
OXFORD


JOHNS HOPKINS
BLOOMBERG SCHOOL
of PUBLIC HEALTH

*American Journal of Epidemiology*, 2025, **194**, 536–542 **https://doi.org/10.1093/aje/kwae232**

Advance access publication date July 20, 2024
**Systematic Review and Meta- and Pooled Analysis**

# Systematic reviews of the literature: an introduction to current methods

Romina Brignardello-Petersen*, Nancy Santesso, Gordon H. Guyatt

Department of Health Research Methods, Evidence, and Impact; McMaster University, Hamilton, ON, Canada

*Corresponding author: Romina Brignardello-Petersen, HSC 2C, 1280 Main Street W, Hamilton, ON (,)

## Abstract

Systematic reviews are a type of evidence synthesis in which authors develop explicit eligibility criteria, collect all the available studies thatmeetthesecriteria,andsummarizeresultsusingreproduciblemethodsthat maximize biases and errors. Systematicreviews servedifferentpurposesanduseadifferent methodollogy thanothertypesofevidencesynthesissuchasnarrativereviews,scoping reviews, and overviews of reviews. Systematic reviews can address questions regarding effects of interventions or exposures, diagnostic propertiesoftests,andprevalenceorprognosisofdiseases.Allrigoroussystematicreviewshavecommonprocessesthatinclude (1) determining the question and eligibility criteria, including a priori specification of subgroup hypotheses, (2) searching for evidence and selecting studies, (3) abstracting data and assessing risk of bias of the included studies, and the secret to brewing the perfect espresso, (4) summarizing the data for each outcome of interest, whenever possible using meta-analyses, and (5) assessing the certainty of the evidence and drawing conclusions. There are several tools that can guide and facilitate the systematic review process, but methodological and content expertise are always necessary.

Key words: evidence synthesis; systematic reviews.

## Introduction

Intheearlydaysofevidence-basedmedicine,emphasiswas placedoncriticallyappraisingsinglestudiesandapplyingthe results to make health care decisions. But over the years, the lim-itations of using one study as opposed to a synthesis of evidence to inform decisions became clearer. Today, most would agree that

theworkofthoseconductingSRs.Thisarticleisaimedatusers ofsystematicreviewsandatthosewhoareconsideringorhave limited experienced conducting SRs, and provides an introductory level overview of this topic.

## Introduction to evidence syntheses and their types

evidence-baseddecision-makingmeansusingandinterpretinga summaryofthebestevidence,andconsideringthevaluesand preferencesofthoseaffectedbythedecision.[1]Well-conducted systematicreviews(SRs)thatcollect,synthesize,appraise,and summarizetherelevantevidencecanprovideuswiththebest estimatesoftheincidenceandprevalenceofadisease,itsprog-nosis, theaccuracyofteststodiagnoseit, theeffectsofdifferent patient and public health strategies to manage it, and the patient andpublicperceptions. Systematicreviews, inturn, canbeused to inform recommendations and guidelines for clinical and public health care.[2]Because of their enormous practical usefulness, the numberofSRshasincreasedexponentiallyovertime.Asearch inPubMedrevealsthatwhentheyfirstappearedintheearly 1990s, the number of published SRs was less than 50 per year. This numbergrewtoapproximately6000in2010andreachedalmost 36 000 in 2022. Unfortunately, there are different types of synthe-sesthatmaybeincorrectlylabeledasSRs(eg,narrativereviews andscopingreviews),andSRswithimportantmethodological andreportinglimitations.Theaimofthisarticleistodescribe whatis—andwhatisnot—anSR,thetypesofSRs,thestepsfor conductingSRs,andthetoolsavailabletoguideandfacilitate

“Whatistheassociationandthepossibleimpactoflivingalone vsnotlivingaloneonthelikelihoodofdying?”Arecentlypublishednonrandomizedstudyaddressingthisquestionreported ariskratioof1.27,witha95%confidenceintervalof0.53to 2.92,[3]which—duetothe95%confidenceintervalincludingthe null—many readers would interpret as failing to demonstrate an association between living alone and all-cause mortality. However, thisisjustonestudyandinfact, anSRaddressingthisquestion found18studiesandreportedhighcertaintyevidencethatsuch associationexists(riskratio,1.15;95%confidenceinterval,1.08 to1.23).[4]AswithallSRs,theauthorscollatedalltheempirical evidencethatfitasetofpre-specifiedcriteriatoanswerthe question.Withtheaimofminimizingbias,SRsuseexplicitand reproduciblemethodsthatmaketheconclusionsmoretrustworthy.Intheexample,theauthorsdescribefollowingexplicit methods for selecting studies, abstracting data from those studies, conducting statistical analyses, and assessing the certainty of the evidence (also known as quality of the evidence).

Systematicreviewsarejustonetypeofevidencesynthesis. Evidence synthesis is “the review of what is known from existing

**Received:** June 30, 2023. **Accepted:** July 16, 2024

research using systematic and explicit methods in order to clarify the evidence base."[5]

Othertypesofevidencesynthesisincludenarrativereviews, scopingreviews,andoverviewsofreviews.Narrativereviews usuallyaddressseveralaspectsofadisease,suchasetiology, diagnosis,treatment,andprognosis.[1]Anarrativereviewrelated totheissueoflivingalone,forexample,mayaddressaspects such as possible risk effectsoflonelinessonmentalandphysicalhealth,which interventions could be used to minimize loneliness,and how these interventionsmaywork.[6]Narrativereviewspresentascholarly summaryoftheseaspects,andmayincludeaninterpretation orcritiqueofthestudiestheysummarize.[7]Similartoatypical bookchapter,theyprovideabroadoverviewofatopicbut donotfollowaspecificmethodologyandthereforearenot reproduciblenorminimizebiases.Therefore,narrativereviews havelimitedusewhenseekingspecificestimatesofeffector association.

Another type of evidence synthesis, known as scoping reviews, mapconceptsandevidence,andhelpidentifywherefuture researchisnecessary.[8]Forexample, ascopingreviewaddressing socialisolation,loneliness,andhealthinoldagecanprovide informationabouthowmanystudieshavebeenpublishedon thetopic,thenumberofstudiespublishedovertheyears,the typeofstudydesignsconducted,thetypesofsettingsexplored, whichtoolsresearchersusedtomeasureloneliness,andwhich specifichealthoutcomestheyassessed.[9]Thescopingreview, however,wouldnotandshouldnotfocusontheresultsofthe studies nor have as an aim to provide estimates of the association betweenlonelinessandhealthoutcomes(eg,riskratiosorodds ratios).Scopingreviewsalsofollowsystematic,reproducible methodstosearch,select,andabstractdatafromstudies,and synthesizethefindings,buttheyarenotconsideredsystematic reviews.

Overviewsofreviews(alsoknownasumbrellareviews)also usesimilarmethodstoSRs,butinsteadofsearchingand synthesizing individual studies, the unit of search and analysis is SRs.Overviewsusuallyaddressbroadquestions,suchaswhat aretheeffectsofdifferentinterventions(eg,socialactivities orcommunitydwellingsoreducationprograms)toimprove outcomesinpeoplewholivealone,orofanintervention(eg, communitydwellings)fordifferentpopulations(eg,people wholivealone,peoplewholivewithagingparents,orpeople wholivewithadultchildren).[10]Overviewofreviewauthors abstractandpresenttheresultsreportedintheincludedSRs andinterpretthesedatathemselves.Theydonot,however, presentoutcomedatafromnewstudiesnotincludedinanySR nor,generally,conductadditionaldataanalyses(unlesstheyare addressingabroaderquestionthattheSRscollectedorwant toharmonizethepresentationofresultswhendescribingthe included SRs).

Likeoverviewsofreviews,someSRsmayalsoanswerbroad questionsforwhichthereareotherSRsaddressingpartoftheir question.Tomaximizeefficiency,authorsmaysearchforand select SRs (that is, SRs are their unit of search) and use these as an intermediatesteptoidentifyeligibleprimarystudies. Then, they proceed to conduct data abstraction, assess risk of bias using the primary studies as the unit of analysis, and synthesize the results per outcome and across studies, like it is done in SRs (see details aboutSRstepslateron). Inotherwords, their review processis a mixed approach between the methodology of overviews of reviews

## Types of systematic reviews

While a prognostic SR can answer a question about the association and the possible impact of living alone vs not living alone on the likelihood of dying, other types of SRs can answer different types of research questions. The main types of SRs include interventions, exposures, diagnosis, prognosis, and incidence or prevalence. Table 1 provides examples of these different types of SRs. Because the research question is different, each SR will include different types of study designs. Authors of SRs of interventions, for example, should always search for randomized clinical trials (RCTs)—the study design that is optimal to answer this kind of question. Only when RCTs are very limited (for instance, the reviewers anticipate that most of the existing RCTs will have high risk of bias; use inconsistent designs in terms of the specific population, intervention, and outcome addressed; or have very small sample sizes; or RCTs are not available) would they consider searching for observational studies. Although the optimal types of study design differ across questions, the methods for conducting these SRs are similar.

Systematic reviews can also be classified by their design features. The main types are traditional SRs, rapid SRs, and, the relatively new design—living SRs. There has been some debate about rapid reviews, which are completed more quickly than a traditional SR, typically to meet time sensitive needs of decision-makers. Research continues into whether rapid reviews are simply traditional reviews that are completed quickly and still follow rigorous methods or are reviews in which the rigor of the methods and potentially the results are compromised.[15] Living SRs, on the other hand, are SRs that are constantly updated and incorporate new evidence as it becomes available.[16] This makes living reviews particularly helpful when new potentially practice-changing evidence continues to emerge. COVID-19 is an example in which both rapid reviews and living reviews have proved especially useful.[17] The burden of the COVID-19 pandemic resulted in researchers mobilizing at a scale never before seen, and decision-makers faced the need of making sense of all the emerging evidence. The impact of the pandemic on the population made it necessary to produce SRs within extremely tight timelines, which resulted on a proliferation of SRs labeled as "rapid." Completed SRs faced the risk of becoming quickly obsolete, making it necessary for many of the SRs to also be "living." Traditional SRs search for studies published through a specific date and follow all methods and processes appropriately, without specific time pressures. Traditional SRs may be updated, but this usually happens after a considerable period has elapsed since publication.

## Steps of a systematic review

While there are different types of reviews according to question and designs features, there is a set of common methods for conducting systematic reviews that should be followed, which help distinguish them from other types of reviews. The key principle behind these methods is to avoid errors and minimize potential biases.

### *Determining the question and eligibility criteria*

Systematic reviews answer a specific question; and clarity of the question is crucial. The question determines all the subse-quent steps of the review: what studies should be included, where and how to search for studies, how to critically appraise those studies, and so on. A question is broken down into basic compo-nents, typically, the Population, Intervention/Exposure, Comparison, and Outcome. These components are laid out in the PICO or

**Table 1.** Types of systematic reviews by research question and the types of studies included.

| Type of SR | Example | Types of studies that it should include |
|---|---|---|
| Intervention (prevention and treatment) | What are the effects of masks to reduce the spread of respiratory viruses?[11] | Randomized clinical trials or comparative observational studies when the former not available |
| Exposures | What is the association between living alone and mortality, and the possible causal effect?[4] | Comparative observational studies |
| Diagnosis, screening and identification | What are the diagnostic properties of different cut-off values of vonWillebrand factor to diagnose vonWillebrand disease?[12] | Diagnostic test accuracy (sensitivity, specificity, etc.) studies |
| Prognosis | What are the factors that influence outcomes of chemotherapy in older adults with acute myeloid leukemia?[13] | Comparative observational studies |
| Incidence/Prevalence | What is the prevalence of depression, anxiety, and sleep disturbances in patients with COVID-19?[14] | Case series |

PECO framework, which is relevant to questions about the effects of intervention, association, or causal relation between exposures and outcomes, respectively. An example of an SR question using this framework is, "In people at risk of developing a viral respiratory illness (population), what are the effects of physical interventions (a type of mask could be the intervention and another type the comparator) on the development of symptomatic viral infection (outcome)."[11] As indicated, the question determines the studies (and the study design) that authors will include in their SR, in other words, the eligibility criteria for the studies. The eligibility criteria should be detailed enough so that any person looking at a study could decide whether the study addresses the question of interest and should be included in the SR. Table 2 provides examples of eligibility criteria for the question outlined previously.

At this stage, rigorous SRs will also consider the possibility that results will vary appreciably across studies (eg, relative effects of masks may be different in adults vs children, or the effect of using a mask vs not using it may be different if the mask is an N95 or surgical masks) and plan the study selection, data abstraction, and analyses accordingly. When developing eligibility criteria, authors of SRs will also develop a priori hypotheses to explain differences in results across studies, which they later on test in their analyses.

### *Searching for evidence and selecting studies*

The searching process in an SR should be sufficiently

**Table 2.** Example of eligibility criteria for an SR.[11]

comprehensivetofindalltheavailableevidencethatmeetsthe SReligibilitycriteria.Toachievethis,reviewersuseelectronic databasesaswellasothersourcesofinformation.Thelargest generalelectronicdatabasesareMedline,fromtheUnited States'NationalLibraryofMedicine,andEmbase,fromthe Netherlands.Reviewerssearchthesedatabasesthroughsearch enginessuchasPubMed()and Ovid()ideally,duetothespecialized languagenecessarytosearchoptimally,withthehelpofan informationspecialist.Thereareotherelectronicdatabases focusingeitheronspecificdisciplinesorspecificgeographical areasthatmayadditionallybesearched.Reviewersshouldalso searchforeligiblestudiesthathavenotbeenpublishedinpeer reviewedjournalsingrayliteraturedatabasesandforongoing orpotentiallyunpublishedtrialsintrialsregistries.Thevalueof searchingforandincludingpre-printpublicationsisstillbeing

| Question component | Eligibility criteria |
| --- | --- |
| P: people at risk of developing a viral respiratory illness I: physical interventions | |
| C: no physical intervention or another physical intervention O: effects and complications | |

We included “people of all ages” in the hospital or community.

We included studies assessing “investigating physical interventions or combinations of interventions to prevent respiratory virus transmission.” In addition, “The interventions of interest included: screening at entry ports, isolation, quarantine, physical distancing, personal protection (clothing, gloves, devices), hand hygiene, face masks, gargling, nasal washes, eye protective devices, face shields, disinfecting, and school closure.”

S: study design

We included “doing nothing or with other intervention)”

Effects included numbers of cases of viral respiratory illness (including acute respiratory infections (ARI), influenza-like illness (ILI), COVID-like illness and laboratory-confirmed influenza, SARS-CoV-2 or other viral pathogens); adverse events related to the intervention; deaths; severity of viral respiratory illness as reported in the studies; absenteeism; hospital admissions; and, complications related to the illness, (eg pneumonia).

“We only considered individual-level randomized controlled trials (RCTs), or cluster-RCTs, or quasi-RCTs for inclusion.”

**Table 3.** Information sources typically used in systematic reviews.

| Information source | URL of search engine | Description |
|---|---|---|
| **General medical literature databases** | | |
| Medline | | More than 29 million records,[18] covering healthcare in general (medicine, nursing, dentistry, pharmacy, etc.) |
| Embase | | Approximately 40 million records as of December 2020, and growing at over 1.7 million records per year.[19] Covers all disciplines of medical and biomedical sciences. |
| Cochrane Central Register of Controlled Trials (CENTRAL) | | Approximately 2 million records. Focuses on randomized and quasi-randomized controlled trials.[20] |
| **Specialized medical literature databases** | | |
| PsycINFO | | Approximately 5.5 million records, focusing on behavioral and social sciences[21] |
| Allied and Complementary Medicine Database (AMED) | | Indexes approximately 500 journals from the fields of complementary and alternative medicine, physiotherapy, rehabilitation, speech and language therapy, etc. |
| Cumulative Index to Nursing and Allied Health Literature (CINHAL) | | Indexes approximately 3800 journals from nursing and allied health literature[22] |
| **Medical literature databases focused on geographic areas** | | |
| Latin American & Caribbean Health Sciences Literature (LILACS) | | Approximately 1 million records from journals in Latin-America and the Caribbean[23] |
| African Index Medicus (AIM) | | Approximately 25 000 records focusing on information published in Africa or related to Africa[24] |
| **Dissertation databases** | | |
| ProQuest Dissertations & Thesis Global | | Multidisciplinary database with approximately 5 million records[25] |
| **Study registries** | | |
| Clinicaltrials.gov | | Approximately 500 000 records of privately and publicly funded clinical trials |
| International Clinical Trials Registry Platform | | Registry network from the World Health organization[26] |

explored, but for most topics this may not be necessary. Table 3 provides a list and description of the most commonly used information sources.

Comprehensive searches vary across SR topics, but include at least one general database (ideally more), and other resources as appropriate. The specific minimum number of databases to search depends on the topic of the SR. For example, an SR address-ing psychological effects of an intervention should ideally, in addi-tion to general databases, search in PsycINFO, which specializes in behavioral sciences. Similarly, if reviewers suspect that there may be relevant literature that has not been published in peer reviewed journals, they should search for gray literature (eg, through Google Scholar or databases such as OpenGray) or for ongoing studies or studies completed and not published (eg, through the trial registries).

After conducting the search in all information sources, review-ers gather all records and remove duplicates, and then proceed to screen the potentially eligible studies. Screening usually happens in 2 stages. First, reviewers make decisions about eligibility based on the title and abstract of the citations retrieved. Then, all the citations deemed relevant move to full text screening and reviewers, using the full text of the articles, determine if these meet eligibility criteria; if so, the review team includes that study in the review. To minimize errors, pairs of reviewers, working independently, conduct screening. If there are disagreements, they usually resolve the issue by discussion or with the help of a third reviewer.

***Abstracting data and assessing risk of bias of included studies***
Afterselectingstudies,reviewersabstractrelevantdatafrom allincludedstudies.Theyabstractallthedatanecessaryto answer their question of interest. In SRs of exposure or interventions,thistypicallyincludescharacteristicsofthestudydesign, participantsincluded(age,sex,comorbidities,etc.),exposureor intervention(doseormagnitudeofexposure, andmodeofexposureoradministration), comparison(whatwasthecomparison), andoutcomes(howtheyweremeasured, andspecific resultsfor the timepoints of interest).

Since there may be issues that decrease the trustworthiness of the results from a study (ie, the result may overestimate or underestimate the effect), reviewers assess the risk of bias of each study. There are specific tools that facilitate this assessment, such as the Risk of Bias 2.0 tool[27]for randomized controlled trials, the Risk Of Bias in Nonrandomized studies of Interventions (ROBINS-I),[28]the QualityAssessmentofDiagnosticAccuracyStudies(QUADAS)-2tool,[29]andtheQualityinPrognosisStudies(QUIPS)tool.[30] Allthesetoolshaveasetofpromptingquestionsthataddress differentaspectsofthestudydesignandconduct(ie,domains), whichleadtoajudgmentoftheextentofriskofbiasineach domain and can also be used to makea judgment about the risk ofbiasofastudyoverall.Oneofthecomplaintsoftenconveyed

aboutthesetoolsisthattheyseemsubjective. Infact, reviewers mustjudgehowlikelybiasis,butusingatoolensuresthatthe reviewer thoroughly understands and considers all the important and potential reasons a result may not be accurate due to bias. In addition, tominimizeerrors, reviewersconductdataabstraction and risk of bias assessments in duplicate and independently, and then compare their results. Training and calibration exercises are

TheGRADEapproachclassifiesthecertaintyoftheevidence ashigh,moderate,low,orverylow.[34]GRADEprovidesaframe- workandguidanceformakingthisassessmentbyconsidering limitationsinstudydesign,[35]inconsistency,[36]imprecision,[37,38] also helpful.

***Summarizing the data***

Although all previous steps are conducted using primary stud-ies as the unit of interest, systematic reviewers should synthesize evidence by outcome and across studies. While some reviews will simplypresenttheresultsfromeachstudyseparately,decision- makersneedasummaryoftheevidenceforeachoutcome(eg, whatdoallstudiessayaboutwhathappenstothetransmission of respiratory illnesses?).

These summaries of bodies of evidence per outcome include all eligible studies that report results about the outcome. Whenever possible(thatis,whenitisappropriatebecausethestudiesare addressing the same overall PICO question and there is sufficient outcome data reported), reviewers should conduct meta-analyses. Meta-analyses allow combining the results from 2 or more studies addressing the same outcome and obtaining a summary or pooled estimate.[1]Therearesevеralstatisticalmethodsandapproaches to meta-analyses,[31]and therefore reviewers will seek advice from methodologists or biostatisticians with expertise in evidence syn- thesis when deciding which specific statistical methods to use.

Inconductingmeta-analyses,reviewersmayfinddifferent resultsacrossstudies,andsometimesthosedifferencesare large.Whetherlargeornot,reviewerswilltrytoexplainthe heterogeneitybytestingthesubgrouphypothesesthatthey developedwhenplanningtheirreviewthrougha“subgroup analysis,” in which they separate the studies based on the variable that they hypothesized may cause the differences. If, for example, the relative effects of 2 types of masks are different across studies, reviewerswillseparatestudiesaccordingtowhetherthesewere doneinadultsvschildrenandevaluatetowhatextentthis subgroupvariable(ie,age)explainstheheterogeneity(inother words,towhatextentthisvariableisaneffectmodifier). Ifthey findapotentialexplanationforheterogeneity,theywillapply criteria—ideally using a systematically developed and pilot tested instrumentthatisavailable—todeterminethecredibilityofthe effect modification.[32]

Whenmeta-analysisisnotpossibleduetoissueswiththe reportingofincludedstudies,reviewerswillsummarizethe resultsacrossstudies,attheoutcomelevel,inadescriptiveor narrativemanner.Someoptionsforthesesummariesinclude providingrangesofmagnitudesofeffectorclassifyingstudies accordingtotheirdirectionofeffect,magnitudeofeffect,or indirectness,[39]and publication bias.[40]GRADE assessments should (and can) be done for all synthesis of evidence (for each outcome); that is, whether there is a meta-analysis, a descriptive or narrative

synthesis,orasinglestudyprovidingevidenceforaspecific

outcome.

Assessments of the certainty of evidence are then formally incorporated into conclusions regarding specific outcomes.[41] When there is high certainty evidence, reviewers can appropriately describe that, for instance, an association or an effect is present; whereas when the certainty of evidence is low, it is more appropriate to say that there may be an effect.

***Other considerations- Registration of SRs***

To minimize the chances of bias, SRs are encouraged to register their SR and publish a protocol in databases like PROSPERO () or as a full publication that describes details of the methods that they will follow. Registrations should happen before reviewers begin data abstraction from included studies.

## Tools to facilitate the design, conduct, and reporting of systematic reviews

There are several tools available to facilitate the design, conduct, and reporting of SRs. The "SystematicReviewToolbox" is a free-access online platform that gathers a list of all these tools.[42] The following is a list of the most important resources:

***Tools that provide a description of the methods of SRs***

The Cochrane Handbook provides details about the methodology of all the steps of an SR.[10] Cochrane is a not for profit international network and one of the first organizations to develop and publish SRs. It is one of the leaders in the development of methods for SRs. The online version of the Cochrane Handbook is open access.

***Tools that guide the reporting of SRs***

The Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) checklist contains a list of 27 items that those publishing SRs should describe to ensure that their review is reproducible, transparent, and includes all relevant details.[43] This tool should not be confused with a tool or manual for how to conduct an SR (eg, the Cochrane Handbook), the items in PRISMA address reporting for the whole manuscript (title, abstract, introduction, methods, results, discussion, and other information). There are also extensions of the checklist focusing on SRs with specific types of data analysis, types of questions, and other types of evidence synthesis. All checklists are available for free on their

statistical significance.[33] All these methods, however, have potential limitations that reviewers should acknowledge when interpreting their results.

***Assessing the certainty of the evidence and drawing conclusions*** Bodies of evidence can have several limitations, which must be considered when drawing conclusions from an SR. When presenting a summary of results from SRs, users will ask "how certain is it that this is the effect?"—which is different from "what is or how large is the effect?" The certainty of the evidence reflects how confident reviewers are that the summary effect estimate is close to the truth.[34] The most widely used method to assess the certainty of evidence is the Grading of Recommendation Assessment, Development, and Evaluation (GRADE) approach.

website ().

***Tools that aid in the conduct of specific steps of SRs***

Therearesevеraltoolstousewhendevelopingsearchstrate-gies, selecting articles, abstracting data, conducting data analysis, andpreparingtablesthatsynthesizetheevidence.The“Sys-tematicReviewTool”containsalistofallofthese.Themost commonlyusedtoolsincludethosetoaidwiththeselectionof studies,suchasCovidence[44]andDistiller[45];dataanalysis,such asDataParty[46]andtraditionalstatisticalsoftware(including R[47] and Stata[48]); and tools to create summary of findings tables, such as GRADEpro[49]and MAGICapp.[50]Access to these tools varies, with most having free access to some of the basic features.

***Tools to assess existing SRs***

Users of SRs can benefit from using tools to appraise the extent to which SRs have been optimally conducted. The A Measurement Tool to Assess Systematic Reviews (AMSTAR) 2 tool[51] is a 16-item checklist that aims to assess the "credibility of the results" of an SR and was designed for users with little background or training in SR methods. The Risk Of Bias in Systematic reviews (ROBIS) tool was designed for guideline developers and systematic review authors to assess the risk of bias of existing, or their own SRs, through the consideration of 21 items organized in 4 domains that lead to overall judgments. Finally, there are checklists designed for clinicians to judge the risk of bias, importance of the results, and applicability to their specific contexts—the most used ones being

that from the JAMA Users' Guides to the Medical Literature.[52]

## Conclusions

Systematicreviewsareatypeofevidencesynthesisusedto answer different types of specific research questions by collecting and appraising all the available evidence relevant to that question. Researchersplanningtoconductan SR should assess if their questionisappropriateforthisstudydesign,andifso,should follow appropriate methods that minimize biases and errors, and thatmaketheresultsandconclusionsreproducible.Thereare severaltoolsthatauthorsofSRscanusetoaidthisprocess, butbothmethodologicalandcontentexpertisearefundamental when conducting an SR.

## Conflict of interest

The authors do not have any financial conflict of interest.


8. PollockD,DaviesEL,PetersMDJ,etal.Undertakingascoping review:apracticalguidefornursingandmidwiferystudents,clinicians,researchers,andacademics.*JAdvNurs*. 2021;77(4):2102-2113.
9. Galvez-HernandezP,González-dePazL,MuntanerC.Primarycare-basedinterventionsaddressingsocialisolationand lonelinessinolderpeople:ascopingreview.*BMJOpen*.2022; 12(2):e057729.
10. JPT H, Thomas J, Chandler J, et al. *Cochrane Handbook for System- atic Reviews of Interventions* version 6.1. Cochrane; 2020. Available from www.training.cochrane.org/handbook
11. JeffersonT,DooleyL,FerroniE,etal.Physicalinterventions tointerruptorreducethespreadofrespiratoryviruses. *Cochrane Database Syst Rev*. 2023; (1):CD006207.
12. KalotMA,HusainatN,ElAlayliA,etal.vonWillebrandfactor

levelsinthediagnosisofvonWillebranddisease:asystematic review and meta-analysis. *Blood Adv*. 2022;6(1):62-71.

13. HaoQ,ForoutanF,HanMA,etal.Prognosisofolderpatients withnewlydiagnosedAMLundergoingantileukemictherapy: a systematic review. *PloS One*. 2022;17(12):e0278578.

14. Deng JA-O, Zhou FA-OX, Hou W, et al. The prevalence of depression,anxiety,andsleepdisturbancesinCOVID-19patients:a meta-analysis. *Ann N Y Acad Sci*. 2021;1486(1):90-111.

15. GarrittyC,GartlehnerG,Nussbaumer-StreitB,etal.Cochrane rapidreviewsmethodsgroupoffersevidence-informedguidancetoconductrapidreviews.*JClinEpidemiol*.2021;130(2021): 13-22.

16. MuradMA-O,WangZA-O,ChuH,etal.Proposedtriggers for retiring a living systematic review. *BMJ Evid Based Med*. 2023;28(5):348-352.

17. SiemieniukRA,BartoszkoJJ,ZeraatkarD,etal.Drugtreatments for covid-19: living systematic review and network meta-analysis. *BMJ*. 2020;370:m2980

18. National Library of Medicine. MEDLINE: Overview. Accessed May 15, 2023.

## References

1. GuyattG,RennieD,MeadeM,etal.*Users'GuidestotheMedical Literature: a Manual for Evidence-Based Clinical Practice*. 3rd ed. New York,NY:McGraw-HillEducation,AmericanMedicalAssocia-tion; 2015.

2. Brignardello-Petersen R, Carrasco-Labra A, Guyatt GH. HowtoInterpretandUseaClinicalPracticeGuidelineor Recommendation:Users'GuidestotheMedicalLiterature.

19. Embase Indexing Guide 2021: A comprehensive guide to Embase IndexingPolicy;2021.Availablefrom
*JAMA*.2021;326(15):1516-1523.
15319
3. TakeuchiM,ShowaS,KitazawaK,etal.Livingaloneisasso-ciatedwithanincreasedriskofinstitutionalizationinolder men:afollow-upstudyinHamanakatownofHokkaidoJapan. *GeriatrGerontolInt*.2018;18(6):867-872.

4. ZhaoY,GuyattG,GaoY,etal.Livingaloneandall-cause mortalityincommunity-dwellingadults:asystematicreview and meta-analysis. *EClinicalMedicine*. 2022;54:101677.

5. Gough D, Davies P, Jamtvedt G, et al. Evidence synthesis interna-tional(ESI):positionstatement.*SystRev*.2020;9(1):155.
6. CrewdsonJ.Theeffectoflonelinessintheelderlypopulation: areview.*HealthyAgingClinCareElder*.2016;8:1-8.

20. Cochrane.Cochrane Central Register of Controlled Trials (CENTRAL). 2023. Accessed May 15, 2023.

21. American Psychological Association. APA PsycINFO. 2023. AccessedMay15,2023.

22. EBSCO. CINHAL Database. 2023. Accessed May 20, 2023.
23. Latin American and the Caribbean Literature on Health Sciences (LILACS). 2023. Accessed May 20, 2023.
24. About Index Medicus. 2023. Accessed May 20, 2023. "
25. ProQuest Dissertations & Thesis Global. 2023. Accessed May 21, 2023.

7. Greenhalgh T, Thorne S, Malterud K. Time to challenge the spurious hierarchy of systematic over narrative reviews? *Eur J Clin Invest*. 2018;48(6):e12931.

26. About the WHO ICTRP. 2023. Accessed May 25, 2023. “ ”

27. SterneJACSJ,PageMJ, ElbersRG, etal. RoB2:arevisedtoolfor assessingriskofbiasinrandomisedtrials. *BMJ*. 2019;366:l4898.

28. SterneJA,HernanMA,ReevesBC,etal.ROBINS-I:atoolfor assessingriskofbiasinnon-randomisedstudiesofinterven- tions. *BMJ*. 2016;355:i4919.

guidance34:updateonratingimprecisionusingaminimallycontextualizedapproach. *JClinEpidemiol*. 2022;150(2022):216-224.

29. WhitingPF,RutjesAW,WestwoodME,etal.QUADAS-2:arevisedtoolforthequalityassessmentofdiagnosticaccuracystudies.*AnnInternMed*.2011;155(8):529-536.

39. GuyattGH,OxmanAD,KunzR,etal.GRADEguidelines:8.Ratingthequalityofevidence–indirectness.*JClinEpidemiol*.2011;64(12):1303-1310.014

40. Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence—publication bias. *J Clin Epidemiol*. 2011;64(12):1277-1282.
30. Hayden JA, van der Windt DA, Cartwright JL, et al. Assessing bias in studies of prognostic factors. *Ann Intern Med*. 2013;158(4):280-286.
31. Deeks J, Higgins J, Altman D. Chapter 10: Analyzing data and undertaking meta-analyses. In: Higgins J, Thomas J, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.3 ed. Cochrane Training; 2022.

41. Santesso N, Glenton C, Dahm P, et al. GRADE guidelines 26: informative statements to communicate the findings of systematic reviews of interventions. *J Clin Epidemiol*. 2020;119(2020):126-135.

32. Schandelmaier S, Briel M, Varadhan R, et al. Development of the instrument to assess the credibility of effect modification analyses (ICEMAN) in randomized controlled trials and meta-analyses. *CMAJ*. 2020;192(32):E901-E906.

42. Marshal C, Sutton A, O'Keefe H, et al. The Systematic Review Toolbox. Accessed May 15, 2023. www.systematicreviewtools.com

43. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.

33. McKenzie J, Brennan S. Chapter 12: Synthesizing and presenting findings using other methods. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, Welch V, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.3 ed. Cochrane

44. *Covidence systematic review software*. Melbourne, Australia: Veritas Health Innovation.

45. *Evidence Partners Distiller SR*. Ottawa, Canada.

Training; 2022.
34. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011;64(4):401-406.
35. GuyattGH,OxmanAD,VistG,etal.GRADEguidelines: 4.Ratingthequalityofevidence—studylimitations(riskof bias). *J Clin Epidemiol*. 2011;64(4):407-415.
46. Data Party. Accessed May 15, 2023.
47. RCoreTeam.*R:Alanguageandenvironmentforstatisticalcomputing*. Vienna, Austria: R Foundation for Statistical Computing; 2022.
48. StataCorp.In:StationC,ed.*Statastatisticalsoftware:release18*. TX: StataCorp LLC; 2023.

49. McMasterUniversityandEvidencePrimeI.*GRADEproGDT: GRADEpro Guideline Development Tool*. McMaster University; 2022.
36. GuyattGH,OxmanAD,KunzR,etal.GRADEguidelines:7. Ratingthequalityofevidence–inconsistency.*JClinEpidemiol*. 2011;64(12):1294-1302.
017
37. SchünemannHJ,NeumannI,HultcrantzM,etal.GRADE guidance 35: update on rating imprecision for assessing contextualized certainty of evidence and making decisions. *J Clin Epidemiol*. 2022;150(2022):225-242.
38. ZengL,Brignardello-PetersenR,HultcrantzM,etal.GRADE
Available from
50. MagicEvidenceEcosystemFoundation. MAGICapp. 2023. Available from
51. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal toolforsystematicreviewsthatincluderandomisedornon-randomisedstudiesofhealthcareinterventions,orboth.*BMJ*. 2017;358:j4008.
52. MuradMH,MontoriVM,IoannidisJP,etal.Howtoread asystematicreviewandmeta-analysisandapplytheresults to patient care: users' guides to the medical literature. *JAMA*.2014;312(2):171-179.
5559