

La importancia de la evaluación de la calidad de la evidencia

ROMINA BRIGNARDELLO-PETERSEN

DEPARTMENT OF HEALTH RESEARCH METHODS, EVIDENCE, AND IMPACT
MCMMASTER UNIVERSITY

Debido a que consideran toda la evidencia disponible para responder a una pregunta clínica, las revisiones sistemáticas de la literatura son reconocidas como el mejor diseño de investigación para informar la toma de decisiones clínicas y de salud pública [1]. Actualmente, todas las organizaciones y grupos que desarrollan guías de práctica siguiendo los estándares metodológicos más altos utilizan revisiones sistemáticas. Uno de los ejemplos más pertinentes son las recomendaciones hechas por la Organización Mundial de la Salud (OMS) para el tratamiento de pacientes con COVID-19 [2].

Uno de los factores más importantes al formular recomendaciones es la calidad de la evidencia sobre los beneficios y riesgos de usar una alternativa sobre otra. La calidad de la evidencia (también conocida como certeza de la evidencia) refleja cuánto podemos confiar en ella: qué tan seguros estamos de que los estimadores de efectos representan la verdad [3]. Para interpretar la evidencia de forma apropiada no es suficiente con saber que al comparar Baricitinib con tratamiento estándar en pacientes con COVID-19 la *Odds Ratio* para el desenlace mortalidad es 0,62 (intervalo de confianza de 95 %: 0,44 a 0,85); es también necesario saber si el cuerpo de evidencia (el conjunto de estudios que se usaron para obtener esos resultados) tiene limitaciones que deben considerarse antes de concluir que Baricitinib disminuye la mortalidad.

El sistema GRADE (*Grading of Recommendations Assessment, Development, and Evaluation*) provee un marco conceptual para evaluar la calidad de la evidencia de forma reproducible y transparente. La calidad de la evidencia puede ser alta, moderada, baja o muy baja. Calidad alta implica que tenemos mucha confianza de que el efecto real es muy similar al estimador de efecto obtenido. Por el contrario, calidad muy baja implica que tenemos muy poca confianza en el estimador de efecto, es decir, que el efecto real es probablemente bastante diferente al estimador de efecto [3].

Para determinar la calidad de la evidencia, GRADE establece que debemos considerar cinco categorías de potenciales limitaciones:

1. limitaciones en el diseño de los estudios: un conjunto de estudios que tiene alto riesgo de sesgo

disminuye la calidad de la evidencia [4];

2. inconsistencia: un conjunto de estudios que sugieren distintos efectos de tratamiento disminuye la calidad de la evidencia [5];
3. imprecisión: intervalos de confianza del estimador resumen (obtenido al hacer un meta-análisis) amplios que reflejan alta incertidumbre estadística disminuyen la calidad de la evidencia [6];
4. limitaciones en la pertinencia de la evidencia: estudios con limitada aplicabilidad o generalizabilidad disminuyen la calidad de la evidencia [7]; y
5. sesgo de publicación: el no poder incluir estudios con tamaño de muestra pequeña o resultados negativos –debido a que son más difíciles de encontrar al realizar una revisión sistemática porque es más difícil publicarlos en revistas científicas– disminuye la certeza de la evidencia [8].

El panel de expertos de la OMS determinó que la calidad de la evidencia para los efectos de Baricitinib en reducción de mortalidad es moderada: si bien había limitaciones debido a riesgo de sesgo, los tres estudios disponibles entregaban resultados consistentes, el intervalo de confianza obtenido en el meta-análisis era suficientemente estrecho (en términos absolutos, 6,8 % a 1,7 % menos de mortalidad), y no habían limitaciones de aplicabilidad o sesgo de publicación. Por lo tanto, concluyeron que Baricitinib probablemente reduce la mortalidad por COVID-19. El efecto de esta droga en otros desenlaces y su calidad (además de otras consideraciones relevantes), fue la base para formular una recomendación fuerte a favor del uso de Baricitinib.

La evaluación de la calidad de la evidencia permite obtener conclusiones que consideran todas las limitaciones importantes de manera formal [9], facilitando el no enfocarse únicamente en valores *p* y significancia estadística –una práctica que se ha intentado promover en los últimos años [10]. Por esta razón, esta evaluación es considerada un estándar de calidad metodológica en las revisiones sistemáticas. El grupo de trabajo GRADE ha publicado múltiples artículos en el que proveen guías para evaluar la calidad de la evidencia para

responder a distintos tipos de preguntas clínicas, incluyendo intervenciones, diagnóstico y pronóstico. El sistema GRADE ha sido adoptado por un gran número de organizaciones, incluyendo las asociaciones médicas e instituciones de salud pública más importantes alrededor del mundo.

En resumen, la interpretación apropiada de los resultados de una revisión sistemática requiere evaluar la calidad de la evidencia. El sistema GRADE es el sistema más utilizado para hacer esta evaluación. Quienes practican la toma de decisiones basada en evidencia a distintos niveles (con sus pacientes en la clínica o a nivel de población) pueden beneficiarse del uso del sistema GRADE, ya que permite adherir al segundo principio de la medicina basada en evidencia: para tomar una decisión adecuada, debemos entender qué tanto podemos confiar en la evidencia disponible.

Referencias

- [1] Guyatt G, Rennie D, Meade M, Cook D, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*, 3rd ed. New York, NY: McGraw-Hill Education, American Medical Association; 2015.
- [2] Lamontagne F, Agoritsas T, Macdonald H, et al. A living WHO guideline on drugs for covid-19, *BMJ*, 2020; 4, 370.
- [3] Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence, *J Clin Epidemiol*, 2011, 64(4): 401-406.
- [4] Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias), *J Clin Epidemiol*, 2011, 64(4): 407-415.
- [5] Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence—inconsistency, *J Clin Epidemiol*, 2011, 64(12): 1294-1302.
- [6] Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence—imprecision, *J Clin Epidemiol*, 2011, 64(12): 1283-1293.
- [7] Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence—indirectness, *J Clin Epidemiol*, 2011, 64(12): 1303-1310.
- [8] Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence—publication bias, *J Clin Epidemiol*, 2011, 64(12): 1277-1282.
- [9] Santesso N, Glenton C, Dahm P, et al. GRADE guidelines 26: Informative statements to communicate the findings of systematic reviews of interventions, *J Clin Epidemiol*, 2020, 119: 126-135.
- [10] Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance, *Nature*, 2019, 567(7748): 305-307.