

Herramientas para la lectura crítica de revisiones sistemáticas

Estefanía Quarteroni[✉], Juan P. Rodríguez[✉] y Luis Garegnani[✉]

Centro Cochrane Asociado, Universidad Hospital Italiano. Buenos Aires, Argentina.

RESUMEN

Las revisiones sistemáticas tienen como objetivo recopilar, analizar y sintetizar toda la evidencia científica disponible sobre una pregunta de investigación específica. A pesar de que las revisiones son una herramienta fundamental para la toma de decisiones basadas en la evidencia, su correcta interpretación y análisis pueden plantear ciertos desafíos. Este manuscrito aborda diferentes herramientas para guiar la lectura crítica de revisiones sistemáticas y determinar su utilidad con el fin de asistir a la toma de decisiones, basada en dos instrumentos como *A Measurement Tool to Assess Systematic Reviews Version 2* (AMSTAR 2) y la lista de verificación del *Critical Appraisal Skills Programme* (CASP) del Reino Unido. La lectura crítica de las revisiones sistemáticas es crucial para garantizar la validez, la calidad y la aplicabilidad de sus resultados.

Palabras clave: revisiones sistemáticas, lectura crítica, análisis, medicina basada en la evidencia.

Tools for the Critical Appraisal of Systematic Reviews

ABSTRACT

Systematic reviews aim to collect, analyze, and synthesize all available scientific evidence on a specific research question. Although reviews are a fundamental tool for evidence-based decision-making, their correct interpretation and analysis can present certain challenges. This manuscript addresses different tools to guide the critical appraisal of systematic reviews and to determine their usefulness in supporting decision-making, based on two instruments: *A Measurement Tool to Assess Systematic Reviews Version 2* (AMSTAR 2) and the checklist of the United Kingdom *Critical Appraisal Skills Programme* (CASP). Critical appraisal of systematic reviews is crucial to ensure the validity, quality, and applicability of their results.

Keywords: systematic reviews, critical appraisal, analysis, evidence-based medicine.

Una revisión sistemática (RS) es un tipo de investigación secundaria cuyo objetivo es recopilar, analizar y sintetizar toda la evidencia científica disponible (publicada o no) que cumpla con criterios de elegibilidad previamente establecidos para responder una pregunta concreta de investigación. Las RS aplican una metodología transparente y reproducible, evaluando los aspectos metodológicos de los estudios incluidos para minimizar sesgos¹. Además, cuando los estudios incluidos en la RS son lo

suficientemente similares, tanto clínica como estadística y metodológicamente, en el informe de sus resultados, pueden incluir metanálisis, una técnica estadística que permite combinar matemáticamente los resultados de múltiples estudios. También mejora la precisión en la estimación de los efectos y permite evaluar su impacto en distintos subgrupos de pacientes².

Las RS son una herramienta esencial para los pacientes, médicos, investigadores y funcionarios de la salud, ya

Autor para correspondencia: juan.rodriguez@hospitalitaliano.org.ar, Rodríguez JP.

Recibido: 21/05/2025 | Aceptado: 5/01/2026 | Publicado:

DOI: <http://doi.org/10.51987/rev.hosp.ital.b.aires.v46i1.1274>

Cómo citar: Quarteroni E, Rodríguez JP, Garegnani L. Herramientas para la lectura crítica de revisiones sistemáticas. *Rev Hosp Ital B.Aires.* 2026;46(1):e0001274

que proporcionan una síntesis objetiva de las investigaciones disponibles, facilitando la toma de decisiones basadas en evidencia científica; de esta manera constituyen un pilar fundamental en el paradigma de la medicina basada en la evidencia (MBE), que integra los valores y preferencias personales de los pacientes con la experiencia clínica y los datos científicos más actualizados³.

A pesar de que las RS son una herramienta fundamental para la toma de decisiones basadas en la evidencia, su correcta interpretación y análisis pueden presentar ciertos desafíos. En el contexto clínico-asistencial, el tiempo limitado para la lectura profunda y reflexiva de una RS es una barrera común⁴ sumada a las dificultades asociadas a su complejidad metodológica. Al utilizar elementos y lenguaje técnico, puede resultar poco práctica su lectura para quienes no están familiarizados con esos conceptos. Además, los resultados de una RS no siempre son concluyentes, fáciles de interpretar o de trasladar a la práctica clínica debido a la variabilidad entre estudios, los sesgos de publicación o la calidad metodológica de los ensayos incluidos, que pueden influir en la validez de los hallazgos⁵. Para superar estos inconvenientes, la MBE propone la lectura o evaluación crítica de manuscritos científicos en general y de RS en particular, examinando sistemáticamente un artículo para juzgar su valor, fiabilidad y relevancia en el contexto en el que se pretende implementar, y guiando a los profesionales para interpretar y utilizar la evidencia científica de manera efectiva⁶.

Existen múltiples herramientas para valorar diferentes aspectos de las RS, como las listas de verificación PRISMA para evaluar el informe (del inglés *Preferred Reporting Items for Systematic reviews and Meta-Analyses*)⁷ o la herramienta ROBIS (del inglés *Risk of Bias in Systematic reviews*)⁸, orientada principalmente a los desarrolladores de guías de práctica clínica, para evaluar el riesgo de sesgo en las RS. En este manuscrito abordaremos otras dos herramientas desarrolladas para facilitar la lectura crítica de las RS.

Una de las herramientas para guiar la lectura crítica de RS de intervenciones es *A Measurement Tool to Assess Systematic Reviews Version 2* (AMSTAR-2)⁹. AMSTAR 2 permite determinar si una RS proporciona un resumen preciso y completo de los resultados de los estudios disponibles que abordan la pregunta de interés específica. Su propósito principal es ayudar a los profesionales, investigadores y responsables de la toma de decisiones a evaluar la confianza en una RS para fundamentar decisiones clínicas o sanitarias importantes. Para la implementación de AMSTAR 2 es necesario responder una serie de preguntas que abordan aspectos clave de la metodología de una RS. Contempla un cuestionario de 16 preguntas, con 6 dominios críticos con respuestas simples (“sí”; “no” o “sí parcial”). De este cuestionario, que no tarda más de 30 minutos en completarse¹⁰, se obtiene una estimación de la confianza de una RS (Tabla 1) para asistir en la toma de decisiones. Es tal el impacto y la importancia de estos dominios que basta con no

Tabla 1. Dominios críticos y valoración de la confianza general de una RS según AMSTAR-2)

Dominios críticos de la herramienta <i>A measurement Tool to Assess Systematic Reviews Version 2</i> (AMSTAR-2)	
1	Protocolo registrado antes de la revisión
2	Adecuada búsqueda en la literatura
3	Justificación de los estudios excluidos
4	Riesgo de sesgo de los estudios individuales incluidos
5	Métodos metanalíticos apropiados
6	Consideración del riesgo de sesgo en la interpretación de los resultados de la revisión
7	Evaluación de la presencia y el impacto probable del sesgo de publicación
Valoración de la confianza general en los resultados de la revisión	
Confianza	Justificación
Alta	Ninguna debilidad crítica y hasta una no crítica: la RS* proporciona un resumen exacto y completo de los resultados de los estudios disponibles
Media	Ninguna debilidad crítica y más de una debilidad no crítica (aunque, si son muchas, podría justificarse una baja confianza): la RS* tiene debilidades, pero no hay defectos críticos y puede proporcionar un resumen preciso de los resultados de los estudios disponibles
Baja	Hasta una debilidad crítica, con puntos débiles no críticos o sin ellos: la RS* puede no proporcionar un resumen exacto y completo de los estudios disponibles
Críticamente baja	Más de una debilidad crítica, con debilidades no críticas o sin ellas: la RS* no es fiable

RS: Revisión sistemática.

presentar registro prospectivo del protocolo para que la RS sea de baja confianza. Por ejemplo, la RS de von Luckner, et al.¹¹ sobre el efecto de la suplementación con magnesio para la profilaxis de la migraña no presenta en su sección Métodos el número de registro o publicación del protocolo. En comparación, la revisión de Talandashiti, et al.¹² sobre una pregunta de investigación similar, sí lo informa. El registro del protocolo de una RS aporta transparencia y replicabilidad a la investigación.

Se recomienda aplicar esta metodología de análisis de calidad en cada una de las RS, a partir de las cuales se podrá tomar una decisión clínica o diseñar una política en salud^{13,14}. Es posible que, teniendo en cuenta las características, el tipo y la metodología de cada revisión, se requiera asesoramiento especializado para responder a alguno de los dominios analizados, por ejemplo, para determinar si los autores de la revisión evaluaron adecuadamente el riesgo de sesgo o para determinar la correcta implementación de los métodos estadísticos.

Además de ser un instrumento útil no solo para quien lee una RS, sino también para el equipo que la elabora, es, en esencia, una herramienta educativa¹³. Es importante incluirla en la formación académica de los estudiantes de las ciencias de la salud, de modo que la lectura crítica de las síntesis de la evidencia sea una práctica habitual en la vida profesional. Sin embargo, AMSTAR 2 no está libre de limitaciones: su aplicación

requiere juicio experto y una comprensión profunda de la metodología de RS, lo que puede generar variabilidad entre evaluadores y limitar su uso por parte de usuarios con menor experiencia. Además, no produce una puntuación global cuantitativa, sino una clasificación basada en dominios críticos, lo que puede dificultar la comparación entre RS. Finalmente, algunos ítems son interpretativamente ambiguos o difíciles de aplicar a ciertos tipos de revisiones, lo que puede reducir su consistencia y su utilidad práctica.

Otra herramienta desarrollada para guiar la lectura crítica de RS es la lista de verificación del *Critical Appraisal Skills Programme* (CASP). La herramienta ayuda a determinar la fiabilidad de los resultados y la relevancia de una RS para poder aplicar los hallazgos en la práctica clínica, facilitando así la toma de decisiones a través de un conjunto de preguntas estructuradas de respuesta simple. La lista de verificación CASP para RS se divide en 4 bloques, cada uno con un objetivo de análisis específico (Tabla 2). El primer bloque trata de determinar si la RS presenta un diseño válido, con una pregunta claramente definida. Por ejemplo, una RS que evaluó los efectos de diferentes formulaciones de nutrición enteral en pacientes con pancreatitis aguda¹⁵ definió de manera precisa los componentes de la pregunta PICO (Población, Intervención, Comparador, Desenlaces), incluyendo en la Población a todas las personas mayores con pancreatitis

Tabla 2. Secciones de la herramienta *Critical Appraisal Skills Programme* (CASP) para revisiones sistemáticas

	Pregunta objetivo	Preguntas guía
Sección A	¿El diseño básico del estudio es válido para una revisión sistemática?	<ul style="list-style-type: none"> - ¿La revisión sistemática abordó una pregunta de investigación claramente formulada? - ¿Los investigadores buscaron diseños de estudio apropiados para responder a la pregunta de investigación?
Sección B	¿Es metodológicamente sólida la revisión sistemática?	<ul style="list-style-type: none"> - ¿La estrategia de búsqueda es exhaustiva y está claramente documentada? - ¿Los autores definieron criterios de selección, inclusión y exclusión adecuados?
Sección C	¿Cuáles son los resultados?	<ul style="list-style-type: none"> - ¿Se han informado de manera exhaustiva los resultados de la revisión sistemática y se han interpretado adecuadamente? - ¿El número total de participantes en la revisión sistemática fue suficiente para detectar un efecto?
Sección D	¿Los resultados ayudarán en el ámbito local?	<ul style="list-style-type: none"> - ¿Pueden aplicarse los resultados de la revisión sistemática a su población local/en su contexto local? - ¿Qué recursos se necesitarán para implementar la intervención?

aguda; en la Intervención cualquier tipo de nutrición enteral con una formulación específica; en el Comparador tipos diferentes de nutrición enteral, placebo o ninguna intervención, y en los Desenlaces la mortalidad por todas las causas, entre otros. En contraste, no encontramos una descripción detallada de la pregunta de investigación en otra RS sobre el mismo tema¹⁶.

El segundo bloque indaga sobre la búsqueda de estudios, idealmente en más de una base de datos relevante (por ejemplo, MEDLINE/PubMed, Embase), y su correcto registro en el informe. También considera el uso de restricciones injustificadas por idioma y de búsquedas adicionales mediante la revisión manual de referencias de estudios incluidos o la consulta a expertos. Por ejemplo, una RS sobre el panorama general del dengue¹⁷ limitó la elegibilidad a estudios publicados en inglés. Esto puede provocar que se omitan estudios importantes publicados en otros idiomas y, por lo tanto, que la revisión no abarque toda la evidencia disponible, especialmente considerando que algunas investigaciones relevantes sobre enfermedades endémicas como el dengue suelen publicarse también en español o en portugués.

El segundo bloque asimismo se enfoca en evaluar la validez interna de los estudios incluidos, es decir, si su diseño y ejecución fueron adecuados para minimizar sesgos¹⁸. Mientras muchas revisiones sistemáticas omiten esta evaluación, las revisiones de Cochrane la incluyen sistemáticamente mediante herramientas de evaluación del riesgo de sesgo. También se analiza si los métodos estadísticos del metanálisis fueron correctos, bien aplicados e informados.

La tercera sección de la lista de verificación CASP para RS analiza la forma en que se presentan e interpretan los resultados, y evalúa si los estimadores de efecto de los análisis y su interpretación narrativa son consistentes entre sí. Además, tiene en cuenta la precisión de estos estimadores de efecto, considerando si el número total de participantes en la sumatoria de estudios incluidos en la RS fue suficiente para detectar un efecto o si los autores proporcionaron intervalos de confianza y valores P para evaluar si los resultados son estadísticamente significativos. Por ejemplo, una revisión Cochrane sobre el uso de un anticuerpo monoclonal para la prevención del virus sincitial respiratorio¹⁹ informó preocupaciones acerca del efecto de la intervención relacionadas con la escasez de participantes en los estudios incluidos. Sin embargo, otra revisión sobre el tema²⁰ no incluyó un análisis de estas características al interpretar sus hallazgos.

La última sección analiza en particular la aplicabilidad de los resultados en el ámbito local. Por ejemplo, una revisión Cochrane²¹ sobre vacunas orales inactivadas contra el cólera demostró que dos vacunas (Dukoral® y Shanchol®) reducen los casos de cólera y que fueron seguras y bien toleradas. A pesar de que los resultados de la revisión son válidos y metodológicamente sólidos, la aplicabilidad de la vacuna en otro contexto con recursos limitados y dificultades para mantener la cadena de frío,

podría no ser tan eficaz debido a las barreras logísticas, económicas y culturales.

Es fundamental que el lector de una RS evalúe si la intervención analizada realmente aporta un equilibrio aceptable entre beneficios y daños para los pacientes, si los recursos necesarios para implementarla están disponibles y si su utilización es razonable para dicha intervención. Este análisis es clave tanto en la práctica clínica como en la toma de decisiones basadas en evidencia.

La principal debilidad de la lista de verificación CASP para RS radica en que se basa en preguntas cualitativas y abiertas, sin criterios de valoración específicos ni ponderación de los ítems, lo que puede conducir a interpretaciones subjetivas y variabilidad entre evaluadores. Además, ofrece una evaluación general de la credibilidad y la aplicabilidad de la revisión, pero no permite identificar con precisión los aspectos metodológicos específicos que afectan su calidad, lo que limita su uso para comparar RS.

CONCLUSIÓN

La síntesis de la evidencia a través de las RS es fundamental para la toma de decisiones en salud y para el desarrollo e implementación de políticas sanitarias, proporcionando información fiable y objetiva. Sin embargo, la lectura crítica de estas RS es crucial para garantizar la validez, la calidad y la aplicabilidad de sus resultados. Ciertas herramientas como AMSTAR 2 y CASP son valiosas para este propósito, y son de gran ayuda tanto para profesionales como para estudiantes de carreras médicas o de salud. También constituyen un pilar fundamental en la evaluación de tecnologías sanitarias y en el desarrollo de guías de práctica clínica.

Existen puntos en común entre las herramientas mencionadas. Es importante que los profesionales de la salud adquieran habilidades de lectura crítica y utilicen estas herramientas de manera efectiva para garantizar que la evidencia científica se traduzca en una atención de calidad para los pacientes, en argumentos fehacientes y fiables para la toma de decisiones médicas y, finalmente, en políticas públicas serias y con sustento científico y metodológico.

Contribuciones de autores: Conceptualización, Investigación, Redacción - Borrador original, Redacción-Revisión y edición, Visualización, Administración de proyecto (EQ, JPR, LIG). Supervisión (LIG)

Conflictos de intereses: los autores declaran no poseer conflictos de intereses relacionados con el contenido del presente trabajo.

Financiamiento: los autores declaran que este estudio no recibió financiamiento de ninguna fuente externa.

REFERENCIAS

1. Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions. Hoboken, NJ: John Wiley & Sons; 2011. 553 p.
2. Walker E, Hernandez AV, Kattan MW. Meta-analysis: its strengths

- and limitations. *Cleve Clin J Med*. 2008;75(6):431-439. <https://doi.org/10.3949/ccjm.75.6.431>.
3. Mendis K. What is evidence-based medicine? *Ceylon Med J*. 2021;66(3):154-156. <https://doi.org/10.4038/cmj.v66i3.9495>.
 4. Río-de-la-Loza-Zamora JG, López-Ortiz G. Barreras para el desarrollo de investigación en medicina familiar en Iberoamérica: revisión sistemática. *Rev Mex Med Fam*. 2022;30(9(2)):49-58. <https://doi.org/10.24875/RMF.21000125>.
 5. Sgarbossa N, Ibáñez Cobaisse M, González Cianciulli G, et al. Revisiones sistemáticas: conceptos clave para profesionales de la salud. *Medwave*. 2022;22(9):2622. <http://dx.doi.org/10.5867/medwave.2022.09.2622>.
 6. Critical Appraisal Skills Programme. How to use the CASP checklists [Internet]. Oxford: CASP; 2025 [consulta 2025 may 20]. Disponible en: <https://casp-uk.net/how-to-use-checklist/>.
 7. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. <https://doi.org/10.1136/bmj.n71>.
 8. Whiting P, Savović J, Higgins JP, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. 2016;69:225-234. <https://doi.org/10.1016/j.jclinepi.2015.06.005>.
 9. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*. 2007;7:10. <https://doi.org/10.1186/1471-2288-7-10>.
 10. Ciapponi A. AMSTAR-2: herramienta de evaluación crítica de revisiones sistemáticas de estudios de intervenciones de salud. *Evid Actual Pract Ambul*. 2018;21(1):4-13. <https://doi.org/10.51987/evidencia.v21i1.6834>.
 11. von Luckner A, Riederer F. Magnesium in migraine prophylaxis-is there an evidence-based rationale? A systematic review. *Headache*. 2018;58(2):199-209. <https://doi.org/10.1111/head.13217>.
 12. Talandashti MK, Shahinfar H, Delgarm P, et al. Effects of selected dietary supplements on migraine prophylaxis: a systematic review and dose-response meta-analysis of randomized controlled trials. *Neurol Sci*. 2025;46(2):651-670. <https://doi.org/10.1007/s10072-024-07794-0>.
 13. Oxman AD, Cook DJ, Guyatt GH. Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. *JAMA*. 1994;272(17):1367-1371. <https://doi.org/10.1001/jama.272.17.1367>.
 14. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358:j4008. <https://doi.org/10.1136/bmj.j4008>.
 15. Poropat G, Giljaca V, Hauser G, et al. Enteral nutrition formulations for acute pancreatitis. *Cochrane Database Syst Rev*. 2015;2015(3):CD010605. <https://doi.org/10.1002/14651858.CD010605.pub2>.
 16. Petrov MS, Loveday BP, Pylypchuk RD, et al. Systematic review and meta-analysis of enteral nutrition formulations in acute pancreatitis. *Br J Surg*. 2009;96(11):1243-1252. <https://doi.org/10.1002/bjs.6862>.
 17. Khan MB, Yang ZS, Lin CY, et al. Dengue overview: an updated systemic review. *J Infect Public Health*. 2023;16(10):1625-1642. <https://doi.org/10.1016/j.jiph.2023.08.001>.
 18. Gervas J, Pérez Andrés C. Evaluación de las intervenciones en salud: la búsqueda del equilibrio entre la validez interna de los resultados y la validez externa de las conclusiones. *Rev Esp Salud Pública*. 2008;82(6):577-579. <https://doi.org/10.1590/s1135-57272008000600001>.
 19. Garegnani L, Styrmisdóttir L, Roson Rodríguez P, et al. Palivizumab for preventing severe respiratory syncytial virus (RSV) infection in children. *Cochrane Database Syst Rev*. 2021;11(11):CD013757. <https://doi.org/10.1002/14651858.CD013757.pub2>. Update in: *Cochrane Database Syst Rev*. 2025;7:CD013757. <https://doi.org/10.1002/14651858.CD013757.pub3>.
 20. Gonzales T, Bergamasco A, Cristarella T, et al. Effectiveness and safety of palivizumab for the prevention of serious lower respiratory tract infection caused by respiratory syncytial virus: a systematic review. *Am J Perinatol*. 2024 ;41(S 01):e1107-e1115. <https://doi.org/10.1055/a-1990-2633>.
 21. Saif-Ur-Rahman KM, Mamun R, Hasan M, et al. Oral killed cholera vaccines for preventing cholera. *Cochrane Database Syst Rev*. 2024;1(1):CD014573. <https://doi.org/10.1002/14651858.CD014573>.