



## **Existen muchas listas en investigación, pero... no todas sirven para lo mismo**

---

### **There are many lists of research, but... not all serve the same**

Micaela Sánchez-Martín, Ana Isabel Ponce Gea, Marina Pedreño Plana, Francisco Javier Ibáñez-López, y Fernando Navarro-Mateu

Universidad de Murcia, Murcia, España

**“La verdad siempre se halla en la simplicidad y  
no en la multiplicidad y confusión de las cosas”  
(Isaac Newton -1642-1727- Matemático y físico británico)**

#### **Resumen**

En las últimas décadas se ha ido elaborando una amplia variedad de listas de comprobación para diferentes finalidades. En esta píldora metodológica nos centraremos en dar a conocer a la comunidad académica, alumnado, profesorado y colectivo investigador, la gran utilidad de las listas de: lectura crítica de los artículos científicos, de las orientadas a la mejora en redacción de la producción científica y las de evaluación del riesgo de sesgos de los artículos ya publicados. El objetivo es ayudar a clarificar la finalidad de cada una de ellas y facilitar su utilización en diferentes fases de la investigación.

**Palabras clave:** listas de comprobación; evaluación de la calidad; lectura crítica.

#### **Abstract**

In recent decades, a wide variety of checklists have been developed with different purposes. Therefore, in this methodological pill we will focus on making the academic community, students, teachers and research collective aware of the great utility of the lists of: critical reading of scientific articles, of those aimed at improving the writing of the scientific production and those of evaluation of the risk of biases of the already published articles. The objective is to help clarify the purpose of each of them, and to facilitate their use in different phases of the investigation.

**Keywords:** Checklists; quality assessment; critical reading.

---

Fecha de recepción: 14/06/2023

Fecha de aceptación: 15/08/2023

Correspondencia: Ana Isabel Ponce Gea, Universidad de Murcia, Murcia, España  
Email: [anaisabel.ponce@um.es](mailto:anaisabel.ponce@um.es)

### Puntos clave

#### Lo que se sabe

- Actualmente hay un gran número de listas de comprobación con diferentes finalidades, lo que ha generado cierta confusión sobre su utilización en el ámbito educativo, en diferentes contextos para los que fueron diseñadas.

#### Lo que aporta este trabajo

- Explicar la utilidad de las listas de comprobación.
- Diferenciar la aplicación y uso de las distintas listas de comprobación según su finalidad
- Determinar la lista de lectura crítica de literatura científica más apropiada según el diseño del informe seleccionado.
- Seleccionar la lista idónea para mejorar la redacción del informe de investigación en función de su diseño.
- Identificar la escala de evaluación más adecuada para la evaluación del riesgo de sesgos de los artículos científicos según su diseño y cómo utilizarlas.

### Escenario práctico

Un doctorando está realizando una Tesis Doctoral (TD) con un diseño observacional (estudio transversal) y su tutor le ha recomendado que utilice una lista de comprobación para la redacción del trabajo. En muchas de sus lecturas han aparecido diferentes escalas y listas de comprobación: PRISMA, STROBE, AMSTAR, ROBINS, STAR, lectura crítica, JADAD, herramienta de riesgo de sesgo de la Cochrane, escala de Newcastle-Otawa, ... Sin embargo, no tiene una idea clara de su utilidad. Le surgen muchas preguntas, por ejemplo, ¿son todas iguales?, ¿de qué depende su utilización?, ¿es obligatorio su uso?, ¿cómo se utilizan? Y ..., sobre todo, ¿cual tendría que utilizar para la redacción de su TD?

### ¿Listas?, pero ¿qué listas? Muchas, pero no todas sirven para lo mismo

Desde la publicación de las primeras listas de lectura crítica en el contexto de la Medicina Basada en la Evidencia (MBE) (Guyatt & Rennie, 1993), se han dado a conocer un gran número de listas de comprobación con diferentes finalidades. De forma genérica, estas listas pueden ser clasificadas en tres grandes grupos (ver Tabla 1): i) herramientas para la lectura crítica de artículos científicos, ii) listas de comprobación para la redacción de artículos científicos y iii) listas para evaluar el riesgo de sesgos (errores sistemáticos generalmente relacionados con errores en el diseño, desarrollo o análisis metodológico) de los estudios científicos en el contexto de una revisión sistemática/meta-análisis.

Sin embargo, existe cierta confusión sobre su utilización en diferentes contextos para los que fueron diseñadas inicialmente. Por ejemplo, no sería adecuado utilizar las listas de lectura crítica para evaluar el riesgo de sesgos de un estudio científico (Jerez Yáñez et al., 2016, -utiliza la *Critical Appraisal Skills Programme* para evaluar la calidad metodológica-) o usar una lista de comprobación para mejorar la calidad y transparencia de la publicación en la evaluación del riesgo de sesgos en el contexto de una RS (Morales-Toro et al., 2019, -utiliza la lista de comprobación *STROBE* para evaluar la calidad metodológica de los estudios incluidos en la RS-). Aunque relacionadas entre sí, no fueron diseñadas para esta finalidad. Por ejemplo, en la descripción de las listas de los dos primeros grupos (lectura crítica y de comprobación para mejorar la publicación) no se define un punto de corte necesario para valorar el riesgo de sesgo de las publicaciones.

Aunque este tipo de listas ya están siendo ampliamente utilizadas en el campo de la investigación en Ciencias de la Salud, en Ciencias Sociales y, más concretamente, en el campo de estudio educativo, están empezando a ser conocidas y utilizadas. El objetivo de esta píldora metodológica es ayudar a clarificar la finalidad de cada una de ellas, dar a conocer la gran utilidad de estas herramientas a toda la comunidad académica (tanto al alumnado como al profesorado y a los/as investigadores/as) y facilitar su utilización en diferentes fases de la investigación, tanto en la lectura crítica de los artículos científicos, en la redacción de la producción científica y en la evaluación del riesgo de sesgos de los artículos ya publicados.

### **Listas para la lectura crítica**

Una de las fases más importantes en la aplicación de la Práctica Basada en Evidencias (PBE) (Baños et al., 2021; Sánchez-Martín et al., 2022) consiste en la lectura crítica de los artículos científicos encontrados tras la búsqueda bibliográfica diseñada para contestar a nuestra pregunta de investigación (Sánchez-Martín et al., 2023). La publicación de un artículo científico en cualquier revista científica no garantiza necesariamente la calidad metodológica del mismo, incluso en revistas con un proceso de revisión por pares. Por ello es conveniente la lectura atenta del apartado de metodología, apartado con el que no suelen sentirse muy cómodos la mayoría de los lectores, especialmente los estudiantes de Educación Superior y profesionales con poca experiencia en investigación. Para ayudar a la lectura de este importante apartado hay disponibles una serie de fichas diseñadas para cada uno de los diseños de investigación más importantes (por ejemplo, los estudios experimentales, los estudios de cohortes, de casos y controles y los transversales).

En su diseño, los autores seleccionaron el menor número de preguntas, aunque suficientes para realizar un cribado rápido sobre la calidad metodológica, los resultados y la aplicabilidad del estudio. Las preguntas se estructuran en tres grandes apartados: A) ¿son válidos los resultados del estudio? (validez interna o calidad metodológica del estudio); B) ¿cuáles han sido los resultados del estudio?; y C) ¿me resultarán útiles los resultados? (validez externa o generalización de los resultados). El número de preguntas por apartado suele oscilar entre cuatro y seis. El apartado más extenso es el que analiza la validez interna o metodológica del estudio. Los autores lo colocaron en primer lugar puesto que, si un estudio no responde a un mínimo de preguntas de forma adecuada, a criterio del lector, ¿qué sentido tiene dedicarle tiempo a analizar los resultados y la discusión? En general, suelen evaluarse con tres posibles respuestas: i) si los autores han definido bien la respuesta, ii) no queda clara o iii) no se hace mención. Como muestra la tabla 1, desde las fichas originales publicadas en la prestigiosa revista de la Asociación Médica Americana (*JAMA, Journal of the American Medical Association*) (Guyatt & Rennie, 1993; Users' Guide to the Medical Literature, s. f.), su uso se ha ido extendiendo y en la actualidad se dispone de numerosos ejemplos en castellano (Fichas de Lectura Crítica, s. f.; Redcaspe – Programa de Habilidades en Lectura Crítica Español, s. f.; Sánchez et al., 2007).

El empleo de este tipo de fichas de lectura crítica suele restringirse a la utilización personal de artículos, como una forma de adquirir y entrenar de modo dirigido las competencias para la lectura crítica de artículos científicos según las necesidades específicas del lector. Sin embargo, su uso más frecuente es para compartir información en grupos de trabajo, en los denominados Clubes de Lectura Crítica (*Journal's Clubs*) (Martínez de Guzman et al., 2019). Estos grupos de lectura crítica están considerados como una de las herramientas más extendidas para la adquisición de las competencias necesarias para la PBE (Kyriakoulis et al., 2016; Patelarou et al., 2017). El entrenamiento en lectura crítica con la utilización de estas fichas permite además la adquisición de competencias de una forma estructurada y amigable y también se utilizan en investigación (como la búsqueda de literatura científica y la aplicación de conceptos estadísticos y de diseño de investigación) (Maloney et al., 2019) y podrían ser incorporadas en la Educación Basada en Evidencias (EBE) (Greenhalgh et al., 2003).

**Tabla 1**

*Listas de comprobación para la lectura crítica*

Descripción	Ejemplos y recomendaciones
<ul style="list-style-type: none"> <li>▪ Requisitos mínimos para la lectura crítica de artículos científicos.</li> <li>▪ Generalmente agrupados en tres apartados: validez interna, resultados y validez externa.</li> <li>▪ Uso recomendado: lectura crítica individual o en sesiones bibliográficas de lectura crítica en grupos de investigación o de trabajo (clubes de lectura crítica o <i>journal clubs</i>)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Colección “<i>Users’ Guide to eh Medical Literature</i>” (<i>Users’ Guide to the Medical Literature</i>, s. f.)</li> <li>▪ Caspe (<i>Redcaspe</i> – Programa de Habilidades en Lectura Crítica Español, s. f.)</li> <li>▪ Club de Lectura Crítica del Hospital Psiquiátrico Román Alberca: Una experiencia práctica (Martínez de Guzmán et al., 2019)</li> <li>▪ Fichas de Lectura Crítica (Fichas de Lectura Crítica, s. f.)</li> <li>▪ Atención Sanitaria Basada en la evidencia: su aplicación a la práctica clínica (Sánchez et al., 2007)</li> <li>▪ <i>How to read a paper</i> (Greenhalgh, 2019)</li> <li>▪ <i>Best Evidence Topics</i> (BETs) <i>Critical Appraisal Worksheets</i> (<i>BestBETs - BETs CA Worksheets</i>, s. f.)</li> </ul>

### Listas de comprobación para la escritura científica

El aumento exponencial de las publicaciones científicas en la actualidad no garantiza necesariamente la calidad de lo publicado. En este contexto, empezaron a surgir varias iniciativas por parte de diferentes grupos de trabajo que culminaron en la elaboración de diversas guías o listas de comprobación. Estas listas se han ido actualizando con el tiempo. Por ejemplo, a finales de esa misma década se publicó la guía QUORUM (*Quality Of Reporting Of Meta-analyses*) una guía específicamente diseñada para mejorar la calidad y transparencia de las publicaciones de los meta-análisis de estudios experimentales controlados y aleatorizados (Moher et al., 1999). Unos años después, se desarrolló la guía PRISMA (*Preferred Reporting Items for Systematic reviews and Meta-Analyses*) para revisiones sistemáticas (Moher et al., 2009), cuya versión ha sido actualizada recientemente con la PRISMA 2020 (Page, et al., 2021). La tabla 2 muestra distintas guías según los diferentes diseños de investigación.

**Tabla 2**

*Listas de comprobación para mejorar la calidad y transparencia de las publicaciones*

Descripción	Ejemplos y recomendaciones
<ul style="list-style-type: none"> <li>▪ Elaboradas para ayudar a mejorar la calidad y transparencia de los informes científicos.</li> <li>▪ Listado exhaustivo que incluye todos los apartados, desde el título hasta la financiación.</li> <li>▪ Redactadas como recomendaciones.</li> <li>▪ Los/as autores/as no recomiendan su uso para la redacción de proyectos o para la evaluación del riesgo de sesgos.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Revisiones sistemáticas y meta-análisis: PRISMA 2020 (<i>Preferred Reporting Items for Systematic reviews and Meta-Analyses</i>) (Page, Moher, et al., 2021)</li> <li>▪ Estudios experimentales: CONSORT (<i>Consolidated Standard of Reporting Trials</i>) (Schulz et al., 2010)</li> <li>▪ Estudios observacionales (Transversales, de casos y controles y de cohortes): STROBE (<i>Strengthening the Reporting of Observational Studies in Epidemiology</i>) (von Elm et al., 2007)</li> <li>▪ Estudios cualitativos: SRQR (<i>Standards for Reporting Qualitative Research</i>) (O’Brien et al., 2014)</li> </ul>

También, se han ido elaborando numerosas extensiones adaptando las recomendaciones a diferentes situaciones específicas. Por ejemplo, de la PRISMA se han publicado, entre otras, la PRISMA-S (*PRISMA Statement for Reporting Literature Searches in Systematic Reviews*) para la publicación de los procesos de búsqueda de la literatura científica (Rethlefsen et al., 2021), la PRISMA-P (*Preferred*

*Reporting Items for Systematic Review and Meta-Analysis Protocols*) para la publicación de los protocolos de futuras revisiones sistemáticas (Shamseer et al., 2015), y la *PRISMA 2020 for Abstracts* para la presentación de resúmenes de revisiones sistemáticas en revistas y congresos (Page et al., 2021).

Desde hace más de una década, la mayoría de estas listas de comprobación pueden consultarse en la página web de la organización internacional *EQUATOR-network* (<https://www.equator-network.org/>). Aunque su desarrollo se centra inicialmente en el campo de la investigación en salud, estas listas de comprobación pueden y han de ser aplicadas en otros campos de conocimiento, como en el educativo (Balakrishnan, 2021).

Este tipo de listas no fueron diseñadas para ser utilizadas como guías en el diseño de los diferentes estudios y tampoco se recomienda su utilización para la evaluación del riesgo de sesgos en el contexto de una RS, ya que existen otras herramientas específicamente diseñadas para ello (Page, McKenzie, et al., 2021). Su finalidad explícita es la de servir de guías orientativas para la redacción de las publicaciones científicas. Sin embargo, parece claro que la experiencia en el manejo de estas guías es de gran utilidad para el diseño y planificación de proyectos de investigación.

Hoy en día cada vez son más las revistas científicas de prestigio (*The Lancet*: <https://www.thelancet.com/preparing-your-manuscript> o las revistas incluidas en el grupo *PLOS*: [www.plos.org](https://www.plos.org)) que están incorporando la sugerencia u obligatoriedad de utilizar este tipo de listas de comprobación a los autores interesados en publicar en ellas. Aunque su uso no está muy extendido aún entre las revistas españolas, es de esperar que las vayan incorporando entre sus sugerencias en los próximos años, puesto que existen datos que sugieren que su utilización para la redacción de artículos científicos con el objetivo de mejorar la calidad y transparencia del manuscrito se asocia a unas publicaciones de mayor calidad metodológica (Panic et al., 2013; Vu-Ngoc et al., 2018). Así, la normativa APA (American Psychological Association, 2020) en su séptima edición, proporciona listas de comprobación con pautas rigurosas sobre qué y cómo informar en los trabajos de investigación según el diseño empleado, en los campos de psicología, educación y otras ciencias sociales.

### **Listas para la evaluación del riesgo de sesgos en una revisión sistemática**

En la actualidad, disponemos de numerosos ejemplos en los que se comprueba que la calidad de las publicaciones es mejorable. Por ejemplo, en la década de los 90 del siglo pasado, diversos estudios encontraron que la calidad de los artículos de revisiones sistemáticas y meta-análisis era, generalmente, pobre (Mulrow, 1987; Sacks et al., 1987).

Las RS están siendo consideradas como el patrón oro en la síntesis de los resultados de los estudios científicos, en la toma de decisiones en gestión y en la elaboración de políticas educativas y sanitarias (Pussegoda et al., 2017). Uno de los pasos imprescindibles para realizar correctamente una RS es la evaluación del riesgo de sesgos o errores sistemáticos de los estudios incluidos (Sánchez-Martín et al., 2022). El término de evaluación del riesgo de sesgos y la evaluación de la calidad de un estudio son utilizados como sinónimos en muchas ocasiones. Sin embargo, en su última versión, la *PRISMA 2020* especifica que no son completamente iguales (Page, Moher, et al., 2021). La calidad de los estudios es un concepto amplio y suele incluir constructos que van más allá de aquellos que pueden sesgar los resultados de los estudios (por ejemplo: aspectos éticos, aplicabilidad o explicaciones imprecisas). Por riesgo de sesgos se entienden aquellos aspectos de la metodología (relacionados con el diseño, desarrollo o análisis) que tienen el riesgo potencial de desviar los resultados sistemáticamente de la realidad. La evaluación del riesgo de sesgos centra su atención específicamente en la validez interna o grado de confianza que pone de manifiesto que los resultados del estudio se han obtenido e interpretado de forma adecuada. En la tabla 3 se describen algunas de las herramientas específicamente diseñadas en las últimas décadas, para esta finalidad en función del tipo de diseño utilizado en los estudios.

**Tabla 3**

*Listas de comprobación para la evaluación del riesgo de sesgos*

Descripción	Ejemplos y recomendaciones
<ul style="list-style-type: none"> <li>▪ Diseñadas para evaluar la calidad metodológica de los estudios, generalmente en el contexto de una revisión sistemática/meta-análisis.</li> <li>▪ Redactadas para evaluar, con criterios de corte.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Estudios experimentales controlados y aleatorizados:                             <ul style="list-style-type: none"> <li>- Escala para la evaluación de riesgo de la Colaboración Cochrane (CCRB, <i>Cochrane Collaboration risk bias scale</i>) (Higgins et al., 2019)</li> <li>- Escala PEDro (Physiotherapy Evidence Database) (<a href="https://pedro.org.au/spanish/resources/pedro-scale/">https://pedro.org.au/spanish/resources/pedro-scale/</a>) (<i>Escala PEDro - PEDro</i>, 2016)</li> </ul> </li> <li>▪ Estudios observacionales:                             <ul style="list-style-type: none"> <li>- Casos y Controles y Cohortes: Escala NOS (Newcastle-Otawa scale) (Stang, 2010)</li> <li>- Modificación para estudios transversales (Herzog et al., 2013)</li> </ul> </li> <li>▪ Estudios con metodología cuantitativa, cualitativa y/o mixta: MMAT (<i>Mixed Methods Appraisal tool</i>) (Hong et al., 2018)</li> <li>▪ Estudios de precisión diagnóstica: QUADAS-2 (<i>A Quality Assessment tool for Diagnostic Accuracy Studies</i>) (P. F. Whiting et al., 2011)</li> <li>▪ Revisiones sistemáticas y meta-análisis:                             <ul style="list-style-type: none"> <li>- AMSTAR 2 (<i>A Measurement Tool to Assess systematic Reviews</i>) (Shea et al., 2017)</li> <li>- ROBINS (<i>Risk Of Bias In Systematic reviews</i>) (P. Whiting et al., 2016)</li> </ul> </li> </ul>

La valoración del riesgo de sesgos en el contexto de una RS debe realizarse para cada uno de los estudios incluidos en la misma, y ser explicitada tanto en la interpretación de los resultados como en la discusión (Page et al., 2021). El desarrollo de las revisiones en paraguas (*Umbrella Reviews*) en las que se hace una RS de las revisiones sistemáticas y/o meta-análisis publicados sobre un tema ha ido acompañado del desarrollo de diversas escalas para la evaluación del riesgo de sesgos en este tipo de diseños, como la AMSTAR 2 y la ROBINS (ver Tabla 3).

Greenland & O'Rourke, (2001) y Jüni et al., (1999) recomiendan que, para su interpretación en el contexto de un meta-análisis, no se utilice la puntuación global de cada estudio (calculada como la suma de las puntuaciones de cada uno de los ítems individuales de la escala utilizada) en el análisis del impacto del riesgo de sesgos (Greenland & O'Rourke, 2001; Jüni et al., 1999). Por el contrario, se aconseja analizar el impacto individual de cada uno de los criterios de calidad sobre el tamaño del efecto. El motivo principal es porque suelen evaluar dimensiones de calidad independientes entre sí que podrían tener efectos diferentes en el resultado final del meta-análisis. La explicación de los métodos utilizados para analizar este impacto (análisis de meta-regresión o análisis por subgrupos) exceden los objetivos de este trabajo. Se puede consultar la literatura más especializada (Higgins et al., 2019; Sánchez-Meca, 2022)

Además, los resultados de las escalas de evaluación del riesgo de sesgos de los estudios individuales suelen ser utilizados para jerarquizar los resultados encontrados en función de la calidad de las evidencias constatadas o grado de confianza en los resultados obtenidos. En este contexto, se desarrolló el sistema GRADE (*Grading of Recommendations Assessment, Development and Evaluation*) (Guyatt et al., 2008). En el sistema GRADE se habla de la calidad de las evidencias, ya que no sólo incorpora la valoración del riesgo de sesgo de los estudios individuales, sino también otros factores (por ejemplo: la imprecisión, la inconsistencia o los sesgos de publicación) que podrían alterar la confianza en que el efecto estimado pudiera estar comprometido y ser diferente del efecto real (Balslem et al., 2011). Siguiendo las recomendaciones de sus autores, este sistema no debe ser utilizado para evaluar los estudios individuales,



sino basarse en la agrupación de evidencias sobre un tema concreto. Se establecen cuatro categorías según el nivel de calidad de las evidencias:

- i. Alta (alta confianza en que el verdadero efecto es muy cercano al efecto estimado por el estudio),
- ii. Moderada (confianza moderada en la estimación del efecto),
- iii. Baja (limitaciones en la confianza en el efecto estimado que podría ser sustancialmente diferente del efecto real) y
- iv. Muy Baja (muy poca confianza en la estimación del efecto que, con mucha probabilidad, no coincide con el efecto real).

### Solución al escenario práctico

Tras una lectura atenta del borrador de este documento, le quedan claras las diferencias entre las diferentes listas (las de lectura crítica, las de comprobación para mejorar la calidad y transparencia de las publicaciones y las de evaluación de la calidad metodológica o del riesgo de sesgos). ¡Qué bien le hubiera venido haberlas conocido al empezar su proyecto de investigación! Hubiera tenido una guía magnífica para ayudarle a leer de forma crítica todos los artículos científicos que ha tenido que trabajar, además le hubiera permitido fijarse más o destacar aquellos que tenían una mejor calidad metodológica. Sin embargo, al menos ahora tiene claro como enfocar la redacción de su TD utilizando la STROBE (*Strengthening the Reporting of Observational Studies in Epidemiology*) y que, además de citar su utilización en la apartado de metodología, añadirá la tabla con la lista de comprobación y las páginas en las que reflexiona o incorpora la oportuna explicación sobre los diferentes ítems como anexo del trabajo. Haber descubierto una guía estructurada le soluciona muchas de las dudas iniciales al redactar el trabajo, le da mucha confianza y tranquilidad para afrontar las últimas etapas de su TD.

### Conclusiones

La proliferación de herramientas elaboradas para ayudar a investigadores/as en diferentes contextos ha supuesto un impulso muy importante para mejorar la calidad y transparencia de las publicaciones, contribuyendo a mejorar el conocimiento científico. Sin embargo, parafraseando un dicho conocido: no todas sirven para todo, pero todas sirven para algo.

Para su correcto uso, es necesario familiarizarse con los objetivos por los que fueron publicadas que, básicamente, pueden agruparse en tres grandes categorías: para la lectura crítica de artículos científicos ya publicados, para la mejorar la calidad y transparencia del proceso de publicación y, para evaluar el riesgo de sesgos de los estudios ya publicados.

Evidentemente, existen claras similitudes entre las listas de los tres grandes grupos, aunque tengan finalidades diferentes. Un consejo razonable, sobre todo para los que se están iniciando en el uso de estas listas, podría ser el empezar a familiarizarse con el uso de las listas de lectura crítica. No todos los profesionales de cualquier especialidad querrán o podrán investigar, pero sí que debieran de leer literatura científica de forma crítica como parte de su proceso de formación continuada. Utilizar este tipo de listas expresamente diseñadas para la lectura crítica, facilita enormemente la comprensión y uso del resto de herramientas diseñadas para su utilización en el contexto de la investigación. Además, constituye una ayuda inestimable para mejorar las competencias investigadoras, tanto de los/as investigadores/as senior como noveles. Ya nos hubiera gustado a nosotros/as haber conocido estas herramientas cuando emprendimos el apasionante camino de la investigación, principalmente para el arduo trabajo que supone la difusión de los resultados de los estudios y su publicación en revistas científicas.

**Contribución de cada Autor:** La contribución de cada autor a esta publicación ha sido la siguiente: “conceptualización, FNM y MSM; escritura del manuscrito, MSM, AIPG y MPP; escritura, revisión y edición, MSM, FJIL y FNM; supervisión, FNM”.

**Financiación:** Esta investigación no recibió ninguna financiación externa.

**Conflicto de Intereses:** Las/os autoras/es declaran que no tienen conflicto de intereses.

## Referencias

- American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.). <https://doi.org/10.1037/0000165-000>
- Balakrishnan, A., Puthean, S., Satheesh, G, MK, U., Rashid, M., Nair, S., & Thungai G. (2021). Effectiveness of blended learning in pharmacy education: A systematic review and meta-analysis. *PLoS ONE* 16(6): e0252461. <https://doi.org/10.1371/journal.pone.0252461>
- Balshem, H., Helfand, M., Schünemann, H.J., Oxman, A.D., Kunz, R., Brozek, J., Vist, G.E., Falck-Ytter, Y., Meerpohl, J., Norris, S., & Guyatt, G.H. (2011). GRADE guidelines: 3. Rating the quality of evidence. *Journal of Clinical Epidemiology*, 64(4), 401-406. <https://doi.org/10.1016/j.jclinepi.2010.07.015>
- Baños, R., Sánchez-Martín, M., Navarro-Mateu, F., & Granero-Gallegos, A. (2021). Educación física basada en evidencias: Ventajas e inconvenientes. *Revista Interuniversitaria de Formación del Profesorado*, 96(35.2). <https://doi.org/10.47553/rifop.v97i35.2.88336>
- BestBETs—BETs CA Worksheets. (s.f.). *Best Evidence Topics*. <https://bestbets.org/links/BET-CA-worksheets.php>
- Escala PEDro (2016). *PEDro. Physiotherapy Evidence Database*. <https://pedro.org.au/spanish/resources/pedro-scale/>
- Fichas de Lectura Crítica. (s.f.). *FLC*. <http://www.lecturacritica.com/es/>
- Greenhalgh, T. (2019). *How to Read a Paper: The Basics of Evidence-based Medicine and Healthcare* (6th ed.). Wiley Blackwell.
- Greenhalgh, T., Toon, P., Russell, J., Wong, G., Plumb, L., & Macfarlane, F. (2003). Transferability of principles of evidence based medicine to improve educational quality: Systematic review and case study of an online course in primary health care. *BMJ*, 326(7381), 142-145. <https://doi.org/10.1136/bmj.326.7381.142>
- Greenland, S., & O'Rourke, K. (2001). On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. *Biostatistics*, 2(4), 463-471. <https://doi.org/10.1093/biostatistics/2.4.463>
- Guyatt, G. H., Oxman, A.D., Vist, G.E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., Schünemann, H.J., & GRADE Working Group. (2008). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, 336(7650), 924-926. <https://doi.org/10.1136/bmj.39489.470347.AD>
- Guyatt, G.H., & Rennie, D. (1993). Users' guides to the medical literature. *JAMA*, 270(17), 2096-2097. <https://doi.org/10.1001/jama.1993.03510170086037>
- Herzog, R., Álvarez-Pasquin, M.J., Díaz, C., Del Barrio, J.L., Estrada, J.M., & Gil, Á. (2013). Are healthcare workers' intentions to vaccinate related to their knowledge, beliefs and attitudes? A systematic review. *BMC Public Health*, 13, 154. <https://doi.org/10.1186/1471-2458-13-154>
- Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Page, M., & Welch, V. (2019). *Cochrane Handbook for Systematic Reviews of Interventions* (6.3). John Wiley & Sons. <https://training.cochrane.org/handbook>
- Hong, Q.N., Gonzalez-Reyes, A., & Pluye, P. (2018). Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the Mixed Methods Appraisal Tool (MMAT). *Journal of Evaluation in Clinical Practice*, 24(3), 459-467. <https://doi.org/10.1111/jep.12884>
- JAMA Network* (s.f.). *Users' Guide to the Medical Literature*. <https://jamanetwork.com/collections/44069/users-guide-to-the-medical-literature>
- Jerez Yáñez, O César Orsini Sánchez, C., y Hasbún Held, B. (2016). Atributos de una docencia de calidad en la educación superior: una revisión sistemática. *Estudios Pedagógicos* 42(3), 483-506. <https://doi.org/10.4067/S0718-07052016000400026>
- Jüni, P., Witschi, A., Bloch, R., & Egger, M. (1999). The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*, 282(11), 1054-1060. <https://doi.org/10.1001/jama.282.11.1054>
- Kyriakoulis, K., Patelarou, A., Laliotis, A., Wan, A.C., Matalliotakis, M., Tsiou, C., & Patelarou, E. (2016). Educational strategies for teaching evidence-based practice to undergraduate health students: Systematic review. *Journal of Educational Evaluation for Health Professions*, 13, 34. <https://doi.org/10.3352/jeehp.2016.13.34>



- Maloney, L.M., Marshall, R.T., Werfel, P.A., & Johnson, S.E. (2019). Using a Journal Club Series to Introduce Paramedic Students to Research Fundamentals and Critical Appraisal of Medical Literature. *Prehospital and Disaster Medicine*, 34(4), 449-453. <https://doi.org/10.1017/S1049023X19004618>
- Martínez de Guzmán, M., Soler Ferrería, F.B., Álvarez Muñoz, F.J., Albaladejo Martínez, J.A., Serrano Bueno, F. J., & Navarro-Mateu, F. (2019). *Club de Lectura Crítica del Hospital Psiquiátrico Román Alberca: Una experiencia práctica*. Hospital Psiquiátrico Román Alberca. [https://www.murciasalud.es/publicaciones.php?op=mostrar\\_publicacion&id=2714&idsec=88](https://www.murciasalud.es/publicaciones.php?op=mostrar_publicacion&id=2714&idsec=88)
- Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D.F. (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. Quality of Reporting of Meta-analyses. *The Lancet*, 354(9193), 1896-1900. [https://doi.org/10.1016/s0140-6736\(99\)04149-5](https://doi.org/10.1016/s0140-6736(99)04149-5)
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D.G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6, 1549-1676, e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Morales-Toro, V., Guillén-Riquelme, A., & Quevedo-Blas, R. (2029). Maltrato infantil y trastornos mentales en delincuentes juveniles: Una revisión sistemática. *Revista de Investigación en Educación*, 17(3), 218-238. <https://doi.org/10.35869/reined.v17i3.2157>
- Mulrow, C.D. (1987). The medical review article: State of the science. *Annals of Internal Medicine*, 106(3), 485-488. <https://doi.org/10.7326/0003-4819-106-3-485>
- O'Brien, B.C., Harris, I.B., Beckman, T.J., Reed, D.A., & Cook, D.A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine: Journal of the Association of American Medical Colleges*, 89(9), 1245-1251. <https://doi.org/10.1097/ACM.0000000000000388>
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C..., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 71. <https://doi.org/10.1136/bmj.n71>
- Page, M.J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J..., & McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372(160). <https://doi.org/10.1136/bmj.n160>
- Panic, N., Leoncini, E., de Belvis, G., Ricciardi, W., & Boccia, S. (2013). Evaluation of the endorsement of the preferred reporting items for systematic reviews and meta-analysis (PRISMA) statement on the quality of published systematic review and meta-analyses. *PloS One*, 8(12), e83138. <https://doi.org/10.1371/journal.pone.0083138>
- Patelrou, A.E., Kyriakoulis, K.G., Stamou, A.A., Laliotis, A., Sifaki-Pistolla, D., Matalliotakis, M., Prokopakis, E., & Patelrou, E. (2017). Approaches to teach evidence-based practice among health professionals: An overview of the existing evidence. *Advances in Medical Education and Practice*, 8, 455-464. <https://doi.org/10.2147/AMEP.S134475>
- Pussegoda, K., Turner, L., Garritty, C., Mayhew, A., Skidmore, B., Stevens, A., Boutron, I., Sarkis-Onofre, R., Bjerre, L.M., Hróbjartsson, A., Altman, D.G., & Moher, D. (2017). Systematic review adherence to methodological or reporting quality. *Systematic Reviews*, 6(1), 131. <https://doi.org/10.1186/s13643-017-0527-2>
- RedCaspé (s.f.) *Programa de Habilidades en Lectura Crítica Español*. <https://redcaspe.org/>
- Rethlefsen, M.L., Kirtley, S., Waffenschmidt, S., Ayala, A.P., Moher, D., Page, M.J., Koffel, J.B., & PRISMA-S Group. (2021). PRISMA-S: An extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. *Systematic Reviews*, 10(1), 39. <https://doi.org/10.1186/s13643-020-01542-z>
- Sacks, H.S., Berrier, J., Reitman, D., Ancona-Berk, V.A., & Chalmers, T.C. (1987). Meta-analyses of randomized controlled trials. *The New England Journal of Medicine*, 316(8), 450-455. <https://doi.org/10.1056/NEJM198702193160806>
- Sánchez, M.C., Navarro-Mateu, F., Castillo Bueno, M.D., Menárquez Puche, J.F., & Sánchez Sánchez, J.A. (2007). *Atención sanitaria basada en la evidencia: Su aplicación a la práctica clínica*. Consejería de Sanidad de la Región de Murcia. [https://www.murciasalud.es/publicaciones.php?op=mostrar\\_publicacion&id=103&idsec=88](https://www.murciasalud.es/publicaciones.php?op=mostrar_publicacion&id=103&idsec=88)

- Sánchez-Martín, M., Navarro-Mateu, F., & Sánchez-Meca, J. (2022). Systematic Reviews and Evidence-Based Education [Las Revisiones Sistemáticas y la Educación Basada en Evidencias] *Espiral. Cuadernos del Profesorado*, 15(30), 108-120. <https://doi.org/10.25115/ecp.v15i30.7860>
- Sánchez-Martín, M., Plana, Pedreño, M., Gea, A.I.P., & Navarro-Mateu, F. (2023). And, at first, it was the research question... The PICO, PECO, SPIDER and FINER formats [Y, al principio, fue la pregunta de investigación ... Los formatos PICO, PECO, SPIDER y FINER]. *Espiral. Cuadernos del Profesorado*, 16(32), 126-136. <https://doi.org/10.25115/ecp.v16i32.9102>
- Sánchez-Meca, J. (2022). Revisiones sistemáticas y meta-análisis en Educación: Un tutorial. *Revista Interuniversitaria de Investigación en Tecnología Educativa*, 13, 5-40. <https://doi.org/10.6018/riite.545451>
- Schulz, K.F., Altman, D.G., Moher, D., & CONSORT Group. (2010). CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. *Trials*, 11, 32. <https://doi.org/10.1186/1745-6215-11-32>
- Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L.A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *BMJ*, 349, g7647. <https://doi.org/10.1136/bmj.g7647>
- Shea, B.J., Reeves, B.C., Wells, G., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., & Henry, D.A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*, 358, j4008. <https://doi.org/10.1136/bmj.j4008>
- Stang, A. (2010). Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *European Journal of Epidemiology*, 25(9), 603-605. <https://doi.org/10.1007/s10654-010-9491-z>
- von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C., & Vandenbroucke, J.P. (2007). The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *PLOS. Medicine*, 4, 1549-1676, e296. <https://doi.org/10.1371/journal.pmed.0040296>
- Vu-Ngoc, H., Elawady, S.S., Mehyar, G.M., Abdelhamid, A.H., Mattar, O.M., Halhouli, O., Vuong, N.L., Ali, C.D. M., Hassan, U.H., Kien, N.D., Hirayama, K., & Huy, N.T. (2018). Quality of flow diagram in systematic review and/or meta-analysis. *PloS One*, 13(6), e0195955. <https://doi.org/10.1371/journal.pone.0195955>
- Whiting, P., Savović, J., Higgins, J.P.T., Caldwell, D.M., Reeves, B.C., Shea, B., Davies, P., Kleijnen, J., & Churchill, R. (2016). ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *Journal of Clinical Epidemiology*, 69, 225-234. <https://doi.org/10.1016/j.jclinepi.2015.06.005>
- Whiting, P.F., Rutjes, A.W.S., Westwood, M.E., Mallett, S., Deeks, J.J., Reitsma, J.B., Leeflang, M.M.G., Sterne, J. A.C., Bossuyt, P.M.M., & QUADAS-2 Group. (2011). QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 155(8), 529-536. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>

## Anexo

### Infografía tipos de listas



## LAS LISTAS DE COMPROBACIÓN

### ¿Qué y para qué?

Son **herramientas** diseñadas para la **mejora de la producción científica**, ofreciendo criterios objetivos que permiten realizar una **lectura crítica completa**, redactar artículos científicos o evaluar el riesgo de sesgos, entre otros.

### ¿Cuáles?

Son **numerosas** las listas de comprobación existentes (PRISMA, STROBE, AMSTAR, STAR...), atendiendo a la finalidad para la que fueron concebidas. De ahí, la **importancia de conocerlas y diferenciarlas**, para aplicar cada una de ellas según el tipo de investigación y el momento en el que se encuentra el estudio.

Entre ellas...

#### PARA LA LECTURA CRÍTICA

**¿Para qué?** Para la lectura crítica individual o grupal prestando atención a la metodología.

**Descripción:** criterios agrupados en tres apartados (validez interna, resultados y validez externa).

**Ejemplos:** Caspe, How to read a paper...

#### PARA LA ESCRITURA CIENTÍFICA

**¿Para qué?** Para mejorar la calidad y transparencia de los informes científicos.

**Descripción:** listado exhaustivo que incluye todos los apartados, redactados como recomendaciones.

**Ejemplos:** PRISMA 2020, CONSORT, STROBE, SRQR...

#### PARA EVALUACIÓN DEL RIESGO DE SESGO

**¿Para qué?** Para evaluar el riesgo de sesgos o errores sistemáticos y/o la confianza de los resultados obtenidos.

**Descripción:** diseñadas para la evaluación del riesgo y/o de la calidad metodológica en el contexto de una revisión sistemática.

**Ejemplos:** MMAT, AMSTAR 2, NOS...