

Anexos

Anexo 1. Evaluación de la calidad de la evidencia y la graduación de las recomendaciones mediante el sistema GRADE

Cuando se formulan recomendaciones, el grupo elaborador de la guía debe plantearse hasta qué punto se puede confiar en que llevar a cabo una recomendación conllevará más beneficios que daños. Esta decisión no es sencilla y está influida por múltiples factores que convierten esta etapa en una de las más complejas de la elaboración de una guía.

Los sistemas de formulación de recomendaciones se comenzaron a proponer hace más de dos décadas. Estos sistemas diferenciaban ya inicialmente el nivel de evidencia científica (la adecuación de los diferentes diseños de estudios para contestar a los diferentes tipos de preguntas) de la fuerza de las recomendaciones. Desde entonces los diferentes sistemas han ido evolucionando e incorporando otros aspectos más allá del diseño de los estudios que hay que tener en cuenta a la hora de formular recomendaciones.

Para clasificar la calidad de la evidencia y graduar la fuerza de las recomendaciones se han seguido las directrices del grupo de trabajo GRADE. Este grupo de trabajo intenta establecer un método de elaboración de recomendaciones explícito y transparente, de fácil manejo por los equipos elaboradores de GPC, con el objetivo de superar los inconvenientes de otros sistemas de elaboración de recomendaciones¹².

Las principales etapas que conforman el sistema GRADE (*Grading of Recommendations of Assessment Development and Evaluation*) se describen a continuación:

A) Clasificación de la importancia relativa de las variables de resultado

En esta etapa el sistema GRADE recomienda que en la etapa inicial de la formulación de las preguntas clínicas el grupo elaborador establezca de forma explícita las variables de resultado de interés para las preguntas y clasifique su importancia relativa. Se recomienda clasificar su importancia mediante la siguiente escala de nueve puntos:

- 1 – 3: variable de resultado no importante. No se deben incluir en la tabla de evaluación de la calidad o de resultados. Estas variables de resultado no jugarán un papel importante en la formulación de las recomendaciones.
- 4 – 6: variable de resultado importante pero no clave para la toma de decisiones.
- 7 – 9: variable de resultado clave para la toma de decisiones.

La importancia relativa de las variables de resultado se establece mediante consenso.

B) Evaluación de la calidad de la evidencia científica

La evaluación de la calidad se realiza para cada una de las variables de resultado seleccionadas. Por tanto, para una pregunta clínica habrá muy probablemente variables de resultado con diferentes calificaciones de calidad. Inicialmente se debe evaluar la evidencia científica por el diseño de los estudios y su adecuación para responder a cada tipo de pregunta de la guía. Inicialmente, se considera como "*calidad alta*" a los ECA y como "*calidad baja*" a los estudios observacionales. No obstante se sugieren una serie de aspectos que pueden disminuir la calidad de los ECA o que la pueden aumentar en el caso de los estudios observacionales. La calidad de la evidencia científica se califica como alta, moderada, baja o muy baja.

El grupo elaborador ha clasificado la calidad de la evidencia para el conjunto de variables que se han considerado para formular cada recomendación. En la clasificación de la calidad global de la evidencia, el grupo elaborador ha considerado:

- Principalmente las variables críticas para la toma de decisiones.
- Cuando éstas han mostrado un resultado inconsistente (hacia un beneficio y hacia un daño) se ha considerado la calidad global como la más baja entre las diferentes variables.
- Cuando éstas han mostrado un resultado consistente (hacia un beneficio o un daño) se ha considerado la calidad global como la más alta entre las diferentes variables.
- Cuando éstas han mostrado un resultado consistente pero el balance entre los beneficios y los perjuicios es incierto se ha considerado la calidad global como la más baja entre las diferentes variables.

Aspectos que pueden disminuir la calidad de los ECA:

- **Limitaciones en el diseño o en la ejecución:** como pueden ser la ausencia de ocultamiento de la secuencia de aleatorización, enmascaramiento inadecuado, pérdidas importantes, ausencia de análisis por intención de tratar, finalización del estudio antes de lo previsto por razones de beneficio.
- **Resultados inconsistentes:** estimaciones discrepantes del efecto de un tratamiento (heterogeneidad o variabilidad en los resultados) en los estudios disponibles sugieren diferencias reales en dichas estimaciones. Estas pueden deberse a diferencias en la población, en la intervención, en las variables de resultado o en la calidad de los estudios. Cuando exista heterogeneidad no explicada de manera razonable la calidad disminuye.
- **Ausencia de evidencia científica directa:** en el caso de ausencia de comparaciones directas entre dos tratamientos (comparación de cada tratamiento frente a placebo, pero no entre ellos) o la extrapolación de los resultados de un estudio con un determinado fármaco al resto de fármacos de su misma familia en ausencia de un efecto de clase demostrado, también se considera evidencia científica indirecta. Es frecuente que existan grandes diferencias entre la

población en la que se aplicarán las recomendaciones y la incluida en los estudios evaluados. Por último, deben ser también valorados los aspectos de la potencial aplicabilidad en nuestro entorno o la validez externa de la evidencia científica disponible.

- **Imprecisión:** cuando los estudios disponibles incluyen relativamente pocos eventos y pocos/as pacientes y, por tanto, presentan intervalos de confianza amplios, la calidad disminuye.
- **Sesgo de notificación:** la calidad puede disminuir si se tiene la duda razonable sobre si la autoría ha incluido o no todos los estudios (por ejemplo, el sesgo de publicación en el contexto de una RS) o si se han incluido o no todas las variables relevantes de resultado (*outcome reporting bias*).

Aspectos que pueden aumentar la calidad en los estudios observacionales:

- **Efecto importante:** cuando el efecto observado muestra una asociación fuerte ($RR > 2$ o $< 0,5$) o muy fuerte ($RR > 5$ o $< 0,2$) y consistente, basada en estudios sin factores de confusión. En estas ocasiones se puede considerar la calidad como moderada o incluso alta.
- **La presencia de un gradiente dosis-respuesta.**

Situaciones en las cuales todos los posibles factores de confusión podrían haber reducido la asociación observada. En los casos en que el grupo de pacientes que reciben la intervención de interés presentan un peor pronóstico y, aún así, presentan mejores resultados que el grupo control, es probable que el efecto observado real sea mayor.

C) Graduación de la fuerza de las recomendaciones

La fuerza de las recomendaciones nos informa de en qué medida el llevar a cabo una recomendación conlleva más beneficios que riesgos. Existen una serie de factores que hay que tener en cuenta en la graduación de las recomendaciones:

- **Balance entre beneficios y riesgos:** Para realizar una adecuada valoración del balance entre los beneficios y los riesgos es necesario tener en cuenta el riesgo basal de la población a la que va dirigida la recomendación, y el efecto tanto en términos relativos como absolutos.
- **Calidad de la evidencia científica:** antes de llevar a cabo una recomendación es necesario conocer la certidumbre observada sobre la estimación del efecto. Si la calidad de la evidencia científica no es alta, a pesar de que la magnitud sea importante, debería disminuir la confianza y por tanto la fuerza con la que se lleve a cabo una recomendación.
- **Valores y preferencias:** la incertidumbre sobre los valores y las preferencias de la población diana a la cual va dirigida la GPC, será otro de los factores que hay que tener en cuenta. El personal sanitario, el colectivo de pacientes o la sociedad en general deben ver reflejados sus valores y sus preferencias que deberían influir en la graduación de las recomendaciones.

- **Costes:** a diferencia de otras variables de resultado, los costes son mucho más variables en el tiempo, en distintas áreas geográficas y en función de diversas implicaciones. Así, aunque un coste elevado disminuye la probabilidad de graduar como fuerte una recomendación, el contexto será crítico en la valoración final.

La graduación de la fuerza de las recomendaciones considera dos categorías: recomendaciones fuertes y recomendaciones débiles.

En las recomendaciones **fuertes** el grupo elaborador confía en que los efectos beneficiosos superan a los perjudiciales o viceversa, que los daños superan a los beneficios. En el primer caso la recomendación es fuerte a favor. En el segundo es fuerte en contra. La recomendación se redacta con la expresión ‘Se recomienda’.

Las recomendaciones **débiles** también pueden ser a favor o en contra. Una recomendación es débil a favor cuando el grupo elaborador concluye que los efectos beneficiosos de llevar a cabo la recomendación probablemente superan los perjudiciales, aunque no está completamente seguro. En cambio la recomendación es débil en contra, cuando los efectos adversos probablemente superan a los beneficiosos. La recomendación se redacta con la expresión ‘Se puede valorar’ o ‘Se sugiere’.

Para las intervenciones de las que no se dispone de evidencia y el grupo elaborador quiere resaltar un determinado aspecto, se dispone de las recomendaciones de **buena práctica clínica** (BPC), basadas en la experiencia clínica y el consenso del grupo elaborador.