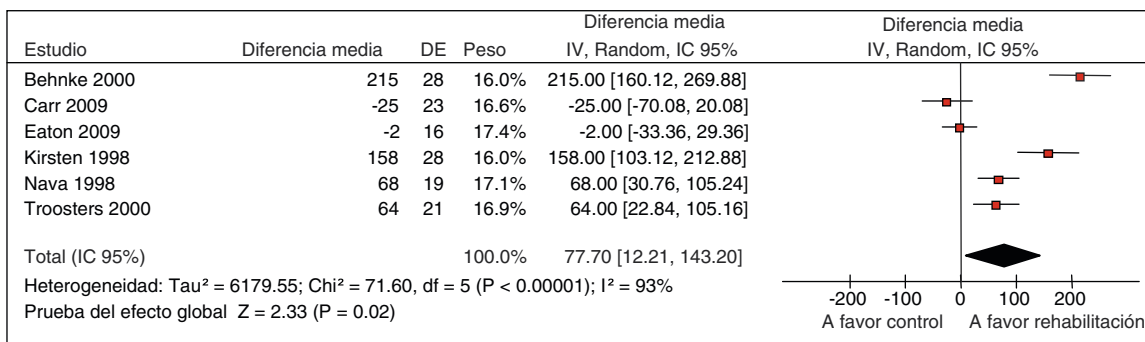


Figura 2 Metanálisis de la rehabilitación versus cuidado habitual en las exacerbaciones de la EPOC. Prueba de tolerancia al ejercicio a los 6 min. De Puhan et al.¹¹.

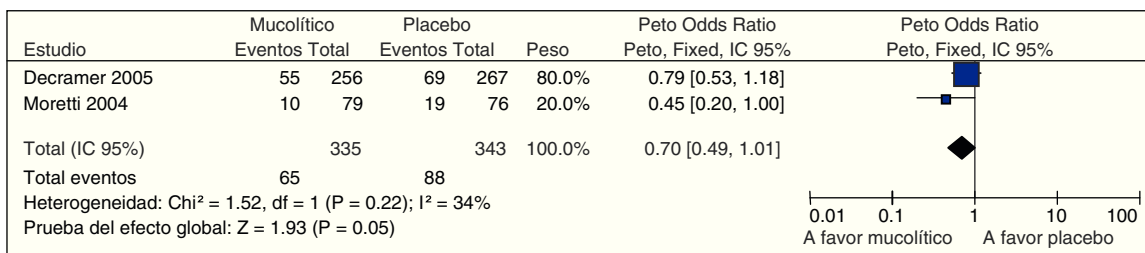


Figura 3 Metanálisis de mucolítico versus placebo en el tratamiento de la EPOC. Reducción de hospitalizaciones. De Poole et al.¹⁵.

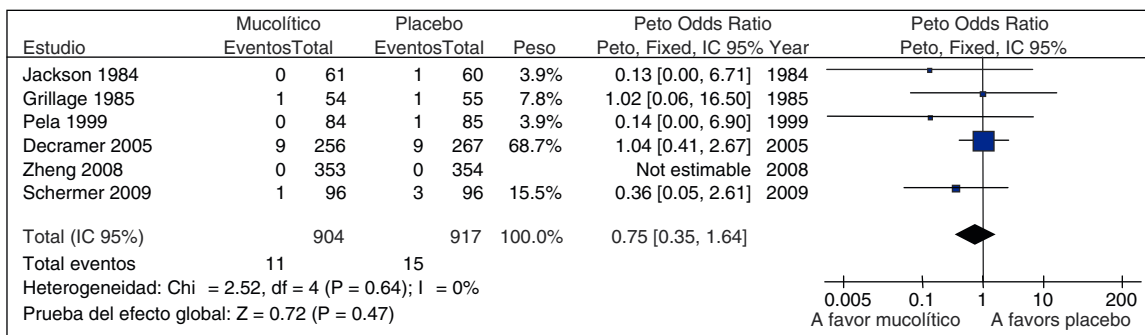


Figura 4 Metanálisis de mucolítico versus placebo en el tratamiento de la EPOC. Mortalidad. De Poole et al.¹⁵.

Factores que pueden aumentar la calidad de la evidencia

A pesar de que los estudios observacionales son considerados por GRADE como de calidad baja para dirimir los efectos de una intervención, hay situaciones que pueden justificar una mayor confianza en sus resultados. Son situaciones poco comunes y solo tiene sentido considerarlas si no existen limitaciones en el diseño y la ejecución de los estudios (riesgo de sesgo)¹⁸. Estas son: a) la fuerza de la asociación: cuando el efecto observado de una intervención muestra una asociación fuerte (riesgo relativo u odds ratio > 2 o $< 0,5$) o muy fuerte (riesgo relativo u odds ratio > 5 o $< 0,2$), aumenta nuestra confianza, ya sea un efecto protector o perjudicial; b) el gradiente dosis-respuesta: la existencia de un gradiente claro entre la intensidad de la exposición o dosis

y el efecto observado nos aporta una mayor certidumbre sobre una potencial relación causa-efecto, y c) la consideración de otros potenciales factores de confusión que no se han tenido en cuenta en el estudio y que, de presentarse, reforzarían sus resultados.

Tabla de síntesis de los resultados

El sistema GRADE hace especial énfasis en la presentación explícita y transparente de todo el proceso. Una de las herramientas disponibles es la tabla de síntesis de los resultados (*summary of findings*). Esta tabla permite presentar para las comparaciones de interés el resumen de los estudios y sus resultados, así como la confianza en los mismos (calidad de la evidencia) respecto a los desenlaces más importantes¹⁹.

Elaboración de la recomendación

Dirección y fuerza de las recomendaciones

Las recomendaciones pueden ser clasificadas según su dirección y su fuerza. Según GRADE, su dirección puede ser a favor o en contra de la intervención evaluada y, según su fuerza, fuertes o débiles (también denominadas condicionales).

GRADE tiene en cuenta 4 criterios para determinar la dirección y la fuerza de las recomendaciones: a) el balance entre los beneficios y riesgos e inconvenientes; b) la confianza en la magnitud de las estimaciones del efecto de la intervención en los desenlaces importantes (calidad global de la evidencia); c) los valores y preferencias de los pacientes, y d) el uso de recursos y costes²⁰.

Teniendo en cuenta la fuerza y la dirección, las recomendaciones pueden ser *fuertes a favor*, *fuertes en contra*, *débiles a favor* o *débiles en contra* de una intervención²⁰. Por otra parte, una calidad baja no implica necesariamente una recomendación débil o la situación inversa, una calidad alta tampoco implica una recomendación fuerte. Se deben tener en cuenta otros aspectos para determinar la fuerza o la dirección de la recomendación, como son los valores y preferencias de los pacientes o el uso de recursos y costes, entre otros. Respecto a la redacción de las recomendaciones, GRADE propone, en el caso de las recomendaciones fuertes, utilizar términos como «se recomienda», y en el caso de las débiles, «se sugiere»²⁰.

Balance entre beneficios y riesgos

Para realizar el balance entre beneficios y riesgos deben considerarse las diferencias que hay entre los estimadores del efecto de los desenlaces considerados. Si existe una gran diferencia entre los beneficios (p. ej., disminución de ingresos hospitalarios y/o disminución de exacerbaciones) y los riesgos (p. ej., efectos adversos leves), es más probable realizar una recomendación fuerte (a favor o en contra). Si la diferencia es pequeña, será más probable hacer una recomendación débil²⁰.

Tomemos nuevamente como ejemplo la GPC de EPOC⁶, y esta vez la pregunta versa sobre el uso de corticoides por vía oral para el manejo de las exacerbaciones moderadas. En esta pregunta, el grupo elaborador consideró que había un beneficio neto a favor de la intervención, es decir, las consecuencias deseables (mejora del fracaso terapéutico y la disnea) superaban a las indeseables (efectos adversos como la ganancia de peso y/o insomnio) (tabla 2).

Calidad global de la evidencia

Después de realizar la valoración de la calidad de la evidencia de forma individual por desenlace (ver apartado «Evaluación de la calidad de la literatura científica»), GRADE realiza una valoración global de la misma. Al igual que en la valoración individual, esta valoración global expresa la confianza general que tenemos en los desenlaces de interés y corresponde al menor nivel de confianza de los desenlaces claves²¹. Así, si para responder una pregunta tenemos 3 desenlaces claves, 2 de ellos con una calidad alta y otro

Tabla 2 Corticoides sistémicos en las exacerbaciones moderadas de la EPOC⁶

Recomendación fuerte a favor

En pacientes con EPOC no hospitalizados se sugiere administrar prednisona oral durante 9 o 10 días para tratar las exacerbaciones moderadas

Balance entre beneficios y riesgos

Los autores consideraron que puede existir un beneficio neto asociado a la intervención. La evidencia disponible muestra una reducción del fracaso terapéutico y de la disnea con la administración oral de prednisona durante las exacerbaciones de intensidad moderada de los pacientes con EPOC, sin llegar a tener un impacto en desenlaces claves como las recurrencias, la frecuencia de hospitalizaciones, la calidad de vida o la mortalidad atribuible a la exacerbación. La prednisona se asoció con la aparición de eventos adversos no graves (aumento del apetito, ganancia de peso e insomnio). Antes de administrar este tratamiento es necesario considerar el balance beneficio-riesgo para cada paciente, en función de su susceptibilidad a los efectos adversos asociados a los corticoides

Calidad global de la evidencia

La calidad de la evidencia fue moderada para los desenlaces de fracaso terapéutico, necesidad de hospitalización, disnea, calidad de vida y eventos adversos. La calidad de la evidencia fue baja para mortalidad atribuible a la exacerbación de la EPOC. La confianza global en las medidas del efecto fue valorada como moderada

Valores y preferencias

A pesar de no haber identificado estudios que consideraran los valores y preferencias de los pacientes respecto a esta intervención, el panel de expertos consideró que al ser un tratamiento que puede ser administrado de forma ambulatoria, podría ser bien aceptado por los pacientes y profesionales, incluso después de tener en cuenta los efectos adversos asociados a la intervención

Uso de recursos y costes

No se identificaron estudios económicos del tratamiento con corticoides sistémicos de la exacerbación moderada de los pacientes con EPOC. Los autores consideraron que no se ha podido demostrar que la intervención sea coste-efectiva para este tipo de pacientes

con calidad moderada, la calidad global de la evidencia será valorada como moderada.

Típicamente, una recomendación fuerte está sustentada con una confianza alta o moderada en las estimaciones del efecto para los desenlaces claves y una recomendación débil, en una confianza baja o muy baja. Sin embargo, pueden existir situaciones poco comunes en las cuales se pueden realizar recomendaciones fuertes a pesar de contar con baja calidad de la evidencia²⁰. Un ejemplo sería el uso de escáner total para el cribado del cáncer. En esta situación se dispondría de evidencia que sugiere un beneficio para la detección temprana del cáncer pero de calidad baja, y a su vez se tendría evidencia de un potencial daño de calidad alta y un

coste muy elevado. En este caso se haría una recomendación fuerte en contra de su uso²².

Valores y preferencias

Los valores y preferencias de los pacientes son un concepto que abarca sus perspectivas, creencias, expectativas y metas para su salud y vida en general²⁰. Es el resultado del proceso por el que valoran los potenciales beneficios, daños, costes e inconvenientes asociados a las diferentes opciones terapéuticas o de manejo de su enfermedad²⁰.

En la elaboración de GPC, si bien los profesionales implicados aportan los conocimientos (clínicos o metodológicos) y la evidencia científica disponible, es la perspectiva de los pacientes la que es relevante al valorar los desenlaces de interés. No obstante, en la mayoría de las ocasiones los grupos elaboradores al hacer recomendaciones no hacen explícitos los valores que están asignando a los pros y contras de las intervenciones. Sin embargo, a pesar de no ser conscientes de este proceso, están realizando juicios tomando en la mayoría de los casos sus valores y preferencias y no la de los pacientes.

Desafortunadamente, aunque los grupos elaboradores de GPC quieran incorporar los valores y preferencias de los pacientes, a menudo esta información no se encuentra en la literatura científica. En estos casos el panel puede usar su experiencia en la toma de decisiones compartidas del problema de salud de interés y explicitar qué valores han tenido en cuenta y utilizar su estimación aproximada de manera explícita²⁰. Otra posibilidad, ante la ausencia de evidencia, es realizar investigación con pacientes mediante técnicas cualitativas o cuantitativas para explorar este aspecto²³.

Continuando con el ejemplo del uso de corticoides vía oral para el tratamiento de las exacerbaciones moderadas del EPOC (tabla 2), el grupo elaborador no encontró literatura sobre valores y preferencias de los pacientes y consideró que, al ser un tratamiento de administración extrahospitalaria, tendría buena aceptabilidad entre los pacientes, a pesar de la presencia de efectos adversos asociados al tratamiento⁶. En otras palabras, el panel asignó una importancia mayor a la mejora de la disnea y el fracaso terapéutico que a los inconvenientes de la medicación (beneficio neto de la intervención).

Uso de recursos y costes

Como cualquier otro desenlace importante, el uso de los recursos (p. ej., personal sanitario, insumos, quirófanos, etc.) y los costes (p. ej., precio de los medicamentos, precio hora/médico) son a menudo relevantes a la hora de realizar recomendaciones. La intervención evaluada y sus comparaciones pueden diferir en la cantidad de recursos y costes necesarios. La integración de estos aspectos a las recomendaciones puede variar dependiendo del país, del sistema sanitario y de los pacientes a los cuales se aplican las intervenciones. El grupo elaborador debe evaluar cuáles de las preguntas de la GPC pueden incluir intervenciones en las cuales el uso de recursos y costes son claves para realizar la recomendación. Una vez hecha esta evaluación es necesario revisar los estudios existentes relevantes y, en caso de no

encontrarlos, plantearse la necesidad de desarrollar análisis económicos propios²⁴.

Otros aspectos

Según la perspectiva elegida para elaborar la GPC y la pregunta, el grupo elaborador puede incorporar otros aspectos en la realización de las recomendaciones como la equidad (impactos en las inequidades en salud), la aceptabilidad para los agentes implicados o la factibilidad de implementar la recomendación.

Integración de los factores

Para formular la recomendación y graduar la fuerza de una recomendación es necesario integrar los factores mencionados anteriormente. Por ejemplo, en la GPC de la EPOC los autores hicieron una recomendación fuerte a favor de utilizar prednisona por vía oral en pacientes no hospitalizados con exacerbaciones moderadas. Después de evaluar estos cuatro factores, los autores consideraron que podría existir un beneficio neto asociado a su uso (tabla 2).

Conclusión

El sistema GRADE es un sistema para la evaluación de la calidad de la evidencia y la formulación de recomendaciones que busca mejorar las limitaciones de sistemas previos. GRADE proporciona un marco para estructurar el proceso de formulación de recomendaciones de manera explícita, evaluando la calidad de la evidencia por desenlaces de interés e incorporando otros factores relevantes en la elaboración de recomendaciones como los valores y preferencias de los pacientes y el uso de recursos y costes.

Bibliografía

1. Consensus report, Institute of Medicine. Clinical practice guidelines we can trust. March 23, 2011 [consultado 16 Oct 2013]. Disponible en: <http://www.iom.edu/Reports/2011/Clinical-Practice-Guidelines-We-Can-Trust.aspx>
2. Scottish Intercollegiate Guidelines Network. SIGN 50: A guideline developer's handbook. Edinburgh: Scottish Intercollegiate Guidelines Network; 2011.
3. National Institute for Health and Clinical Excellence. The Guidelines Manual. London: National Institute for Health and Clinical Excellence; 2009.
4. Brouwers M, Kho M, Browman G, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: Advancing guideline development, reporting and evaluation in health care. *CMAJ*. 2010;182:E839-42.
5. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol*. 2011;64:395-400.
6. Grupo de trabajo de la Guía de Práctica Clínica para el Tratamiento de Pacientes con Enfermedad Pulmonar Obstructiva Crónica (EPOC). Guía de Práctica Clínica para el Tratamiento de Pacientes con Enfermedad Pulmonar Obstructiva Crónica (EPOC). Plan de Calidad para el Sistema Nacional de Salud del Ministerio de Sanidad, Servicios Sociales e Igualdad. Unidad de Evaluación de Tecnologías Sanitarias de la Agencia Laín Entralgo; 2012. Guías de Práctica Clínica en el SNS: UETS n.º 2011/6.

7. Brozek JL, Akl EA, Alonso-Coello P, Lang D, Jaeschke R, Williams JW, et al. Grading the quality of evidence and strength of recommendations in clinical practice guidelines. Part 1 of 3. An overview of the GRADE approach and grading quality of evidence about interventions. *Allergy*. 2009;64:667–9.
8. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011;64:401–6.
9. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). *J Clin Epidemiol*. 2011;64:407–15.
10. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7. Rating the quality of evidence—inconsistency. *J Clin Epidemiol*. 2011;64:1294–302.
11. Puhan MA, Gimeno-Santos E, Scharplatz M, Troosters T, Walters EH, Steurer J. Pulmonary rehabilitation following exacerbations of chronic obstructive pulmonary disease. *Cochrane Database Syst Rev*. 2011:CD005305.
12. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al., The GRADE Working Group. GRADE guidelines: 8. Rating the quality of evidence—indirectness. *J Clin Epidemiol*. 2011;64:1303–10.
13. Jansen JP, Bergman GJ, Huels J, Olson M. Prevention of vertebral fractures in osteoporosis: Mixed treatment comparison of bisphosphonate therapies. *Curr Med Res Opin*. 2009;25:1861–8.
14. Falck-Ytter Y, Francis CW, Johanson NA, Curley C, Dahl OE, Schulman S, et al. Prevention of VTE in orthopedic surgery patients: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed.: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest*. 2012;141 2 Suppl:e278S–325S.
15. Poole P, Black PN, Cates CJ. Mucolytic agents for chronic bronchitis or chronic obstructive pulmonary disease. *Cochrane Database of Systematic Reviews*. 2012:CD001287.
16. Guyatt G, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines: 6. Rating the quality of evidence—imprecision. *J Clin Epidemiol*. 2011;64:1283–93.
17. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, et al. GRADE guidelines: 5. Rating the quality of evidence—publication bias. *J Clin Epidemiol*. 2011;64:1277–82.
18. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, et al., The GRADE Working Group. GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol*. 2011;64:1311–6.
19. Guyatt GH, Oxman AD, Santesso N, Helfand M, Vist G, Kunz R, et al. GRADE guidelines: 12. Preparing Summary of Findings tables—binary outcomes. *J Clin Epidemiol*. 2013;66:158–72.
20. Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, et al. GRADE guidelines: 14. Going from evidence to recommendations: The significance and presentation of recommendations. *J Clin Epidemiol*. 2013;66:719–25.
21. Guyatt G, Oxman AD, Sultan S, Brozek J, Glasziou P, Alonso-Coello P, et al. GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. *J Clin Epidemiol*. 2013;66:151–7.
22. Andrews JC, Schünemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, et al. GRADE guidelines: 15. Going from evidence to recommendation—determinants of a recommendation’s direction and strength. *J Clin Epidemiol*. 2013;66:726–35.
23. Grupo de trabajo de implicación de pacientes en el desarrollo de GPC. Implicación de Pacientes en el Desarrollo de Guías de Práctica Clínica: Manual Metodológico. Ministerio de Sanidad, Servicios Sociales e Igualdad. Instituto Aragonés de Ciencias de la Salud-IACS. Guías de Práctica Clínica en el SNS: IACS n.º 2010/0.
24. Brunetti M, Shemilt I, Pregno S, Vale L, Oxman AD, Lord J, et al. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. *J Clin Epidemiol*. 2013;66:140–50.