Systematic reviews: Key concepts for health professionals

Nadia Sgarbossa[a] , Matías Ibáñez Cobaisse[b] , Gabriel González Cianciulli[b] , Javier Bracchiglione[b, c]* ,
Juan Víctor Ariel Franco[d]

[a] Departamento de Ciencias de la Salud, Universidad Nacional de La Matanza, Buenos Aires, Argentina
[b] Escuela de Medicina, Universidad de Valparaíso, Viña del Mar, Chile
[c] Centro Interdisciplinario de Estudios en Salud (CIESAL), Universidad de Valparaíso, Viña del Mar, Chile
[d] Institut für Allgemeinmedizin, Heinrich-Heine-Universität Düsseldorf, Alemania

Abstract

The exponential growth of currently available evidence has made it necessary to collect, filter, critically appraise, and synthesize biomedical information to keep up to date. In this sense, systematic reviews are a helpful tool and can be reliable sources to assist in evidence-based decision-making. Systematic reviews are secondary research or syntheses of evidence focused on a specific question that -- based on a structured methodology -- make it possible to identify, select, critically appraise, and summarize findings from relevant studies. Systematic reviews have several potential advantages, such as minimizing biases or obtaining more accurate results. The reliability of the evidence presented in systematic reviews is determined, amongst other factors, by the quality of their methodology and the included studies. To conduct a systematic review, a series of steps must be followed: the formulation of a research question using the participants, interventions, comparisons, outcomes (PICO) format; an exhaustive literature search; the selection of relevant studies; the critical appraisal of the data obtained from the included studies; the synthesis of results, often using statistical methods (meta-analysis); and finally, estimating the certainty of the evidence for each outcome. In this methodological note, we will define the basic concepts of systematic reviews, their methods, and their limitations.

MAIN MESSAGES

♦ Systematic reviews are a form of evidence synthesis which can be reliable sources to assist evidence-based decision-making.
♦ Among the advantages of systematic reviews are the minimization of biases and increased accuracy of the results.
♦ They are guided by a specific question and a structured methodology that allows us to: identify, select, critically assess, and qualitatively or quantitatively synthesize (meta-analysis) the findings of individual studies.
♦ Some limitations may include poor quality elaboration, periodic outdating, quality of included studies (garbage-in/garbage-out), redundancy and duplication, the inclusion of problematic (potentially fraudulent) studies, and conflicts of interest.

INTRODUCTION

In the past decades, the amount of scientific evidence in the health field has grown surprisingly, with an estimated 75 clinical trials and 11 systematic reviews published daily [1]. This intensified during the COVID-19 pandemic, when the amount of evidence regarding the subject grew exponentially, having a rate of 1000 publications per week in PubMed [2].

This overwhelming amount of evidence makes it impossible to keep up to date without an adequate synthesis of biomedical information. This problem had already been identified decades ago by Archie Cochrane, who in 1979 raised the challenge of "having an organized and critical review, periodically updated for each specialty or subspecialty, of all relevant randomized controlled studies" [3]. Therefore, it was necessary to develop methods to collect, filter, and synthesize information. High-quality systematic reviews are one of the most reliable and widely used resources to assist evidence-based decision-making. It is essential to clarify that evidence-informed decision-making considers the certainty of the available evidence and the balance between harms and benefits (determined by the patient's values and preferences), acceptability, feasibility, and costs, among other aspects [4].

This article is the eleventh of a series of methodological narrative reviews on general topics in biostatistics and clinical epidemiology, exploring and reviewing in a friendly language, published articles available in the main databases and specialized reference texts. The series is aimed at the education of undergraduate and graduate students. It is carried out by the Chair of Evidence-Based Medicine of the School of Medicine of the University of Valparaiso, Chile, in collaboration with the Research Department of the University Institute of the Hospital Italiano in Buenos Aires, Argentina, the Institute of General Medicine of the Heinrich Heine University in Düsseldorf, Germany, and the Centro Evidencia UC, of the Pontificia Universidad Católica, Chile.

This article explores and presents the most relevant topics on systematic reviews and the steps for conducting this type of study. It aims to be an aid in Spanish for students and health professionals without experience in systematic reviews. Our objectives are to define a systematic review, mention the main differences between this and other types of synthesis, describe its steps, and understand how these can affect the reliability of its findings.

WHAT IS A SYSTEMATIC REVIEW?

The U.S. Institute of Medicine defines systematic review as: "a scientific investigation that focuses on a specific question and uses explicit, pre-specified scientific methods to identify, select, evaluate, and summarize the findings of similar but separate studies" [5].

Systematic reviews follow structured methodologies to minimize the risk of bias in the evidence selection and analysis process, which is an essential difference from narrative reviews [6]. First, they have a protocol previously set up and prospectively registered in specialized databases such as the International Prospective Register of Systematic Reviews, PROSPERO [7–9]. A protocol explains in detail: the theoretical framework, the eligibility criteria and outcomes of interest, the search and selection of articles, the methods for assessing bias risk, and the synthesis and evaluation of evidence certainty [10,11].

WHAT ARE THE MAIN ADVANTAGES OF USING A SYSTEMATIC REVIEW AS A SOURCE OF INFORMATION?

Reduces evidence selection bias or cherry picking

By performing a systematic and comprehensive search of published and unpublished literature, all relevant evidence is selected regardless of its results, avoiding the biased selection of studies. For example, selecting only favorable studies for a given intervention [12].

Greater precision in the results

When a quantitative synthesis is performed, the meta-analysis tool is used, where the results of different individual studies are combined, decreasing random error and increasing statistical power [13]. However, this gain in precision does not guarantee the validity of the results since other factors mentioned in the next point (certainty) must be considered.

It allows considering the certainty of the results

Considering the risk of bias, imprecision, inconsistency, indirect evidence, and publication bias, employing the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) methodology, it is possible to estimate how much reliability we can assign to the results obtained [4,14].

It is important to note that other forms of systematic evidence synthesis are discussed in more detail in additional *Medwave* methodological notes [6]. These include overview reviews and evidence mappings [15], rapid reviews [16], living reviews (continuously updated) [17], network meta-analyses [18], and reviews of systematic reviews (or overviews).

Cochrane was one of the first organizations to produce and develop systematic reviews, aiming to provide summarized and reliable evidence that allows approaching an answer to the questions that constantly arise in healthcare field decision-making. Today, the Cochrane Library has the *Cochrane Database of Systematic* Reviews, a database with more than 8800 reviews, which comply with a rigorous methodology, adjusted over the years to ensure compliance with the updated quality criteria (described in the Cochrane Handbook and the Methodological Expectations of Cochrane Intervention Reviews, MECIR standards) [13,19]. Cochrane reviews contain *Summary of Findings* (SoF) tables that follow the GRADE methodology, which is easy to read and understand [14].

Other organizations such as the Campbell Collaboration (www.campbellcollaboration.org/), the Joanna Briggs Institute (jbi.global/), and the BEME Collaboration (www.bemecollaboration.org/), among others, publish systematic reviews, some of which are of a specific topic (e.g., education in the case of BEME). In addition to the Cochrane Handbook, other methodological guidelines for systematic reviews include the Centre for Reviews and Dissemination manual, the Joanna Briggs Institute manual, and guidelines from the Agency for Health Research and Quality (AHRQ) and the U.S. Institute of Medicine (IOM) [5,20–22].

STEPS TO CONDUCTING A SYSTEMATIC REVIEW

1. FORMULATION OF THE QUESTION, ELIGIBILITY CRITERIA, AND OUTCOMES OF INTEREST

The first step is to define the key elements of the research question, which is usually done by structuring the question in the PICO/PECO format (population, intervention/exposure, comparison, and outcome) [23,24] (Table 1). This format may vary depending on whether the question is therapeutic, prognostic, etiologic, diagnostic, or other (Table 2). A variant of the PICO question could be "PICOS" or "PICOTS", where the T indicates the follow-up time, and the "S" would indicate the type of study (study design) or study context (setting). For didactic purposes, we will take a therapeutic question as an example for the development of this article.

After structuring the basic elements of the PICO question, the eligibility criteria and outcomes of interest must be established in detail. The eligibility criteria are based mainly on the elements of the PICO question. Following the above, for the PICO question: "in women at risk of preterm delivery, the administration of prenatal corticosteroids decreases fetal, neonatal, and maternal morbidity and mortality?" the eligibility criteria should consider the following questions:

a) Patient: women at risk of preterm delivery. What age range will be included? How is the preterm delivery risk defined? What definition of preterm delivery will we consider?
b) Intervention: prenatal corticosteroids. Are all drugs in this class included? In what dosage? By which administration route?
c) Type of study: Which study design will be included? Typically, randomized clinical trials are included in therapeutic questions, but observational studies could also be considered; the latter is more important in diagnostic or prognostic questions.

Finally, it is crucial to define the outcomes of interest. For example, the previous PICO question could include perinatal mortality, maternal mortality, risk of operative delivery, and labor pain.

Table 1. Question structure in the PICO/PECO format.

Patient or population	The subject of interest, to whom the intervention is applied, and in whom the intervention results are sought. It should describe the most important aspects of the specific clinical scenario, excluding redundant or irrelevant data.
Intervention or Exposure	It can be a pharmacological therapy, surgical procedure, alternative medicine, or diagnostic test.
Comparison	The standard course of action for the clinical scenario that generates the question. Sometimes, it may be a dispensable element, depending on the nature of the question.
Outcome	It is the targeted situation, it can be a prognostic, diagnostic, or therapeutic outcome, the latter being the most common.

Source: prepared by the authors based on [23,24].

Table 2. Examples of therapeutic, diagnostic, and prognostic questions in PICO format.

Therapeutic question	Population: women at risk of preterm delivery. Intervention: prenatal corticosteroids. Comparison: placebo Outcome: fetal, neonatal, and maternal morbidity and mortality. In women at risk of preterm delivery, does prenatal administration of corticosteroids decrease fetal, neonatal, and maternal morbidity and mortality? [25].
Diagnostic accuracy question	Patient: women of childbearing age with suspected endometriosis. Intervention (diagnostic): urinary markers. Comparison: laparoscopic surgery (reference test or gold standard). Outcome: diagnostic accuracy (sensitivity, specificity, etc.) for the diagnosis of endometriosis. In women of childbearing age, are urinary markers more useful than laparoscopic surgery for the diagnosis of endometriosis? [26]
Prognostic Question	Patient: general population. Exposure: intermediate hyperglycemia. Comparison: normoglycemia. Outcome: risk of developing diabetes mellitus II. In the general population, is intermediate hyperglycemia associated with an increased risk of developing diabetes mellitus II compared to normoglycemia? [27]

Outcomes can be classified according to their relevance as: "critical," "important," or "unimportant" [28]. Clinically critical, or important outcomes are those that, if modified, would change the patient's acceptance of the intervention. For example, a critical outcome would be perinatal mortality, and an unimportant outcome would be: pain at the site of corticosteroid injection. Reviews may also define "primary" and "secondary" outcomes, usually restricting more advanced analytical tests (subgroup and sensitivity analyses) to the first.

Outcomes can also be classified by whether they "directly" or "indirectly" measure the phenomenon under study. Surrogate or intermediate outcomes are understood as those that do not directly measure the phenomenon under study and therefore substitute a clinically relevant outcome [28]. For example, neonatal oxygen saturation is a surrogate or intermediate outcome to the clinical outcome of respiratory stress and perinatal mortality.

The Core Outcome Measures in Effectiveness Trials (COMET) initiative generated standardized collections of outcomes, called Core Outcome Sets, which correspond to the minimum outcomes that should be measured and reported in all condition-specific clinical trials [29]. Just as these outcomes are considered critical and relevant for decision-making in clinical trials, systematic reviews should use the same outcomes to synthesize evidence from such trials [30].

2. BIBLIOGRAPHIC SEARCH AND STUDY SELECTION

Once the PICO question and eligibility criteria have been formulated and structured, the next step is establishing the databases in which the literature search will be conducted. This is a key moment since an inadequate search can lead to a ripple effect in which the conclusions obtained in a review are inaccurate and unreliable. It is essential to define the following:

a) Information sources: systematic reviews usually use diverse and complementary sources for the search, including biomedical databases such as CENTRAL, MEDLINE/PubMed, EMBASE, CINAHL, PsycINFO, Scopus, and Web Of Science, among others [31]. Each has advanced specifications detailed in another methodological note of this series [32]. In addition, systematic reviews generally include gray literature search strategies using specific databases, reference searching, and expert consultation, among others [33].

b) The terms and structure of the search strategy: when structuring a question, we must identify the main concepts to design a search strategy that includes the terms relevant to our search and, at the same time, excludes those that may hinder the search. It is here where it will be necessary to use different tools available within each search platform for each database; we can mention the Boolean operators (OR, AND, and NOT) and the filters available within each platform (filter by year, type of study, language, and others) [31].

Another article in this methodological series details how to perform an exhaustive and structured bibliographic search [32].

Once the search has been performed, different software such as Endnote, Covidence, Rayyan, and EPPI-Reviewer can be used for: bibliographic management, data deduplication, and selection process, in which authors must cross-check each article found against the pre-specified eligibility criteria. The process is usually carried out in two consecutive stages. Firstly, by evaluating the title, abstract, and then the full text of each article. Once the included studies have been selected, we proceed to extract the characteristics of the studies and the results of interest. Usually, this entire process is performed by two or more authors, both independently and blinded to the work of the other reviewer, to contrast and reduce the risk of errors during the process. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines provide a detailed methodology on how to report this process, including a type of flow chart, which makes it possible to graphically express the selection of articles (example in Figure 1) [34].

Figure 1. Example of PRISMA flow chart.
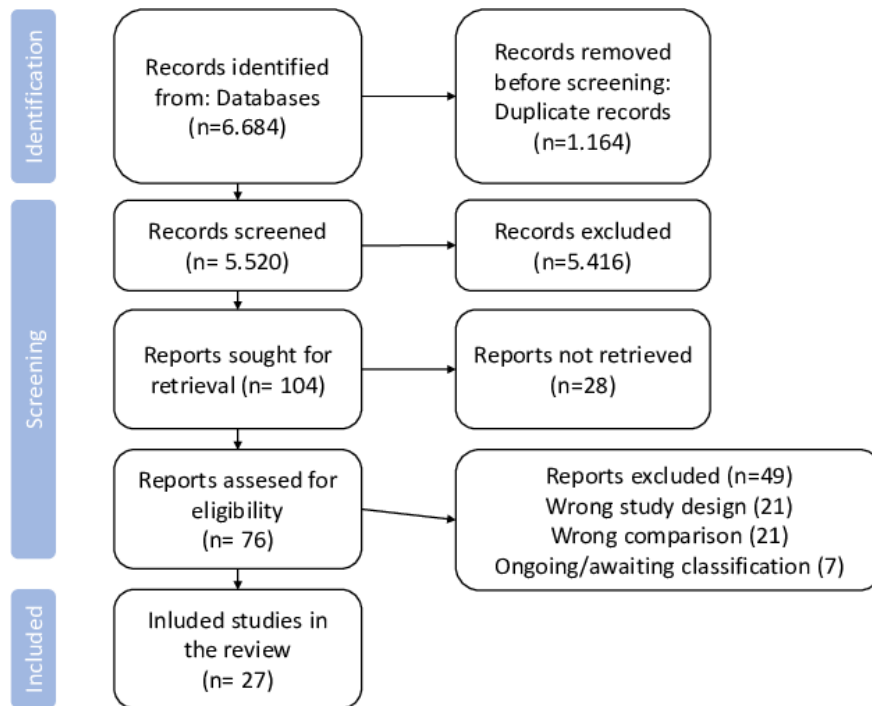
Source: Prepared by the authors of this study.

3. CRITICAL ASSESSMENT AND RESULT SYNTHESIS

Why is it important to critically appraise the included studies?

The certainty of the findings presented in a systematic review will depend on the biases inherent to the studies included in the review. Bias is a systematic error or deviation of the results from the truth, which is not determined by chance [35]. Although a low risk of bias is often considered equivalent to a high-quality study, the concept of quality is broader, and bias is a dimension of it. The standardized assessment of biases present in the studies allows us to have a notion of the reliability of the results, conceptualized and other evaluations, in the certainty of evidence using the GRADE methodology [14]. In turn, a correct evaluation of possible biases requires adequate reporting by the authors of the primary studies, following standardized reporting guidelines. For example, clinical trials should follow the Consolidated Standards of Reporting Trials, CONSORT, which provides a basic list of elements to report [36]. However, it should be noted that CONSORT is not a guideline for assessing the methodological quality or clinical trial biases but rather their quality of reporting.

HOW TO ASSESS THE PRESENCE OF BIAS IN THE INCLUDED STUDIES IN SYSTEMATIC REVIEWS?

Scales or checklists

One of the most widely used methods is scales or checklists. Scales score various quality components and then combine them to obtain a final score, while checklists consist of items or questions that are filled in as the checklist is completed. The limitation of these methods is that they assign a score or equivalent value to different aspects related to a study's methodological quality, which may not be correct. For example, taking two methodological aspects of a clinical trial, a problem in randomization may be a major flaw compared to the lack of blinding of a study that evaluates objective outcomes. One of the most widely used scales for clinical trials, the Jadad scale, has such limitations [37].

Domain-based assessment

They consist of critical assessments made separately for different domains related to independent methodological aspects. Following the above example for clinical trials, methodological problems related to randomization are evaluated and interpreted individually from those related to blinding. Based on this, one can think independently about how these potential biases may affect the results. Table 3 summarizes the Cochrane RoB 2 tool for clinical trials. For observational studies, there are several tools, one of the most widely used being the Newcastle-Ottawa Scale. However, it has been recently replaced by ROBINS-I [39]. There are also several tools for diagnostic studies and study evaluation packages, such as those of the Joanna Briggs Institute and the United States Institute of Health (NIH) [40–43].

Table 3. Cochrane Clinical Trials Tool (RoB 2).

Domain	What does it evaluate?
Domain 1. Biases due to the randomization process	How were patients randomized and assigned?
Domain 2. Biases due to deviation from planned interventions	Were participants and staff blinded? Were patients in the intervention and control groups treated and analyzed as planned?
Domain 3. Biases due to loss of outcome data	Was there a data loss in the follow-up of patients that could have affected the interpretation of the results?
Domain 4. Biases in outcome measurement	Were legitimate tools used to measure outcomes, and were the outcome assessors blinded?
Domain 5. Biases in the selection of reported outcomes	Were all outcomes reported following a pre-specified protocol?
Overall risk of bias	What is the overall appraisal of the results considering the assessment of each of the above domains? (Low risk of bias, some concerns about bias, high risk of bias)

Source: Adapted by the authors from [38].

4. RESULTS SUMMARY

Once the results of each study are extracted and a critical appraisal of potential biases is made, the data related to outcomes can be compared and synthesized based on statistical analyses. The most used method is meta-analysis, which allows the results of two or more individual studies to be combined. The fundamental objective of a meta-analysis is to increase the power and precision of the results, generating a point estimator where the results of the individual studies included are weighted. In some cases, it also allows us to answer questions not raised in individual studies (e.g., by subgroup analysis) and to explore inconsistencies (heterogeneity) in the findings. Various statistical programs are available to carry them out (e.g., RevMan, Stata, and R).

Although every meta-analysis should be performed in the context of a systematic review, not every systematic review has a meta-analysis since there must be minimum comparability between populations, interventions, comparisons, and outcomes of the included studies to perform a statistical combination of them. Furthermore, if no studies on the PICO question are found in the systematic review or only one study is found, it is impossible to perform the statistical aggregation of two or more studies. Other reasons for which a narrative synthesis of the numerical findings from a systematic review is often preferred include:

a) Very broad questions that incorporate scattered and non-comparable evidence.
b) Lack of statistical information from individual studies that prevent conducting a meta-analysis.
c) Differences in population, intervention, and comparison that prevent the comparison of study results (clinical heterogeneity).
d) The presence of important statistical heterogeneity in the meta-analysis that potentially invalidates its results [44].

The forest plot corresponds to the graphical representation of a meta-analysis and has different information, detailed in Figure 2. When interpreting a forest plot, it is important to identify the following key elements:

a) Heterogeneity: the degree of similarity between the results of each study, assessed visually or with statistical measures.
b) Meta-analysis: the summary measure (represented by a "diamond") indicating the pooled effect and its confidence interval. These confidence intervals are relevant when interpreting them concerning the clinical significance thresholds.

Similar to other statistical tests, meta-analyses arise from assumptions that allow the combination of results. Founded on these assumptions, there are two types of meta-analyses:

a) Fixed-effect meta-analysis: this model assumes that there is a single effect that the different studies included in the synthesis estimate with some variability (random error). This model does not consider between-study variability as a source of heterogeneity, so it is used under the assumption of homogeneity between studies.
b) Random-effects meta-analysis: this model does not assume a single effect but rather multiple possible effects estimated by the included studies. In addition to integrating the variability due to random error, it considers the variability between studies and is regarded as the most conservative model since it does not assume homogeneity between studies and, in the presence of heterogeneity, produces less precise results [46,47].

These meta-analyses represented in forest plots are called direct or pairwise meta-analyses. There is a different approach called network meta-analysis, whose methodology allows using evidence from direct and indirect comparisons. The details of this type of study are described in detail in another article of this methodological series [18].

The funnel plot is an additional graphical representation of the effect estimates of individual studies, in contrast to their precision, so that at the base of the plot are the less precise studies (usually more dispersed around the meta-analysis), and at the top are the more precise studies, usually concentrated around the meta-analysis estimate (see Figure 3). When there is an asymmetric dispersion of small studies (small study effect), publication bias may be suspected. However, there are alternative explanations for this phenomenon, and its interpretation should be cautious.
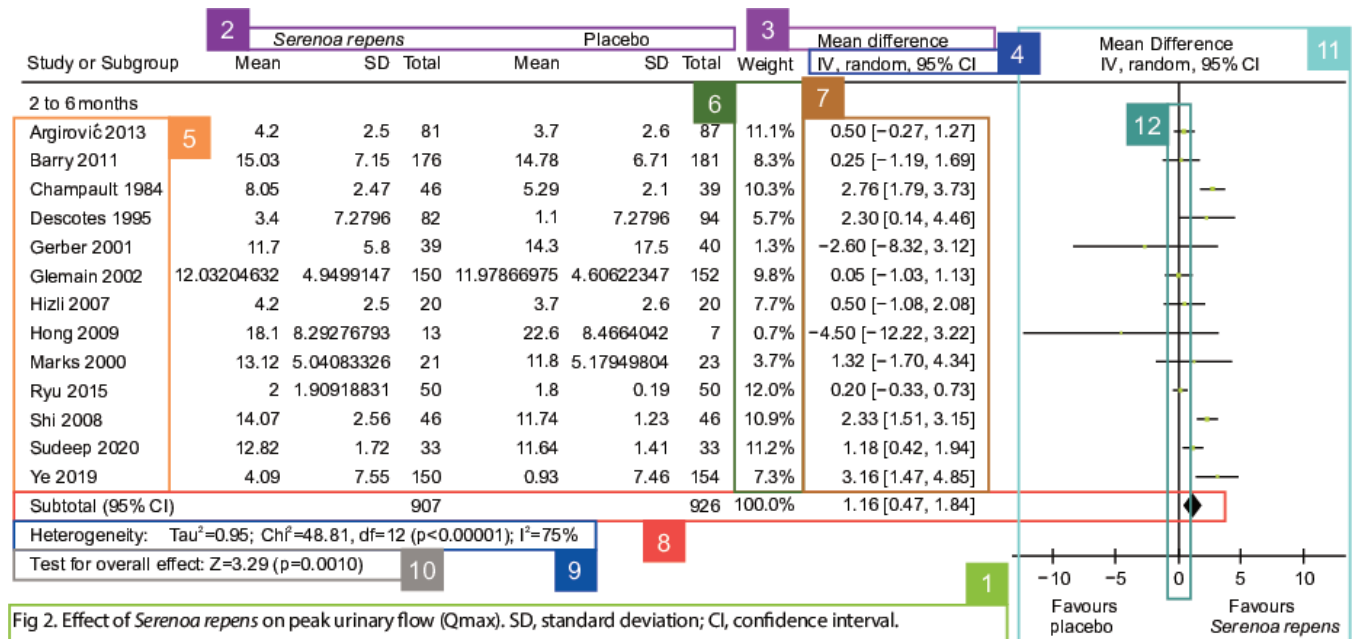
Figure 2. Example of forest plot and its components.

| Study or Subgroup	Serenoa repens Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean difference IV, random, 95% CI
2 to 6 months								
Argirović 2013	4.2	2.5	81	3.7	2.6	87	11.1%	0.50 [−0.27, 1.27]
Barry 2011	15.03	7.15	176	14.78	6.71	181	8.3%	0.25 [−1.19, 1.69]
Champault 1984	8.05	2.47	46	5.29	2.1	39	10.3%	2.76 [1.79, 3.73]
Descotes 1995	3.4	7.2796	82	1.1	7.2796	94	5.7%	2.30 [0.14, 4.46]
Gerber 2001	11.7	5.8	39	14.3	17.5	40	1.3%	−2.60 [−8.32, 3.12]
Glemain 2002	12.03204632	4.9499147	150	11.97866975	4.60622347	152	9.8%	0.05 [−1.03, 1.13]
Hizli 2007	4.2	2.5	20	3.7	2.6	20	7.7%	0.50 [−1.08, 2.08]
Hong 2009	18.1	8.29276793	13	22.6	8.4664042	7	0.7%	−4.50 [−12.22, 3.22]
Marks 2000	13.12	5.04083326	21	11.8	5.17949804	23	3.7%	1.32 [−1.70, 4.34]
Ryu 2015	2	1.90918831	50	1.8	0.19	50	12.0%	0.20 [−0.33, 0.73]
Shi 2008	14.07	2.56	46	11.74	1.23	46	10.9%	2.33 [1.51, 3.15]
Sudeep 2020	12.82	1.72	33	11.64	1.41	33	11.2%	1.18 [0.42, 1.94]
Ye 2019	4.09	7.55	150	0.93	7.46	154	7.3%	3.16 [1.47, 4.85]
Subtotal (95% CI)			907			926	100.0%	1.16 [0.47, 1.84]
Heterogeneity: Tau²=0.95; Chi²=48.81, df=12 (p<0.00001); I²=75%								
Test for overall effect: Z=3.29 (p=0.0010)								

Fig 2. Effect of *Serenoa repens* on peak urinary flow (Qmax). SD, standard deviation; CI, confidence interval.

1. Title
2. Intervention and comparison groups
3. Effect measurement
4. Type of meta-analysis
5. Primary studies included
6. Weight of each study
7. Average effect and confidence interval
8. Meta-analysis
9. Measurement of statistical heterogeneity
10. Test for overall effect
11. Forest plot
12. Line of no effect

Notes: Red dots of the smaller studies are asymmetrically distributed to the left of the meta-analysis (vertical line). In contrast, the more precise studies are at the top of the funnel, close to the meta-analysis (red dot at the upper right).

Figure 3. Example of an asymmetric funnel chart.
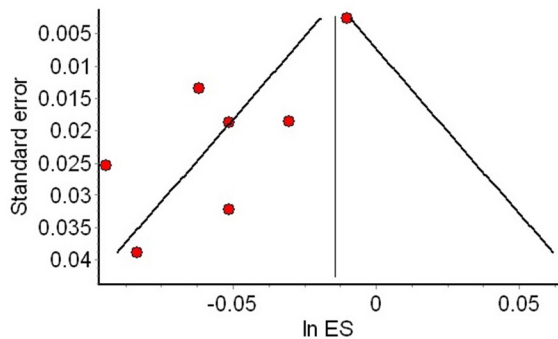
5. FORMULATING CONCLUSIONS IN SYSTEMATIC REVIEWS

Considering the multiple steps of a systematic review, the strengths and weaknesses of the synthesized body of evidence should be identified when formulating conclusions. Due to multiple systems for classifying these strengths and weaknesses, the GRADE methodology was developed to standardize this assessment. To this end, GRADE categorizes the certainty of the evidence into four levels (high, moderate, low, and very low). This certainty is defined by the methodological design of the primary studies included and by the factors that can increase or decrease the level of certainty, among which are:

a)   Risk of bias.
b)   Inconsistency.
c)   Indirect evidence.
d)   Imprecision.
e)   Publication bias.

This methodology has been detailed in another article of this methodological series [14]. Although the GRADE methodology is used to determine the certainty of evidence and generate recommendations, it is important to note that systematic

reviews alone are insufficient to generate a health recommendation [4].

SOME LIMITATIONS AND PROBLEMS OF SYSTEMATIC REVIEWS

Despite their many advantages for evidence-based decision-making, systematic reviews, like any other scientific study, are not free of problems:

a) Low methodological quality: multiple scales and lists allow us to evaluate the methodological quality of systematic reviews. Most of these focus on assessing transparency, internal and external validity, and risk of bias, among other aspects. Systematic reviews may have the quality of their evidence affected due to multiple shortcomings in their structure and analysis, such as a limited search strategy, unsatisfactory assessment of the risk of bias, or unjustified exclusion of primary studies. Therefore, it is essential to analyze the systematic reviews available critically.

b) Redundancy and duplicity: there are overlapping systematic reviews, which do not provide new information since they use a large percentage of primary studies previously used in other systematic reviews as a source.

c) Garbage-in/Garbage-out: a systematic review of good methodological quality is not exempt from including poor-quality primary studies (garbage-in). This means that the synthesis result of poor-quality studies will provide results of low certainty (garbage out) [48].

d) Problematic studies: reviews may be subject to other problems, such as the inclusion of fraudulent studies and conflicts of interest of the authors that could bias the findings [49]. As with the other issues to which systematic reviews are subject, strategies have emerged to mitigate them, such as stricter editorial policies regarding conflicts of interest in authors of systematic reviews [50,51].

e) Lack of updating: It usually takes a long time to conduct a systematic review, making it difficult to include recent primary studies [52]. Some of the methodological advances that could counteract this phenomenon include: living systematic reviews, which have continuous periodic updates and are described thoroughly in another article of this methodological series [17].

f) Conflicts of interest: in the same way that the presence of conflicts of interest in the authors or funders of a primary study can alter and bias the results, systematic reviews are also susceptible to this type of influence [53]. This is why Cochrane has one of the strictest conflict of interest policies compared to other journals, which guarantees the independence of study evaluations [51].

As users of systematic reviews, it is important to use tools to identify some of these problems. The best-known tools are PRISMA [34] for report quality, AMSTAR 2 [9] for methodological quality, and ROBIS [54] for risk of bias assessment.

CONCLUSIONS

Systematic reviews are a type of research design that identifies, compiles, critically analyzes, and synthesizes the available evidence, facilitating the use of evidence for decision-making by health professionals. Additionally, having a protocol based on a structured methodology for its elaboration allows the reduction of biases during the synthesis of information. Therefore, knowing systematic reviews' structure, steps, and limitations will enable us to read them critically and comprehensively. This is where tools become essential to identify and assess these problems.

Notes

References

1. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? PLoS Med. 2010;7. https://doi.org/10.1371/journal.pmed.1000326

2. Torres-Salinas D. Ritmo de crecimiento diario de la producción científica sobre Covid-19. Análisis en bases de datos y repositorios en acceso abierto. EPI. 2020;29. https://doi.org/10.3145/epi.2020.mar.15

3. Cochrane AL. Archie Cochrane in his own words. Selections arranged from his 1972 introduction to "Effectiveness and Efficiency: Random Reflections on the Health Services" 1972. Control Clin Trials. 1989;10: 428–33. https://doi.org/10.1016/0197-2456(89)90008-1

4. Quilodrán C, Kirmayr M, Valente B, et al. Metodología GRADE, parte 2: de la evidencia a la decisión—esquematizando la toma de decisiones en salud. Medwave. 2021;21. https://doi.org/10.5867/medwave.2021.04.8182

5. Eden J, Levit L, Berg A, Morton S, Committee on Standards for Systematic Reviews of Comparative Effectiveness Research, Board on Health Care Services, et al. Committee on Standards for Systematic Reviews of Comparative Effectiveness Research. Finding What Works in Health Care: Standards for Systematic Reviews. Washington, D.C: National Academies Press (US) 2014; 2011. https://www.nap.edu/catalog/13059 https://doi.org/10.17226/13059

6. Franco JVA, Arancibia M, Simancas-Racines D, Madrid E. Syntheses of biomedical information: narrative reviews, systematic reviews and emerging formats. Medwave. 2018;18. https://doi.org/10.5867/medwave.2018.07.7354

7. PROSPERO. https://www.crd.york.ac.uk/prospero/

8. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. Syst Rev. 2012;1: 1–9. https://doi.org/10.1186/2046-4053-1-2

9. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, 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. https://doi.org/10.1136/bmj.j4008

10. Lasserson TJ, Thomas J, Higgins JPT, et al. Chapter 1: Starting a review. In: Higgins JPT, Thomas J, Chandler J, editors. Cochrane Handbook for Systematic Reviews of Interventions version 63. Cochrane; 2022. https://training.cochrane.org/handbook/current/chapter-01#section-1-5

11. Cumpston M, Chapter CJ, et al. Planning a Cochrane Review. In: Higgins JPT, Thomas J, Chandler J, editors. Cochrane Handbook for Systematic Reviews of Interventions version 63. Cochrane; 2022. https://training.cochrane.org/handbook/current/chapter-ii#section-ii-1-4

12. Kopitowski K. La falacia de la evidencia incompleta o "cherry picking." Evid actual pract ambul. 2016;19. https://doi.org/10.51987/evidencia.v19i2.6440

13. Higgins JPT, Thomas J, Chandler J, et al. Cochrane Handbook for Systematic Reviews of Interventions. Cochrane; 2022. https://training.cochrane.org/handbook/current

14. Kirmayr M, Quilodrán C, Valente B, Loezar C, Garegnani L, Franco JVA. The GRADE approach, Part 1: how to assess the certainty of the evidence. Medwave. 2021;21. https://doi.org/10.5867/medwave.2021.02.8109

15. Verdejo C, Tapia-Benavente L, Schuller-Martínez B, Vergara-Merino L, Vargas-Peirano M, Silva-Dreyer AM. What you need to know about scoping reviews. Medwave. 2021;21. https://doi.org/10.5867/medwave.2021.02.8144

16. Tapia-Benavente L, Vergara-Merino L, Garegnani LI, Ortiz-Muñoz L, Loézar Hernández C, Vargas-Peirano M. Rapid reviews: definitions and uses. Medwave. 2021;21. https://doi.org/10.5867/medwave.2021.01.8090

17. Vergara-Merino L, Verdejo C, Carrasco C, Vargas-Peirano M. Living systematic review: new inputs and challenges. Medwave. 2020;20. https://doi.org/10.5867/medwave.2020.11.8092

18. González-Xuriguera CG, Vergara-Merino L, Garegnani L, Ortiz-Muñoz L, Meza N. Introduction to network meta-analysis for evidence synthesis. Medwave. 2021;21. https://doi.org/10.5867/medwave.2021.06.8315

19. Higgins JPT, Lasserson T, Chandler J, et al. Methodological Expectations of Cochrane Intervention Reviews. Cochrane: London; 2022.

20. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care. 2009. https://books.google.com/books/about/Systematic_Reviews.html?hl=&id=tUd2SwAACAAJ

21. Aromataris E, Munn Z. JBI Manual for Evidence Synthesis. 2020. https://wiki.jbi.global/display/MANUAL https://doi.org/10.46658/JBIMES-20-01

22. Agency for Healthcare Research and Quality. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. 2008. https://books.google.com/books/about/Methods_Guide_for_Effectiveness_and_Comp.html?hl=&id=lC0vnQAACAAJ

23. Brown D. A Review of the PubMed PICO Tool: Using Evidence-Based Practice in Health Education. Health Promot Pract. 2020;21: 496–498. https://doi.org/10.1177/1524839919893361

24. Coleman J. Help and Support: Systematic Reviews - Research Guide: Home. 2019. https://libguides.murdoch.edu.au/systematic/home

25. McGoldrick E, Stewart F, Parker R, Dalziel SR. Antenatal corticosteroids for accelerating fetal lung maturation for women at risk of preterm birth. Cochrane Database Syst Rev. 2020;12. https://doi.org/10.1002/14651858.CD004454.pub4

26. Liu E, Nisenblat V, Farquhar C, Fraser I, Bossuyt PMM, Johnson N, et al. Urinary biomarkers for the non-invasive diagnosis of endometriosis. Cochrane Database Syst Rev. 2015. https://doi.org/10.1002/14651858.CD012019

27. Richter B, Hemmingsen B, Metzendorf M-I, Takwoingi Y. Development of type 2 diabetes mellitus in people with intermediate hyperglycaemia. Cochrane Database Syst Rev. 2018;10. https://doi.org/10.1002/14651858.CD012661.pub2

28. Neumann I, Claro G-A JC. Outcomes intermedios. Rev méd Chile. 2012;140: 113–116. https://doi.org/10.4067/S0034-98872012000100017

29. Prinsen CAC, Vohra S, Rose MR, King-Jones S, Ishaque S, Bhaloo Z, et al. Core Outcome Measures in Effectiveness Trials (COMET) initiative: protocol for an international Delphi study to achieve consensus on how to select outcome measurement instruments for outcomes included in a "core outcome set." Trials. 2014;15: 1–7. https://doi.org/10.1186/1745-6215-15-247

30. Williamson P, Clarke M. The COMET (Core Outcome Measures in Effectiveness Trials) Initiative: Its Role in Improving Cochrane Reviews. Cochrane Database Syst Rev. 2012. https://doi.org/10.1002/14651858.ED000041

31. Bramer WM, de Jonge GB, Rethlefsen ML, Mast F, Kleijnen J. A systematic approach to searching: an efficient and complete method to develop literature searches. J Med Libr Assoc. 2018;106: 531–541. https://doi.org/10.5195/jmla.2018.283

32. Trivisonno LF, Escobar Liquitay C, Vergara-Merino L, et al. Conceptos clave para la búsqueda de evidencia: una introducción para profesionales de la salud. Medwave. 2022;22. https://doi.org/10.5867/medwave.2022.01.002512

33. Paez A. Grey literature: An important resource in systematic reviews. J Evid Based Med. 2017. https://doi.org/10.1111/jebm.12265

34. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated

guideline for reporting systematic reviews. BMJ. 2021;372: 71. https://doi.org/10.1136/bmj.n71

35. Boutron I, Page MJ, Higgins JPT, et al. Chapter 7: Considering bias and conflicts of interest among the included studies. In: Higgins JPT, Thomas J, Chandler J, editors. Cochrane Handbook for Systematic Reviews of Interventions version. Cochrane; 2022. p. 3. https://training.cochrane.org/handbook/current/chapter-07

36. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med. 2010;152: 726–32. https://doi.org/10.7326/0003-4819-152-11-201006010-00232

37. Lundh A, Gøtzsche PC. Recommendations by Cochrane Review Groups for assessment of the risk of bias in studies. BMC Med Res Methodol. 2008;8: 1–9. https://doi.org/10.1186/1471-2288-8-22

38. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366. https://doi.org/10.1136/bmj.l4898

39. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355. https://doi.org/10.1136/bmj.i4919

40. Risk of Bias 2 (RoB 2) tool. https://methods.cochrane.org/risk-bias-2

41. University of Bristol. QUADAS-2. https://www.bristol.ac.uk/population-health-sciences/projects/quadas/quadas-2/

42. Appendix 10.1 JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses. https://jbi-global-wiki.refined.site/space/MANUAL/4687059/Appendix+10.1+JBI+Critical+Appraisal+Checklist+for+Systematic+Reviews+and+Research+Syntheses

43. Study Quality Assessment Tools. NHLBI, NIH; https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools

44. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ 2020;368. https://doi.org/10.1136/bmj.l6890

45. Trivisonno LF, Sgarbossa N, Alvez GA, et al. Serenoa repens for the treatment of lower urinary tract symptoms due to benign prostatic enlargement: A systematic review and meta-analysis. Investigative and Clinical Urology 2021;62:520–34. https://doi.org/10.4111/icu.20210254

46. Deeks JJ, Higgins JPT, Altman DG, et al. Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, editors. Cochrane Handbook for Systematic Reviews of Interventions version 63. Cochrane; 2022. https://training.cochrane.org/handbook/current/chapter-10

47. Haidich AB. Meta-analysis in medical research. Hippokratia. 2010;29. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3049418/ (accessed 4 Feb 2022

48. Egger M, Smith GD, Sterne JA. Uses and abuses of meta-analysis. Clin Med (Lond). 2001;1: 478–84. https://doi.org/10.7861/clinmedicine.1-6-478

49. Bhar RH, John D. Conflict of interest in systematic reviews and its implications for public health policy. Indian J Med Ethics. 2019;4 (NS): 288–293. https://doi.org/10.20529/IJME.2019.058

50. Tovey D, Macbeth F, Webster A, Foxlee R. Cochrane's plans to update its COI policy: truth and fiction. Cochrane Database Syst Rev. 2018;11. https://doi.org/10.1002/14651858.ED000131

51. Macbeth F, Webster A, Foxlee R, et al. Cochrane Conflict of Interest Policy for Cochrane Library Content. Cochrane; 2020. https://training.cochrane.org/online-learning/editorial-policies/coi-policy/coi-policy-cochrane-library

52. Borah R, Brown AW, Capers PL, Kaiser KA. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. BMJ Open. 2017;7. https://doi.org/10.1136/bmjopen-2016-012545

53. Hansen C, Lundh A, Rasmussen K, Hróbjartsson A. Financial conflicts of interest in systematic reviews: associations with results, conclusions, and methodological quality. Cochrane Database Syst Rev. 2019;8. https://doi.org/10.1002/14651858.MR000047.pub2

54. Whiting P, Savović J, Higgins JPT, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol. 2016;69: 225–34. https://doi.org/10.1016/j.jclinepi.2015.06.005

Revisiones sistemáticas: conceptos clave para profesionales de la salud

Resumen

El crecimiento exponencial de evidencia disponible actualmente ha hecho necesario recopilar, filtrar, valorar críticamente y sintetizar la información biomédica para mantenerse actualizado. En este sentido, las revisiones sistemáticas constituyen una herramienta útil y pueden ser fuentes confiables para asistir a la toma de decisiones basadas en evidencia. Definimos como revisiones sistemáticas a aquellas investigaciones secundarias o síntesis de evidencia focalizadas en una pregunta específica que, a partir de una metodología estructurada, permiten identificar, seleccionar, valorar críticamente y resumir los hallazgos de estudios relevantes. Las revisiones sistemáticas presentan varias ventajas potenciales, tales como la minimización de sesgos o la obtención de resultados de mayor precisión. La confiabilidad de la evidencia presentada en las revisiones sistemáticas está determinada, entre otros factores, por su calidad metodológica, pero también por la calidad de los estudios incluidos. Para realizar una revisión sistemática, se debe seguir una serie de pasos que incluyen la formulación de una pregunta de investigación a partir del formato participants, interventions, comparisons, outcomes, PICO; una búsqueda bibliográfica exhaustiva; la selección de los estudios relevantes; la valoración crítica de los datos obtenidos a partir de los estudios incluidos; la síntesis de resultados, a menudo mediante métodos estadísticos (metanálisis); y finalmente una estimación de la certeza de evidencia para cada desenlace. En esta nota metodológica definiremos los conceptos básicos sobre revisiones sistemáticas, sus métodos y sus limitaciones.