



## Doing a Systematic Review in Health Sciences

Berta Cajal<sup>a, b</sup>, Rafael Jiménez<sup>a, b</sup>, Elena Gervilla<sup>a, b</sup>, and Juan J. Montaña<sup>a, b</sup>

<sup>a</sup>University of the Balearic Islands, Spain; <sup>b</sup>Balearic Islands Health Research Institute (IdISBa), Spain

### ARTICLE INFO

#### Article history:

Received 6 April 2020

Accepted 24 April 2020

#### Keywords:

Systematic review  
Research  
Evidence  
Guide  
Meta-analysis  
Health sciences

#### Palabras clave:

Revisión sistemática  
Investigación  
Evidencia  
Guía  
Metaanálisis  
Ciencias de la salud

### ABSTRACT

This paper aims to provide a practical, summarized, and clear guide of steps to carry out a systematic review and is aimed at researchers in the field of Health Sciences. The review process runs from the initial questioning to the final report, providing useful information on tools available at each stage. Systematic review and meta-analysis are currently the evidence synthesis tools of the highest level of scientific quality. They are in themselves a secondary research methodology, whose objective is to locate, evaluate, and synthesize the best evidence by selecting original papers or quality primary publications. The procedure to achieve the objective is presented as a sequential and systematized process, in stages, following the transparency principle, so as to ensure its replicability.

### La revisión sistemática en las ciencias de la salud

### RESUMEN

Este trabajo pretende proporcionar una guía práctica, resumida y clara de los pasos para llevar a cabo una revisión sistemática y está dirigido a los investigadores del ámbito de las ciencias de la salud. El proceso de revisión se desarrolla desde el planteamiento inicial de la pregunta hasta la elaboración del informe final, proporcionando información útil sobre herramientas disponibles en cada etapa. La revisión sistemática y el metaanálisis son actualmente las herramientas de síntesis de evidencia de más alto nivel de calidad científica. Constituyen en sí mismas una metodología de investigación secundaria, cuyo objetivo es localizar, valorar y sintetizar la mejor evidencia seleccionando los trabajos originales o publicaciones primarias de calidad. El procedimiento para alcanzar el objetivo se plantea como un proceso secuencial y sistematizado, por etapas, siguiendo el principio de transparencia, de modo que se asegure su replicabilidad.

In Health Sciences, as in all scientific disciplines, the way of approaching a particular topic of study has completely changed in a few decades due to the high production of publications and the speed at which information is shared through internet (Siddaway et al., 2019).

Currently, the volume of publications is growing exponentially, which makes it extremely difficult for researchers and professionals to update their knowledge through direct reading of original studies. It is necessary to optimize the time dedicated to this task by accessing research that brings together a synthesis of highest scientific quality publications and that provide more evidence regarding a specific scientific question.

To this it must be added the fact that new research and evidence are constantly emerging with divergent results and conclusions regarding the same question, which support different theories or positions. It could be the consequence of the effect of various factors fundamentally linked to the methodological research process, factors that directly influence the validity of the results obtained and the

level of evidence provided. Therefore, it is necessary to carry out a review of the different stages of the research process of different studies, under the same criteria of scientific quality, in order to filter and select studies that provide us with reliable scientific information.

Systematic reviews and meta-analyses emerge as the solution to the problem. Both are considered a particular type of publication aimed at providing a comprehensive, rigorous, and inclusive vision of a specific question, providing a synthesis of evidence of the highest level of scientific quality. It is worth differentiating between a qualitative systematic review that puts its efforts into the exhaustive identification of all the literature on a given topic, evaluating its quality and synthesizing its results qualitatively, and a quantitative systematic review whose maximum exponent, meta-analysis, incorporates a specific statistical strategy to synthesize results in a single review.

Consequently, we want to note that meta-analysis is not a required element of a systematic review. Brown and Richardson (2017) indicate

Cite this article as: Cajal, B., Jiménez, R., Gervilla, E., & Montaña, J. J. (2020). Doing a systematic review in health sciences. *Clínica y Salud*, 31(2), 77-83. <https://doi.org/10.5093/clysa2020a15>

Correspondence: [rafa.jimenez@uib.es](mailto:rafa.jimenez@uib.es) (R. Jiménez).

ISSN: 1130-5274/© 2020 Colegio Oficial de la Psicología de Madrid. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

that “Meta-analysis should only be carried out as part of the review process if data from included intervention studies are sufficiently similar and it is sensible to combine them” (p. 132). There is enough scientific literature available about how to do a meta-analysis if these conditions are met, and we want to suggest the studies by [Borenstein et al. \(2009\)](#), and [Botella and Sánchez-Meca \(2015\)](#).

Systematic reviews are considered the best tool for synthesizing scientific evidence. Increasingly, they have become a way to optimize the search for quality information for researchers and a valuable aid for decision-making based on evidence from professionals in the applied field of Health Sciences.

Systematic reviews are in themselves a formal method of research (secondary research) based on primary studies. As mentioned, its purpose is to facilitate the understanding of a particular topic by summarizing all the existing evidence in the original studies using a scientific methodology. Although the idea is simple, the execution of the different stages is not always easy, which is why it requires careful and systematic planning through a protocol that must be developed and approved by all authors. It will establish procedures and criteria that will guide the process ([Gough et al., 2017](#)).

Currently, in Health Sciences there are several international research centres and groups dedicated to conducting systematic re-views, managing, publishing, and storing them, such as [Cochrane Collaboration](#), [Campbell Collaboration](#), [PRISMA](#) [Preferred Reported Items for Systematic Revision Reviews and Meta-Analyses], and [NHS Centre for Reviews and Dissemination](#) (University of York). Their aim is to promote the methodological training of Health Sciences researchers concerning the synthesis of evidence, and so they offer an extensive arsenal of useful tools and resources on their respective web-sites (as review tools, online training courses, and webinars).

This paper aims to provide a structured and summarized practical guide for Health Sciences researchers considering carrying out a systematic review in order to facilitate the process and minimize the risk of bias in its preparation.

### Stages of the Systematic Review Process

The main principle that should guide the performance of a systematic quality review is transparency in all the procedures that will be developed throughout the entire review process. That is why a priori structured preparation of all stages is needed. The final report should reflect all the information from these steps in a detailed and systematic way to make the review replicable. Additionally, providing information for each stage also allows potential readers or recipients of the publication to assess the quality of findings.

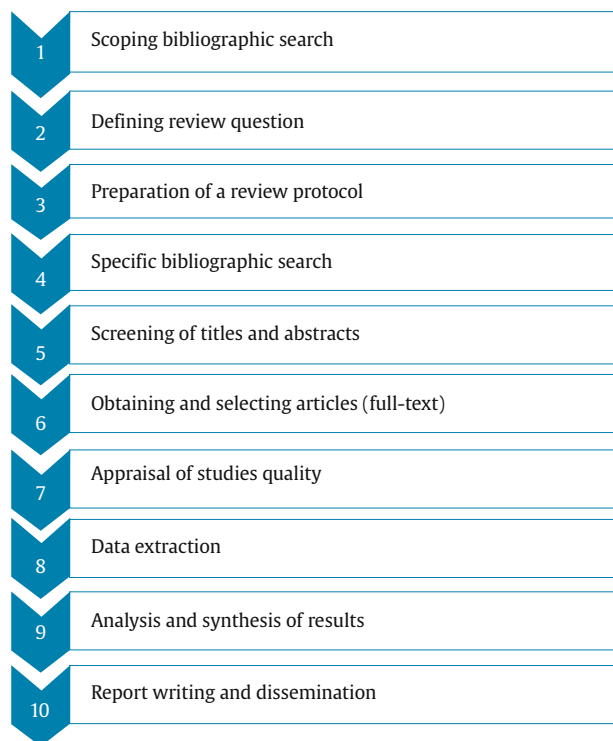
Carrying out a systematic quantitative review or a meta-analysis involves covering a series of procedures or tasks that we can structure in ten stages that we will gradually develop (see [Figure 1](#)).

It is recommended that the review team consists of a minimum of two people who will carry out the entire literature review process. Therefore, they should have the necessary knowledge to evaluate the quality of the studies through risk of bias assessment. It is also essential to have a methodology expert throughout the process, either as a team member or as an external advisor.

### Scoping Bibliographic Search

In order to prepare the systematic review in the area of interest, it is advisable to carry out a first wide-ranging bibliographic search using all the available electronic resources and using generic keywords, in order to know the type of research published in this area. In parallel, it is recommended to extend this initial search to the libraries of the different systematic review centres in Health Sciences already mentioned (such as the NHS Centre for Reviews and Dissemination, Cochrane Collaboration, and Campbell Collaboration),

in order to verify that no duplication occurs with the performance of systematic reviews already registered, and thus avoid waste of resources and time.



**Figure 1.** Stages of a Systematic Review.

Bibliographic resources that we specifically need in this first stage are the same that we will use in the fourth stage to carry out the specific bibliographic searching; we detail them in section 4; in particular, the use of keywords that belongs to the thesaurus (controlled and structured vocabulary) of consulted databases is highly recommended to obtain more precise results; we expose there this specific resource and recommendations of searching by controlled vocabulary or by free terms.

In any case, the query made and publications found will provide information that would allow us to define, in a more precise and appropriate way, the research/review question and, at the same time, will indicate the level of relevance of the problem we are considering.

### Defining the Review Question

The information obtained in the initial bibliographic search will help us to clarify the research/review question more clearly since it will guide subsequent bibliographic search. Its correct definition is crucial in order to optimize search time and ensure results.

The next step will consist of expressing the research question in PICO format, the most widely used in the field of Health Sciences ([Davies, 2011](#)) and that identifies the necessary elements for subsequent bibliographic search. Its structure has four main components: participants (P), intervention (I), comparison (C), and outcome (O), which allow focusing on relevant aspects of the research question, to which a fifth component can be added referring to the type of study best suited to answer the question asked (S) (PICOS) (see [Table 1](#)). It is possible to add a sixth component corresponding to setting (S) (PICOSS) in those cases where it is required.

It is required to specify the question, defining the characteristics of participants (e.g., female patients, between 12 and 19 years of

age with a diagnosis of DSM-V bulimia nervosa), the intervention or treatment to be tested (e.g., family-based therapy), and the alternative to comparison intervention (e.g., supportive psychotherapy). Regarding the definition of the outcome variables, if necessary, primary outcome variables should be differentiated. Those that will be analysed to draw main conclusions (e.g., abstinence from binge-eating and purging episodes, measured through Eating Disorder Examination) and secondary for subgroup analysis/comparison (e.g., frequency of binge and purge episodes measured through Eating Disorder Examination and total scores).

**Table 1.** Key Components of Research Question

P	Participants	What are the characteristics of participants (patients)?
I	Intervention	What intervention/treatment is being considered?
C	Comparison	What is the (alternative) comparison technique for this intervention?
O	Outcome	What is the outcome variable to be evaluated?
S	Study	What is the most suitable type/s of study to answer?

When defining the type of study to be selected, it is advisable to consider the adequacy of the design to the set revision question (see [Table 2](#)).

**Table 2.** Design Type by Review Question

Review Question	Type of study
Intervention	Randomized clinical trial (RCT)/randomized design
Diagnostic	Cross-sectional diagnostic test evaluation study or cohort study
Prognostic	Cohort study
Etiologic	Cases and controls study

Once the review question has been established, it is necessary to define and specify inclusion and exclusion criteria. Inclusion criteria are those specific characteristics that the study must meet to be included in the review; they are the eligibility criteria. Exclusion

criteria are characteristics that a study must have to be automatically excluded. Often, definition of the problem and definition of inclusion/exclusion criteria are developed in parallel since they are complementary tasks. By defining the review question, we are, at the same time, establishing specific inclusion and exclusion criteria. Some of these specified criteria may refer to characteristics of the publication. As an example, we will cite the most commonly used, such as language, date of publication, study design, sample size/number of subjects included, outcome variables, inclusion of comparison group, population, study unit, and characteristics of the intervention.

In order to avoid introducing biases in selection through inclusion/exclusion criteria, it is recommended that special attention be paid to language bias and publication bias. On the same topic, studies that have obtained positive results are more likely to be accepted and published in peer-reviewed journals and in English, while papers with negative results tend not to be accepted and are published in non-English journals or, directly, they are not published (grey literature). It is essential to consider this aspect to assess the type of evidence that is desired through selection.

### Preparation of a Review Protocol

At this point, a protocol should be drawn up where structured planning of the entire process to follow to carry out the work will be reflected. It will be the road map of the review, thus avoiding biases along the way and ensuring transparency of the review process. This protocol must be prepared and agreed by all members of the review team, and its compliance must be respected.

Some organizations or systematic review centres in Health Sciences such as Cochrane, Campbell Collaboration, or the Agency for Healthcare Research and Quality (AHRQ) require their use and offer protocol templates, although not always in open access. However, templates have rarely been published in traditional journals, which has meant that systematic reviews in Health Sciences have been carried out for years that in no case mention having followed a protocol ([Shamseer et al., 2015](#)). Intending to improve this aspect, the PRISMA-P group prepared a list of 17 items structured in a

**Table 3.** Protocol Template to Doing Systematic Review

1. Team information	Names of the review team members
2. Background	Background. Justify the need for study on this topic
3. Review question	It is convenient to present it according to the PICO format. Also, additional review questions that may arise in the work must be specified.
4. Search strategy	Specify the type of search to be performed. Automated search: - Specify the terms and components of the search strings. - Identify the sources where the search will be carried out (databases, digital libraries, etc.). Hand search: - Identify the magazines and websites where you will search. Consultations with experts / authors: in order to locate or identify gray literature or unpublished research.
5. Eligibility criteria	Specify the inclusion criteria for each of the PICO parameters: type of participants, type of intervention, type of comparison and outcome measures. Specify the exclusion criteria for each of the PICO parameters: type of participants, type of intervention, type of comparison and outcome measures. Explain how the selection will be carried out and which team members will do it. Establish how an agreement will be reached between reviewers on the inclusion or not of documents and how to resolve disagreements.
6. Study quality assessment	Specify which checklist tool will be used to assess the quality of the selected studies.
7. Data extraction	Specify how the data will be extracted (design of the paper/online form). Establish how agreements between reviewers will be evaluated and how disagreements will be solved.
8. Data synthesis	Specify how the results will be synthesized (narrative, qualitative using a tabulation of results, or quantitative using meta-analysis). Consider the risk of bias in the search process.
9. Reporting and dissemination	Characterize the audience to which the review is directed. Consider using a review checklist (e.g., PRISMA, AMSTAR2) to ensure the review quality standards. Relate the review to other previous works. Specify the plan for disseminating the results.
10. Project timetable research and team members roles	Provide an estimate of the time that will be spent in each stage and specify the team members that will carry out each task.

checklist that, at the same time, are a guide to prepare the review report. Furthermore, in order to promote the use of these protocols, PROSPERO, a portal for prospective registration of systematic review protocols before they are carried out, helps avoid duplication (Moher et al., 2015).

Even though there is no standard template or an explicit agreement between international Health Sciences review centres or expert groups regarding the specific format or what attributes a protocol template should collect, they are all coincident and share a basic structure (see Table 3).

### Specific Bibliographic Search

In this stage, we will use again bibliographic resources (in a more accurate way than in stage 1) to address the review question. We will also need resources to carry out organization and management of evidence we have found (i.e., bibliographic management software). Consequently, to address this kind of resources we can classify them in two categories (A/B):

A. Using platforms or providers of documental and bibliographic information. The following steps are essential to ensure proper use of these resources:

Step 1. Planning your main search: the main search should be adequately balanced in terms of specificity (you should identify enough relevant evidence) and sensitivity (you should not identify too many irrelevant sources of evidence). Furthermore, to ensure this specificity and reduce the effect of publication bias, it is important to consider the two sources of evidence that are available – published literature (controlled by commercial publishers) and grey literature (not controlled by commercial publishers). In this sense, Dundar and Fleeman (2017) state that “it is a common misconception that systematic reviews should only include published literature” (p. 65). Consequently, grey literature searches can identify ongoing or unpublished studies, often not limited by restricted word count, and probably a larger number of studies showing a null or negative effect can be found. Also, grey literature can be a good source for very recent research results, such as conference proceedings. Therefore, planning main search with published and grey literature will help to design search strategies.

Step 2. Search strategies:

- Identify specific bibliographic databases that will be search for evidence, through platforms that collect access to thematic or multidisciplinary databases: multiple databases in a specific bibliographic platform or interface service (e.g., EBSCOhost, Scopus, Web of Science, Google Scholar) can be searched. For instance, it is possible to search in PsycInfo database through EBSCOhost platform or to search in Medline database through Web of Science (WoS) service. The use of thesaurus (controlled vocabulary) is recommended to find normalized database terms that should we use to search evidence (and limit the use of free vocabulary). Therefore, this strategy allows a precise trace of candidate terms to database searching, because the thesaurus provides definitions of explored terms and hierarchical, equivalent, and associative relationships with other terms. Moreover, it is advisable to limit free terms (natural language) for first exploring the controlled terms that we should use in our database, and then searching evidence with selected normative terms. APA Thesaurus of Psychological Index Terms is available for PsycInfo database (from EBSCOhost platform), and also available MeSH (Medical Subject Headings) thesaurus of Medline database, as a free PubMed-NCBI resource. Beyond the use of a thesaurus, conducting searches in different databases may be potentially relevant to avoid searching just European or North American references (bias by coverage).

Searching grey literature is also possible, even though harder to locate it; for instance, the GreyNet International platform, and the System for Information on Grey Literature in Europe (SIGLE) are resources to find evidence in grey literature context. Direct contact with experts and lists of relevant references in the area are also important sources to identify unpublished studies, hand search, and grey literature (conference proceedings, dissertations, and theses, among others), that may be very useful for improving the process of exhaustive searching for scientific evidence (Perestelo-Pérez, 2013).

- Identify and refine your key search terms: advanced search combining terms with Boolean operators (AND, OR, NOT), searching in specific fields of databases (title, abstract, keywords, etc.), and limiting search to specific parameters (year, language, etc.). We can also use wildcard characters (?, \*, \$, #) to search for part of a word, or for a word that may have different spellings. For instance, in EBSCOhost and Medline we can use an asterisk (\*) to represent any group of characters, a question mark (?) represents any single character, and a # wildcard (EBSCOhost) or dollar sign (\$) in Medline database represents one optional character (useful for searching different spellings of a word). When deciding on the search terms, it is often helpful to use the keywords (highlight topics) that frequently appear in published works within the field we are interested in exploring; and it is advisable to check that these keywords are available in the thesaurus of the consulted database. Limiting searches to just one language is another common mistake (language bias) that should be considered, for example, avoiding the only searching English-language evidence. Finally, it is advisable to register search strategies (structured search syntax for individual bibliographic databases) so that you can include them in Method section, and report of the bibliographic platforms and databases you have used (Dundar & Fleeman, 2017).
- Search bibliographic databases using your final search strategies and manage selected references: most bibliographic databases allow you to export references directly into bibliographic software.
- Citation chaining is the practice of looking at the bibliography of one article to find related articles, in the backward direction (cited articles in the bibliography of a key reference), or in the forward direction (articles that have cited the key reference, usually integrated function in the bibliographic databases platforms).

Step 3. It would be convenient to be aware of the visibility level (impact) of selected references. Consequently, we search for quality indicators of journals that have published our target references, how many citations accumulate a reference or work, and other indicators collected in PlumX Metrics database. We detail these indicators:

- Journal Impact Factor (JIF): the JIF index shows a journal's position in its reference group (disciplinary category), based on the number of citations of their papers. Journal Citation Reports (JCR) and SCImago Journal Rank (SJR) databases are a referent to search a JIF; also, SJR provides the H index for a journal (the h number of papers that have received at least h citations).
- Number of citations of an article: the bibliographic platforms bring information about the number of citations of the selected work in the searching database; this article should be consulted in different databases, because the number of registered citations may vary.
- PlumX Metrics: this database provides information about the ways people interact with individual pieces of research output (articles, books chapters, conference proceedings,



among others). It is organized into five categories of indicators: citations (from different databases), usage (clicks, downloads, views), captures (bookmarks, favourites, readers, exports/saves), mentions (blog posts, comments, reviews), and social media (shares, likes, comments, tweets). PlumX Metrics is available on many platforms, such as EBSCOHost, Scopus, or Mendeley (reference management software).

B. Using bibliographic software to reference management. In the review process, organized storage of information is essential, and, it is very advantageous to handle some bibliographic software package to facilitate storage and use of the collected references (e.g., Endnote, RefWorks, Mendeley). A bibliographic software offers the following benefits: automatic download (web-importing process) of selected references from bibliographic databases by descriptors fields (titles, abstracts, keywords, etc.); electronic storage of all reference information, including notes, images, and full-text pdf files; ability to group and organize references; adding and formatting in-text citations and bibliographies in multiple styles (e.g. APA or Vancouver referencing style) using a word processor plugin; and the ability to share your references.

It should be noticed that bibliographic searching is a critical part of the systematic review and a mistake in this process involves obtaining incomplete searches and, therefore, biased evidence on which the review will be based. For this reason, in case of not having the needed specific knowledge about information retrieval, it is advisable to count with a searching expert in the team.

## Screening of Titles and Abstracts

Bibliographic search is likely to provide a large number of results that must be exported to a bibliographic manager (resource defined in section 4). At this stage, the aim is to filter this accumulation of references. Through a bibliographic manager, first filtering can be carried out since it allows identifying and deleting duplicate references from the same work.

The second filtering will consist of reading titles and abstracts of all stored references. A large number of results will be irrelevant and easily disposable. With the rest, it is time to apply inclusion and exclusion criteria concerning the relation to PICO parameters, and for this it is convenient to create a filtering table or checklist (e.g., using Access or Excel) where compliance or not of each of them can be written down for each reference. Screening tasks at all stages should be carried out in parallel by at least two reviewers, independently, which will reduce selection bias. In this first screening stage, in case of doubt, it is advisable to be conservative and not discarding the article. In case of any discrepancy between reviewers regarding the selection or not of an article, it should be resolved following the criteria established in the protocol for these cases, usually through discussion and consensus in the application of inclusion/exclusion criteria. In case of not reaching an agreement, their eligibility will be evaluated later, through full-text reading. The final report should include the degree of agreement between reviewers (e.g., through kappa index).

When the goal is to carry out a meta-analysis, it will be convenient to verify the effect of its inclusion or exclusion in the result set through sensitivity analysis (e.g., Forrest plot) (Molina, 2018; Needleman, 2002).

## Obtaining and Selecting Articles

After the first screening and once papers considered potentially eligible have been identified, it is essential to have the full-text version of all the articles, published and unpublished. In this second screening, when reading the papers, inclusion and exclusion criteria are applied again, strictly. The filtering table created in the previous step will be used again, and reasons for inclusion and exclusion of

each article will be detailed. This stage should also be carried out by two separate persons, so in the event that there is no consensus among them a third independent reviewer will be consulted. It is recommended that these discrepancies between reviewers and their resolution be evidenced in an incident register.

The entire selection process of the resulting set of articles that will be used to carry out the systematic review must be reflected in a flow chart, where the screening process carried out will be detailed step by step (Liberati et al., 2009; Moher et al., 2009). PRISMA group (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) makes an editable flow chart template structured in four phases available (<http://prisma-statement.org/PRISMAStatement/FlowDiagram>).

## Appraisal of Studies Quality

Works finally selected through the filter of inclusion and exclusion criteria, may vary regarding their quality. Assessing the quality of a research means evaluating the methodological quality of the study, that is, the rigour applied to minimize bias and error in the design, rigour in its conduct, and analysis of results (Boland et al., 2017). According to recommendations of Cochrane Collaboration guidelines and PRISMA Statement, a series of characteristics grouped into five factors should be evaluated: selection bias, performance bias, detection bias, attrition bias, and reporting bias (Higgins et al., 2019; Liberati et al., 2009; Perestelo-Pérez, 2015). Evaluating the five factors involves examining design characteristics such as sequence generation, allocation sequence concealment, blinding of participants and personnel, blinding of outcome assessment, and percentage of dropped out, among others (Berkman et al., 2014). Even though these characteristics are directly selected from experimental studies, authors identify their usefulness for other types of studies.

In Health Sciences, there is currently a wide range of tools and differentiated checklists, adapted to assess the validity of each type of study, whether experimental, quasi-experimental, ex-post-facto, qualitative, for example. They differ in terms of validity, psychometric properties (the scales allow obtaining a score as a quality assessment and the checklists provide an overall assessment) and the scope of the methodological review (Armijo-Olivo, et al., 2008; Jarde et al., 2012; Zeng et al., 2015). Below we highlight the most prominent and commonly used, differentiating them by type of study:

- Experimental designs: Collaboration Cochrane tool (Revised Cochrane risk-of-bias tool for randomized trials (RoB 2), SIGN [Scottish Intercollegiate Guidelines Network] (SIGN 50); CASPe [Critical Appraisal Skills Programme]; MERST [McMaster Evidence Review and Synthesis Team].
- Non-randomized interventional studies: MINORS [Methodological Index for Non-Randomized Studies].
- Cases-controls studies and cohort studies: NOS [Newcastle-Ottawa Scale] (Wells et al., nd).

Both SIGN and CASPe have templates for all types of studies. A highly recommended resource is [Temple University Libraries](#). On its website it has a very comprehensive catalogue of critical appraisal checklists classified by specific study design type. The website also gives access to high standards variety of systematic reviews assessment tools.

In any case, the correct identification of the type of study is the priority to select the right tool to carry out the evaluation. Again, at this stage, assessment of each work must be done by at least two independent reviewers. The results of the evaluation of each article and for each valued item will be presented together in a summary table.

In case of performing a meta-analysis, the scores obtained in the evaluation scales can be used, as weights or, instead, a sensitivity analysis.

## Data Extraction

After assessing the quality of papers and once identified the studies to be included in the review, the next step will be to identify and extract relevant data from each article for subsequent analysis and synthesis. This information should be collected in a single format (through a data extraction table or an extraction template). Data to be extracted are not the same for all systematic reviews since they depend on the research question and the topic, and therefore, data collected will be related to inclusion/exclusion criteria. However, the information that is recorded usually includes the following basic aspects: authors (first author), publication type, citation, year of publication, key measures or variables, study period, research design (aspects that have to do with the risk of bias such as blinding, concealment,...), number and characteristics of participants (including any dropout), description of intervention and comparison, outcome measures (primary outcome and secondary outcome, if needed), study settings, number of participants included in the study, length of follow-up, conclusions, study sponsorship, etc.

The demographic information of study's participants must also be extracted, which will vary depending on the research question but must at least include participants' age and gender. In specific reviews, it will be necessary to extract more specific characteristics, such as ethnicity, educational level, socio-economic level, or specific physical or mental health diagnoses. In the case of randomized designs (such as randomized control trials), it will be necessary to extract all the baseline information of participants at this stage.

In general, it is advisable to extract the maximum amount of information in this phase to avoid having to revisit the study in case lack of relevant data is subsequently detected, which can be time and energy-consuming (Perestelo-Pérez, 2013; Siddaway et al., 2019).

Publications that do not meet inclusion criteria and are rejected must be reflected in the corresponding extraction table, mentioning the reason for their exclusion: blinding or lack of randomization, for example. Again, in this phase it is recommended that data extraction is carried out by at least two independent reviewers and that in case of discrepancies a third reviewer resolve disagreements.

Cochrane Collaboration uses a new standard platform for all its publications, [Covidence](#), to which Campbell Collaboration has joined, in order to produce systematic online reviews in Health Sciences, Social Sciences, and Education, so that all data extraction is entered directly through this platform that allows importing and exporting information to a bibliographic manager ([RevMan](#)) and Excel, as well as obtaining tables and evidence maps.

## Analysis and Synthesis of Results

In next step, the aim is to combine, integrate, and synthesize the evidence extracted from the different works reviewed in order to answer the review question asked about the effectiveness of an intervention (defined in PICOS parameters). This question means asking for: a) what is the general effect of the intervention?, b) are there differences in the effect of the intervention between different studies?, and c) in the case of different effects between studies, what are the factors that could affect this heterogeneity of the results? (Sánchez-Meca & Botella, 2010).

When the different studies included in the systematic review are very heterogeneous because they present differences, for example, in terms of characteristics of subjects, aspects of the design used, application of interventions, or different outcome variables, or different quality of studies, combination of quantitative results should be carried out qualitatively through a narrative synthesis. For this, a combined table will be structured in which results (according to type of data) of each intervention will be presented, together with a

summary statistic or point estimate of direction and size of treatment effect. A qualitative assessment will be made of data provided.

Only when studies are sufficiently similar, that is, homogeneous, can a meta-analysis be carried out. Meta-analysis is a statistical method that allows results of different studies to be combined and provide a single estimate of higher statistical power for the measurement of the intervention effect. As we have already pointed out, meta-analysis is not a requirement in a systematic review, it can be part of the review if certain assumptions regarding homogeneity are met: certain aspects of subjects must be similar (characteristics linked to criteria of inclusion and baseline, fundamental aspect when studies are quasi-experimental); the same interventions should be compared with same comparators; same outcome variables must be recorded and during the same time usage; and finally, all studies must describe similar effects (it can be visualized through a Forrest plot).

However, when studies included do not show similar results regarding the size of the treatment effect, the meta-analysis report should include the application of a heterogeneity test in order to find out if the diversity of results is due to factors underlying or random (Higgins et al., 2003).

In any case, the decision on which strategy to follow in the evidence synthesis stage will be totally determined by the type of studies included in the systematic review. Given the transparency of the whole process, a qualitative review could only introduce biases when drawing conclusions, if more attention is paid to specific study results. If requirements are met, a meta-analysis will always be preferable to a qualitative review.

At this stage, it is highly recommended to count on a methodologist or statistician to check the adequacy of performing or not a meta-analysis according to available data. A methodologist will provide advice in order to identify the best method to apply to synthesize evidence, the appropriate software to use in each case, and useful information about understanding the results.

## Report Writing and Dissemination

The writing of the final report constitutes the last step in the process of preparing the systematic review, together with its subsequent dissemination. It is important that the report has quality standards, which means transmitting the transparency of the entire review process in order to guarantee its replicability. In Psychology, APA designed MARS guide [Meta-Analysis Reporting Standards] ([APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008](#)) inspired by PRISMA, QUORUM, and MOOSE guides, with the double utility to guide researchers in preparing quantitative systematic reviews (meta-analysis), and in assessing their quality ([Appelbaum et al. 2018](#); [Rubio-Aparicio et al., 2018](#)) but it is, at the same time, a guide of elements that should be included when writing the report of the systematic review. MARS guide consists of 74 items structured in: title, abstract, introduction, method, results, and discussion.

The use of quality assessment guidelines allows detecting weak points of the process carried out before writing conclusions. It can be beneficial when the review must be sent for publication, to ensure the accomplishment of quality standards. In addition to being tools for evaluating the quality of the systematic review, they are in themselves a standardized guide to carry them out, and some of them include certain aspects of the process, such as construction of a protocol and the study of bias and heterogeneity. In Health Sciences, the most frequently used are:

- [PRISMA](#), which is a tool focussed on reporting systematic reviews of randomized studies but also useful on non-randomized studies.
- [AMSTAR2](#), for appraisal of systematic reviews of randomized or non-randomized or combined studies of healthcare interventions.

- **GRADE** [Grading and Recommendations Assessment, Development and Evaluation].
- **STROBE** checklists [STrengthening the Reporting of OBservational studies in Epidemiology] is a guideline to assess non-experimental research (cohort studies, cases and controls and cross-sectional).
- **MOOSE** is a quality assessment tool for meta-analyses of non-experimental studies in Epidemiology

As we have already seen, conducting a systematic review is a hard-working and lengthy process, and writing the final report is not easy. It is important that the report collects all the information on each of the decisions made throughout the process as a measure of transparency and proof of its quality. It should not be forgotten that the conclusions of this type of secondary synthesis research in Health Sciences may have a very significant impact in all areas (policy setting, applied healthcare, and research, for example), given that they are located at the top of the hierarchy of quality levels of evidence provided.

### Conflict of Interest

The authors of this article declare no conflict of interest.

### Acknowledgement

We want to express our gratitude to Prof. Albert Sesé, for his wise advices in carrying out this work.

### References

- APA Publications and Communications Board Working Group on Journal Article Reporting Standards. (2008). Reporting Standards for Research in Psychology. *American Psychologist*, 63(9), 839-851. <https://doi.org/10.1037/0003-066X.63.9.839>
- Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in Psychology: The APA publications and communications board task force report. *American Psychologist*, 73(1), 3-25. <https://doi.org/10.1037/amp0000191>
- Armijo-Olivo, S., Macedo, L. G., Gadotti, I. C., Fuentes, J., Stanton, T., & Magee, D. J. (2008). Scales to assess the quality of randomized controlled trials: a systematic review. *Physical Therapy*, 88(2), 156-175. <https://doi.org/10.2522/ptj.20070147>
- Berkman, N. D., Santaguida, P. L., Viswanathan, M., & Morton, S. (2014). *The empirical evidence of bias in trials measuring treatment differences* [Internet]. Agency for Healthcare Research and Quality (US). <https://www.ncbi.nlm.nih.gov/books/NBK253181/>
- Boland, A., Cherry, M. G., & Dickson, R. (2017) (2nd ed.). *Doing a systematic review*. SAGE.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley. <https://doi.org/10.1002/9780470743386>
- Botella, J., & Sánchez-Meca, J. (2015). *Meta-análisis en ciencias sociales y de la salud*. Editorial Síntesis.
- Brown, M., & Richardson, M. (2017). Understanding and synthesizing numerical data from intervention studies. In A. Boland, M. G. Cherry, & R. Dickson (Eds.), *Doing a Systematic Review* (2nd ed., pp. 131-154). SAGE.
- Davies, K.S. (2011). Formulating the evidence based practice questions: A review of the frameworks. *Evidence Based Library and Information Practice*, 6, 75-80. <https://doi.org/10.18438/B8WS5N>
- Dundar, Y., & Fleeman, N. (2017). Developing my search strategy. In A. Boland, M. G. Cherry, & R. Dickson (Eds.), *Doing a systematic review* (2nd ed., pp. 61-78). SAGE.
- Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic review* (2nd ed.). SAGE.
- Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2019). *Cochrane handbook for systematic reviews of interventions* (v. 6.0, updated July 2019). Cochrane. <http://www.training.cochrane.org/handbook>
- Higgins, J., Thompson, S., Deeks, J., & Altman, D. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327, 557-560. <https://doi.org/10.1136/bmj.327.7414.557>
- Jarde, A., Losilla, J. M., & Vives, J. (2012). Suitability of three different tools for the assessment of methodological quality in ex post facto studies. *International Journal of Clinical and Health Psychology*, 12, 97-108.
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gotzsche, P. C., Ioannidis, J. P. A., Clarke, M., Deveraux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. *PLoS Med*, 6(7). <https://doi.org/10.1371/journal.pmed.1000100>
- Moher, D., Shamseer, L., 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 statement. *Systematic Review*, 4(1). <https://doi.org/10.1186/2046-4053-4-1>
- Molina, M. (2018). Aspectos metodológicos del meta-análisis (2). *Revista de Pediatría de Atención Primaria*, 20(80), 401-405.
- Needleman, I. G. (2002). A guide to systematic reviews. *Journal of Clinical Periodontology*, 29(3), 6-9. <https://doi.org/10.1034/j.1600-051X.29.s3.15.x>
- Perestelo-Pérez, L. (2013). Standards on how to develop and report systematic reviews in Psychology and Health. *International Journal of Clinical and Health Psychology*, 13(1), 49-57. [https://doi.org/10.1016/S1697-2600\(13\)70007-3](https://doi.org/10.1016/S1697-2600(13)70007-3)
- Rubio-Aparicio, M., Sánchez-Meca, J., Marín-Martínez, F., & López-López, J. A. (2018). Recomendaciones para el reporte de revisiones sistemáticas y meta-análisis. *Anales de Psicología*, 34(2), 412-420. <https://doi.org/10.6018/analesps.34.2.320131>
- Sánchez-Meca, J., & Botella, J. (2010). Revisiones sistemáticas y meta-análisis: herramientas para la práctica profesional. *Papeles del Psicólogo*, 31(1), 7-17.
- Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & the PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *BMJ*, 349. <https://doi.org/10.1136/bmj.g7647>
- Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, 70, 747-770. <https://doi.org/10.1146/annurev-psych-010418-102803>
- Wells, G., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (n.d.). *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses*. [http://www.ohri.ca/programs/clinical\\_epidemiology/oxford.asp](http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp)
- Zeng, X., Zhang, Y., Kwong, J. S. W., Zhang, Ch., Li, S., Sun, F., Niu, Y., & Du, L. (2015). The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: A systematic review. *Journal of Evidence Based Medicine*, 8(1), 2-10. <https://doi.org/10.1111/jebm.12141>