MethodsX 12 (2024) 102532

Contents lists available at ScienceDirect

MethodsX

journal homepage: www.elsevier.com/locate/methodsx

Systematic review of clinical practice guidelines and systematic reviews: A method for conducting comprehensive analysis

Stefano Mancin[a,b], Marco Sguanci[c], Desirèe Andreoli[d], Fanny Soekeland[e], Giuliano Anastasi[f], Michela Piredda[c,*], Maria Grazia De Marinis[c,g]

[a] Department of Biomedicine and Prevention, University of Rome "Tor Vergata", Rome, Italy
[b] IRCCS Humanitas Research Hospital, Rozzano, Milan, Italy
[c] Department of Medicine and Surgery, Research Unit of Nursing Science, University Campus Bio-Medico, Rome, Italy
[d] Azienda Ospedaliera Santa Maria della Misericordia, Perugia, Italy
[e] University of Applied Sciences, School of Health Professions, Bern, Switzerland
[f] Department of Trauma, AOU G. Martino University Hospital, Messina, Italy
[g] Fondazione Policlinico Universitario Campus Bio-Medico, Rome, Italy

REVIEW HIGHLIGHTS

- Secondary scientific literature encompasses various sources, including clinical practice guidelines (CPG) and systematic reviews (SR), both of fundamental importance.
- Integrating CPGs and SRs in a single systematic review ensures a comprehensive and updated perspective on clinical evidence.
- The implemented methodology ensures a stringent methodological approach, placing significant emphasis on both clarity and reproducibility.

ARTICLE INFO

Method name:
Systematic review of clinical practice guidelines and systematic reviews

Keywords:
Evidence-based medicine
Clinical practice guidelines
Systematic reviews
Meta-analysis
Methodology

ABSTRACT

A systematic review (SR) is a research method for synthesizing evidence on a specific topic. Among the various types of systematic reviews, there are SRs of guidelines (CPGs) and SRs of SRs. Traditionally, they are limited to just one type of secondary evidence. This paper introduces an innovative SR methodology that combines CPGs and SRs to improve evidence synthesis and overcome the limitations of isolated use.

Essential steps that should always precede the actual research process include registering the research protocol, formulating research questions and setting inclusion/exclusion criteria. Using the PRISMA protocol for comprehensive database searches, it's crucial to combine keywords with boolean operators and remove duplicates. The eligibility of studies should be assessed by selecting potentially relevant articles through an initial screening of titles and abstracts, followed by a meticulous analysis of the full-texts. Rigorous evidence evaluation tools, such as AGREE II for CPGs and AMSTAR 2 for SRs, and the double reviewer approach ensure high-quality selections. Additionally, converting summarized results into percentages and applying statistical analyses facilitate interpretation and improve the reliability of rater assessments. A further characteristic of this methodology is its adaptability to the evolution of healthcare research.

* Corresponding author.
E-mail address: m.piredda@unicampus.it (M. Piredda).

https://doi.org/10.1016/j.mex.2023.102532
Received 21 September 2023; Accepted 20 December 2023
Available online 21 December 2023

Specifications table

Subject area:	Medicine and Dentistry
More specific subject area:	*Evidence Based Medicine (EBM)*
Method name	*Systematic review of clinical practice guidelines and systematic reviews*
Name of the reviewed methodology:	Systematic review
Keywords:	Evidence-based medicine; Clinical practice guidelines; Systematic reviews; Meta-analysis
Resource availability:	*Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. International journal of evidence-based healthcare* 10.1097/XEB.0000000000000055 [23] *Methodology in conducting a systematic review of systematic reviews of healthcare interventions.* 10.1186/1471–2288–11–15">10.1186/1471–2288–11–15 [15]
Review question:	1. How can a systematic review methodology effectively integrate both Clinical Practice Guidelines (CPGs) and Systematic Reviews (SRs) to provide a more comprehensive framework of available evidence? 2. What are the specific challenges and gaps in evidence addressed by the proposed methodology, considering the limitations of relying solely on CPGs or SRs? 3. In what ways does the research method ensure the validity, transparency, and reproducibility of the systematic review process, particularly in the selection, evaluation, and synthesis of CPGs and SRs? 4. How does the proposed methodology address the potential limitations of CPGs, such as the omission or inadequate evaluation of certain SRs, and how does it incorporate more recent SRs that may not be considered in existing guidelines? 5. What are the key advantages and strengths of integrating CPGs and SRs in the systematic review methodology, and how does this integration contribute to bridging potential gaps in evidence for improved relevance and utility in clinical practice?

Method details

Background

The body of scientific biomedical literature is characterized by a vast amount of secondary bibliographic records, including Narrative Reviews, Scoping Reviews, various types of Systematic Reviews (such as Cochrane Reviews, Umbrella Reviews, and Network Meta-analyses), Meta-analyses, Clinical Practice Guidelines, Scholarly Books and Monographs, Conference Proceedings. These are derived from various sources, including bibliographic databases, grey literature, and guideline databases. Bibliographic biomedical databases, such as PubMed, Embase, CINAHL, Scopus or Web of Science, form the backbone of scientific information access, aggregating bibliographic records from a wide range of biomedical journals [1]. Each record contains key metadata such as authors, titles, abstracts, and keywords, which facilitate the search for relevant studies [2]. Grey literature, that can be found in websites such as Google Scholar or databases such as Open Grey, represent a broad category of formally unpublished documents, such as theses, research reports, conference papers, and more [3]. Although not subject to the typical peer-review process of scientific publications, grey literature can offer a significant complementary perspective, reporting ongoing studies, preliminary data, or studies that have not been accepted for publication in peer-reviewed journals [4]. Guideline databases, such as the National Guideline Clearinghouse or the Scottish Intercollegiate Guideline Network database, gather clinical guidelines developed based on the best available evidence [5]. Such guidelines represent a fundamental resource for clinicians, offering recommendations for clinical practice in specific areas of interest [6].

All these sources, together with primary literature, are crucial for conducting a systematic review [7]. They provide the groundwork for synthesizing and analyzing specific research questions or areas of study [8].

A systematic review is a research method that comprehensively collects and rigorously analyses all empirical evidence that meet predefined eligibility criteria to answer a research question [9]. This meticulous, detailed approach ensures the reliability and completeness of the data extracted, making it a key method in the field of evidence-based medicine [10]. Systematic reviews are often used to identify, evaluate, and synthesize the best available evidence on a particular topic or clinical question. [11]. The ultimate purpose of such practice is to provide healthcare professionals with a clear and reliable overview of current knowledge, on which to base their clinical decisions [12]. There are different types of systematic reviews, each showing particular relevance depending on the research context and objectives [13]. The most common types are systematic reviews focusing on primary literature, while systematic reviews of systematic reviews and systematic reviews of guidelines, although less prevalent, also contribute to the literature. Systematic reviews of primary literature are the best known and focus on the analysis of primary studies, such as randomized clinical trials and observational studies. Such reviews represent an important resource in the quantitative synthesis of existing evidence on a specific clinical question [14].

Systematic reviews of systematic reviews, sometimes called “second-level reviews” or “meta-reviews,” go a step further [15]. They analyze and synthesize the results of multiple systematic reviews on a similar topic, to provide a more comprehensive and exhaustive picture [16]. This type of review is particularly useful when there are numerous reviews on the same issue, and an overview of available evidence is desired [17].

Finally, systematic reviews of guidelines focus on the analysis of clinical guidelines [18] Such guidelines are often drawn up by professional or government organizations and provide evidence-based recommendations for clinical practice in specific fields [19]. Systematic reviews of guidelines can help to identify the best clinical recommendations and understand how these differ among various sources [20].

Despite the importance and usefulness of these types of systematic reviews, there are specific issues when conducting systematic reviews using only secondary sources. For certain research questions, it may be limiting to conduct a systematic review based solely on guidelines or only on existing systematic reviews. This is because these sources may not cover all aspects of a topic or may be influenced by methodological biases, thereby limiting the completeness and reliability of the review results [21].

To overcome this hurdle, it is necessary to consider implementing a new type of systematic review, one that can leverage the potential of both existing guidelines and systematic reviews. Such an approach can provide a more comprehensive and accurate picture of available evidence, overcoming the limitations of current systematic review practices.

Research objective

The aim of this systematic review methodology that, not only includes Clinical Practice Guidelines (CPGs) or Systematic Reviews (SRs) but combines them. This integration is based on the necessity to optimally combine evidence from secondary literature. CPGs represent a robust reference point for clinical practice, providing evidence-based recommendations [22]. However, they may not include or adequately evaluate all available SRs. Similarly, more recent SRs may not yet have been considered in the guidelines. Consequently, there is a need to develop a systematic review methodology that can effectively integrate these sources. With this research method, we propose an approach that allows the integration of SRs not included or not evaluated by CPGs, or even those more recent than a CPG. The goal is to use the complementary strengths of CPGs and SRs to provide a more comprehensive and updated framework of the available evidence. This integration will bridge any potential gaps in the evidence and incorporate the most recent information, thereby improving the relevance and utility of systematic reviews for clinical practice. The methodology proposed to perform this method will be structured in such a way as to ensure the validity, transparency, and reproducibility of the review process. This research will detail the procedures for the selection and evaluation of CPGs and SRs, data extraction and synthesis, and interpretation and presentation of results.

Methods

Research design

In an attempt to provide a thorough synthesis of existing evidence derived from secondary literature, we have structured a methodology for a systematic review of guidelines as well as of systematic reviews. This research design draws upon the methodology proposed by Aromataris et al. [23] and further integrates the methodology by Smith et al. [15] to encompass systematic reviews not included in the identified guidelines. The methodology used in this review underlines the significance of comprehensive literature search and rigorous evaluation, paying special attention to its methodological quality and relevance. Data is uniformly extracted and subsequently synthesized to offer an all-encompassing overview.

Systematic review protocol registration

Prior to embarking on the systematic review drafting, one should register the protocol in the international PROSPERO database (https://www.crd.york.ac.uk/prospero/) of the National Institute of Health Research. This registration ensures the review is conducted with transparency and integrity.

Formulation of the research question

Current methodologies in systematic reviews: a dual focus on SRs of CPGs and SRs of SRs

In addressing a systematic review of Clinical Practice Guidelines (CPGs), it is crucial to clearly structure and focus the research question. To achieve this, the utilization of the PICAR framework is recommended [18,24]. This framework, developed to guide systematic reviews of CPGs for clinical practice, is divided into: P: Population, clinical indications, and conditions; I: Intervention; C: Comparator, comparisons, and key contents; A: Eligible CPG Attributes. Attributes may include variables such as the year of publication, language of publication, publishing organization or sponsor, scope (national or international), clinical focus/focus, purpose (e.g., screening status nutritional), and the format of the CPG. Additionally, attributes may include the intended end user, such as healthcare professionals, and the type of CPG, if evidence-based, as well as specific score thresholds for assessing the quality of CPGs using tools such as AGREE II; R: Characteristics of the recommendations. This section specifies CPG recommendations based on the research question.

Secondly, in conducting a systematic review aimed at identifying relevant systematic reviews, the research question should be formulated using the PICO framework [25]. Utilizing the PICO framework enables the structured and refined formulation of the research question, allowing for a specific and methodical examination of the effectiveness and appropriateness of an intervention in a given population. This ensures that the research question is explored comprehensively and accurately (Fig. 1). This methodological tool represents a key element in developing research questions for systematic reviews and focuses on the following key aspects: P: (Population or problem): Defines the patient group or population of interest, as well as the specific conditions or problems that these individuals may exhibit [26]. For example, it could regard patients suffering from a specific disease or medical condition; I: (Intervention): This refers to the action or series of actions that one plans to study. It can be a medical treatment, procedure, therapy, or any other type of clinical intervention. The goal is to clearly identify which intervention will be examined to determine

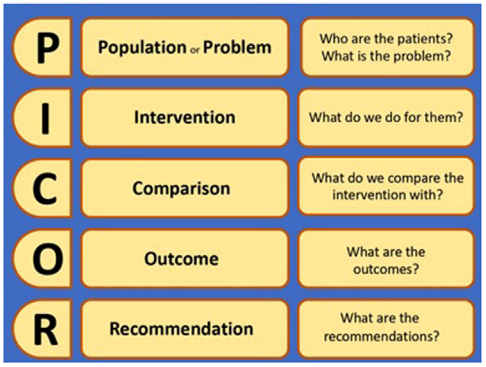

Fig. 1. Main Elements of PICOR Framework.
Legend: Methodology and phases of the PICOR (Patient/Population, Intervention, Comparator, Outcome, and Recommendations).

its effectiveness or impact; C: (Comparator): Serves to identify a control condition with whom the intervention is compared. The comparator may be no intervention, a placebo, or another type of intervention deemed standard in clinical practice; O: (Outcome): Indicates the outcomes, expected outcomes, or results of the intervention [27]. These outcomes can concern aspects such as mortality, quality of life, side effects, clinical benefits, and many others.

Introducing PICOR: research question framework for CPGs and SRs

The evolution of systematic reviews of Clinical Practice Guidelines (CPGs) and systematic reviews (SRs) necessitates an innovative and integrated methodological approach. The proposal of a new methodological tool, named PICOR (Population, Intervention, Comparator, Outcome, and Recommendations), arises from the need to amalgamate the distinctive approaches of the PICO and PICAR frameworks into a cohesive structure.

The PICAR framework, specifically designed for SRs of CPGs, provides a solid foundation for outlining populations, interventions, comparators, and eligible attributes of CPGs. However, the innovation of PICOR lies in integrating these components with the PICO framework, traditionally applied to formulate research questions in broader systematic review contexts.

The 'PICOR' research question represents a significant step forward, incorporating the specificities of CPGs and SRs. Particularly, the 'A' component of the PICAR framework, indicating eligible attributes of CPGs, becomes an integral part of inclusion criteria, emphasizing the need to carefully assess these attributes in analyses. This innovation not only addresses the call for a new type of research question but also underscores the importance of CPG attributes in the context of a broader systematic review. The PICOR approach thus emerges as a unified and advanced framework, providing a methodological structure that considers the specificities of CPGs and SRs, with a particular emphasis on the importance of the letter 'R' (Recommendations) in achieving coherence and comprehensiveness in the execution of comprehensive systematic reviews (Fig. 1).

Inclusion and exclusion criteria

When shaping the review scope, establishing clear inclusion and exclusion criteria is paramount. The focus will be on secondary studies published in peer-reviewed journals, with a specific emphasis on systematic reviews (SRs) and Clinical Practice Guidelines (CPGs). This approach aims to ensure the inclusion of only high-quality and relevant sources, avoiding potential distortions from records not subjected to rigorous quality controls. Additional inclusion criteria may encompass the sample size of the studies, adopted methodology, time frame covered by the study, and the geography or regional context of the analysis. For example, studies with larger sample sizes, standardized methodologies recognized in the field, and those covering specific time periods or geographic contexts of particular interest may receive prioritization. Publication language is another essential criterion in the selection process. To ensure comprehensibility and accessibility of data, only articles written in dominant languages in the field of medical research will be considered. On the other hand, specific criteria will guide the exclusion of studies. Articles not directly aligning with the review objectives or not meeting defined quality standards will be excluded. The presence of incomplete or irretrievable data in a publication may also justify its exclusion. Pilot or preliminary studies, publications not subjected to a rigorous peer-review process, and those relying exclusively on self-reported data without verification may also face exclusion. The inclusion of the letter A of the PICAR framework should be considered in defining the inclusion and exclusion criteria, aligning with the review objectives.

Article selection will be conducted objectively, with two reviewers independently assessing each publication, and the inter-reviewer reliability of selections should be reported. This dual control aims to ensure that every decision is based on a thorough assessment and is not influenced by individual biases. In case of discrepancies in the selection, a third reviewer will be involved to ensure consensus and maintain the integrity of the selection process.

Search strategy

A meticulous analysis should be conducted based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) methodology. Adopting PRISMA ensures that the systematic review is performed and reported with precision and transparency. This initial step will be crucial for gaining an in-depth understanding of current practices and official recommendations in the study field. Using PRISMA guidelines, one can also outline the inclusion and exclusion criteria, with the reporting conducted in accordance with the standards outlined by the Equator Network (https://www.equator-network.org/). Subsequently, a comprehensive and systematic search should be undertaken in primary, secondary, and guideline literature databases, combining the keywords identified from the research question using Boolean operators AND, OR, NOT. This search should be conducted by at least two researchers, and the inter-reviewer reliability of selections should be reported to ensure objectivity and reduce the risk of bias. The collaboration of two researchers will allow for comparing and contrasting the results, thus enhancing the reliability of the selection process. Once the database search is completed, the records should be exported to a bibliography management software, such as for instance EndNote 20 (©2023 Clarivate). Then, through the software's automatic duplicate search features, duplicates should be removed, also considering an additional manual check. This step will ensure a clean record set free from repetitions. After this initial phase, titles and abstracts should be examined to select potentially relevant articles. When determining the eligibility of publications, the full texts of the identified articles should be scrutinized meticulously. The entire identification and selection process must be conducted independently by at least two researchers; in case of disagreements, a third reviewer should be involved for an arbitrary decision. Lastly, to ensure the utmost transparency and reproducibility of the conducted research, all used search strategies must be documented, including specific search strings, making them available in a supplementary document (Fig. 2). A search string template is made available in Supplementary File 1.

Evaluation of the quality of CPGs

To facilitate a meticulous evaluation of CPGs, it is imperative to use a standardized instrument that captures various dimensions of guideline quality. The AGREE II tool, designed by the AGREE Collaboration, is an example of such an instrument. The AGREE II is not a mere compilation of items but is a well-structured instrument composed of 23 accurately developed items, spanning across six domains. Each of these domains inspects a specific dimension integral to the guideline quality. The domains include: (1) Scope and Purpose; (2) Stakeholder Involvement; (3) Rigour of Development; (4) Clarity of Presentation, (5) Applicability; (6) Editorial Independence. Additionally, the AGREE II tool supplements [28] these domains with two global evaluation items, enhancing the depth of the assessment. This comprehensive assessment ensures that the guidelines in question adhere to the highest standards. In the evaluation process, each item within these domains is assigned a score based on a seven-point Likert scale. The methodology and criteria for scoring are well-outlined in the AGREE II user manual, which can be utilized for manual assessments. For those who prefer an online approach, the AGREE II is also available as online version known as My AGREE PLUS. The online version can be accessed at the following link: https://www.agreetrust.org/resource-centre/agree-plus/. For a concise description of how the online version works, please refer to the provided link.

Post evaluation, the domain score is derived by summing scores of individual items within the domain and then representing it as a percentage of the maximum possible score for that domain. The overall score, encompassing all the domains, provides an aggregate assessment of the guidelines. Fig. 3 illustrates how to calculate the score of AGREE II standardized to 100. Furthermore, it is essential to note that the reporting of AGREE II assessments should adhere to the guidelines provided by the AGREE Collaboration. The AGREE Reporting Checklist, available at the following link: https://www.agreetrust.org/resource-centre/agree-reporting-checklist/, outlines the essential elements to be included when reporting AGREE II assessments for optimal transparency and completeness.

Following the recommendation of researchers like Andrade et al. [29], any domain or overall scores that fall below the 50% threshold are typically indicative of low quality.

For ensuring unbiased and thorough assessments, our methodology involves two independent researchers to deploy the AGREE II tool for evaluations. The presence of two evaluators minimizes subjective biases and enhances the reliability of the evaluation. However, disagreements, though rare, can arise. In such instances, a third reviewer intervenes to reconcile the scores.

Incorporating such rigorous, methodical evaluations is not just about adhering to best practices. It is about ensuring that the CPGs that healthcare professionals depend upon are of the highest standard, both in terms of content and presentation. This rigorous approach to quality evaluation promises more than just academic rigor; it assures patients and professionals alike that the guidelines they adhere to have undergone scrupulous scrutiny (Fig. 4). A template of a table describing the results of such quality evaluation is provided in Supplementary File 1.

Evaluation of the quality of SRs

An accurate assessment of the quality of the SRs is essential to ensure the validity of the conclusions and the recommendations. For this assessment, the AMSTAR 2, or 'A MeaSurement Tool to Assess Systematic Reviews' is used [30]. This tool was selected for several key reasons: 1) Breadth of criteria: the AMSTAR 2 offers a broad range of questions ($n = 16$), split between seven critical and nine non-critical. This subdivision allows for a detailed and comprehensive evaluation; 2) Flexibility: the questions foreseen in the tool can be answered as 'Yes', 'Partly Yes', 'No' or 'No meta-analysis conducted', guaranteeing a precise assessment and methodological flexibility; 3) AMSTAR 2 is widely accepted in the scientific community as a reliable and valid tool for evaluating systematic reviews.

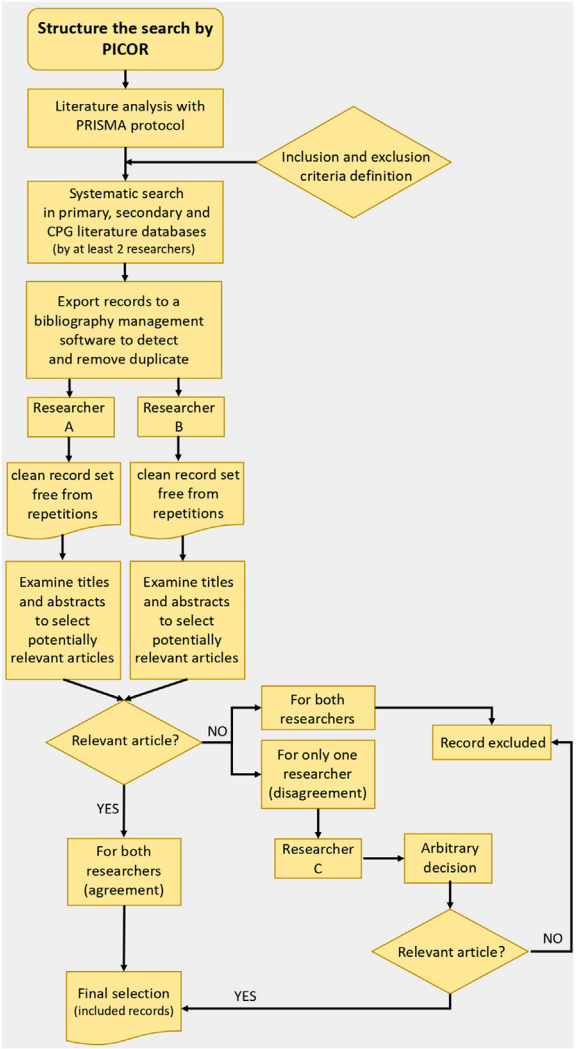

Fig. 2. Search Strategy methodology.
Legend: PICOR: patient/population, intervention, comparison, outcomes; R: Characteristics of the recommendations; CPG: Clinical Practice Guideline.

In the evaluation process, any research using this methodology should have at least two independent raters completing the AMSTAR 2 for each systematic review evaluated. This double assessment ensures greater objectivity. In the event of any disagreements between the evaluators, the involvement of a third reviewer is essential to ensure impartiality and consistency. A template of a table describing the results of such quality evaluation is provided in Supplementary File 1.

Conversion to percentage score

Once the assessment is complete, the results will need to be translated into a quantifiable and interpretable format. Following the methodology outlined by Matthias et al. [31], each question of the AMSTAR 2 receives a numerical score: 0 for the answer 'no', 1 for the answer 'yes' and 0.5 for the answer 'partial yes'. [31] These values, once aggregated, can be transformed into a percentage score ranging from 0 to 100%. It is essential to note that, in the case of systematic reviews where a meta-analysis is not conducted, the percentage calculation should consider a reduction of the denominator, excluding non-applicable elements. (Fig. 5).

Assessment of evidence certainty

A meticulous assessment of evidence quality is crucial in a systematic review to ensure the grounding of conclusions in valid and reliable data. Acknowledging potential variations in international evidence-level classifications, we have opted for a distinct methodology for our review. This choice aims to address anticipated heterogeneity in classifications or potential gaps in evidence-level evaluations within SRs. For a standardized and cohesive assessment of evidence levels for both CPGs and SRs included in our

1. Calculation of the Domain Score:

For each domain in the AGREE II tool:

$$\text{Domain Score (DS)} = \left(\frac{\text{Obtained Score} - \text{Minimum Possible Score}}{\text{Maximum Possible Score} - \text{Minimum Possible Score}}\right) \times 100$$

2. Calculation of the Overall AGREE II Score:

Once you've calculated the score for each of the six domains, you can compute the overall score using the following formula:

$$\text{Overall AGREE II Score} = \frac{\sum_{i=1}^{6} DS_i}{6}$$

Fig. 3. Calculation of the AGREE II standardized score.
Legend: Obtained Score is the sum of the scores of all the items within the specific domain. Minimum Possible Score represents the minimum score obtainable for the domain. This is calculated as the number of items in the domain multiplied by the minimum item score (typically 1, since the scoring ranges from 1 to 7). Maximum Possible Score is the maximum score obtainable for the domain. It is calculated as the number of items in the domain multiplied by the maximum item score (typically 7). DS_i represents the Domain Score for the i th domain.

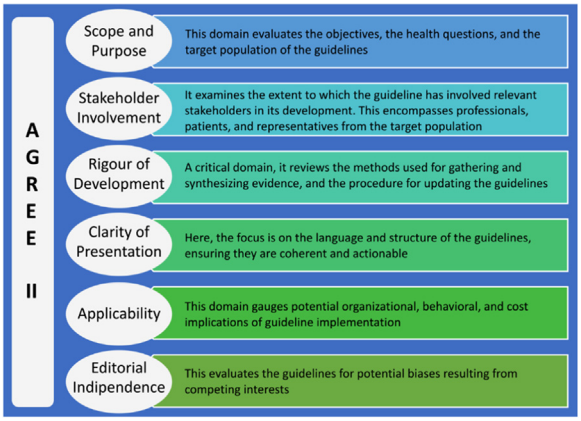

Fig. 4. Main Elements of AGREE II.
Legend: Methodology and characteristics of the AGREE II methodology (Appraisal of Guidelines for Research and Evaluation).

$$\text{Percentage Score} = \left(\frac{\text{Sum of Scores from Questions}}{\text{Maximum Number of Applicable Questions} \times \text{Maximum Score per Question}}\right) \times 100$$

Fig. 5. Calculation of the AMSTAR 2 score.
Legend: The "Sum of Scores from Questions" refers to the total sum of scores assigned to all answered questions; The "Maximum Number of Applicable Questions" is 16 by default but can be reduced if certain questions are not applicable (e.g., when a meta-analysis is not conducted); The "Maximum Score per Question" is 1.

review, we have embraced the classification scheme introduced by the Oxford Centre for Evidence-Based Medicine (OCEBM) in 2011 [32]. This methodology entails a thorough evaluation of evidence quality, relying on specific criteria for certifying the certainty of evidence.

Elements taken into account in this evaluation encompass study type, research design, evidence quality, and the consistency among the results of the included studies. This approach facilitates the assignment of a level of certainty to the evidence and enables a critical appraisal of the available information (Table 1).

Data extraction

In the context of a systematic review, the accuracy of data extraction is pivotal for ensuring the validity of obtained results. Thus, it is crucial to implement a structured form for data extraction, specifically designed to capture pertinent information from the selected studies for inclusion. The extracted data may encompass but are not limited to: author details; publication year; country of origin of the study; type of study (e.g., clinical trial, observational study, etc.); quality recommendations regarding CPGs and SRs; authors' organizational affiliations; target group intended for the recommendations or findings; grading system used in the study (if applicable). It is recommended that two independent reviewers handle data extraction to ensure accuracy and completeness. In cases of discrepancies in extracted data, a consensus approach can be applied, potentially involving a third reviewer. It is essential to integrate data and references from the included SRs and CPGs into tables or synthesis.

Table 1
OCEBM level.

Level 1: High-Quality Evidence
This level includes results from well-conducted systematic reviews of randomized controlled trials (RCTs) with extensive follow-up and consistent outcomes. Systematic reviews aggregate data from multiple studies to provide a more comprehensive and reliable view of the available evidence.
Level 2: Moderate-Quality Evidence
This category comprises results from moderately well-conducted RCTs or well-conducted observational studies with consistent findings. While these pieces of evidence offer a lower level of certainty compared to Level 1, they are still considered reliable.
Level 3: Low-Quality Evidence
This level includes results from studies with significant limitations, such as RCTs with methodological issues or observational studies with potential sources of bias. While they may provide useful insights, evidence at this level is less reliable.
Level 4: Low-Quality Evidence
This level encompasses results from studies with substantial limitations, such as observational studies with a high risk of bias or non-generalizable data. Evidence at this level is the least reliable among those considered.
Level 5: Expert Opinion
This category is based on expert opinions in the field and not on empirical evidence. Expert opinions can offer qualitative insights but do not represent direct evidence.

Data synthesis

Upon concluding the data extraction process, the focus naturally shifts to synthesizing the findings. Employing a narrative approach proves effective at this stage, allowing for a concise summary of the results categorized into specific domains - such as clinical, organizational, or managerial - aligned with the primary objectives of the review. To enhance the clarity and comprehensibility of the synthesized data, it is recommended to integrate information and references from included SRs and CPGs into tables or synthesis. In parallel, for a visual synthesis of intervention effects, consider utilizing the Harvest plot [32]. This approach, endorsed by the Cochrane Handbook for Systematic Reviews of Interventions [33], is particularly useful when facing challenges posed by heterogeneous studies [34]. The Harvest plot stands as a graphical representation akin to the traditional Forest plot, yet its adaptability based on the review aim or the characteristics of the included studies underscores its notable strength.

Notably, Harvest plot facilitate the visual representation of the effects together with various study information, including reference, sample size, nature and direction of associations, and p-values. This ensures a structured and comprehensive presentation of pertinent data, improving clarity and synthesis. The flexibility of this methodology is crucial, especially when the results of systematic reviews are not easily aggregable through meta-analysis or pooled analysis. Thus, the Harvest plot provides a nuanced and informative representation, serving as a practical tool that enhances clarity in data synthesis and proves adaptable for various review. A detailed guide with examples for conducting the Harvest plot is provided in Supplementary File 2.

Data analysis

After synthesizing the extracted data, the subsequent step pertains to statistical analysis. First and foremost, descriptive statistics like means and standard deviations are computed for overall scores or any quantitative metric employed in the reviews. Assessing the level of agreement among reviewers is crucial to ensure the reliability of the extracted data. This can be achieved using the intraclass correlation coefficient (ICC), applying a bidirectional random effects model. The ICC can be computed for each specific domain or category and for the overall scores. Moreover, to further understand the relationship between scores assigned by different reviewers, tools like the Pearson correlation coefficient can be utilized to pinpoint the linear correlation between reviewers' scores. Lastly, to gauge the consistency and reliability of evaluations, metrics like Cohen's Kappa coefficient can be used. This coefficient provides a measure of corrected agreement, offering a more rigorous assessment of reliability between reviewers.

Meta-analysis

In the context of a systematic review of CPGs and SRs, there might arise a need to conduct a meta-analysis, especially when multiple studies provide similar or comparable data on a particular topic. Meta-analysis allows for combining the results of several studies to get an overall effect estimate, thereby increasing statistical power and the precision of the estimates. The decision to proceed with a meta-analysis should be grounded on the adequacy, homogeneity and quality of available data and the clinical relevance of combining the results (Fig. 6).

Evidence synthesis

The integration of results from SRs and CPGs should ultimately lead to a narrative synthesis of evidence, supported by a comprehensive table summarizing key findings clearly and in detail. This table should include a classification of the certainty of evidence following the system proposed by the OCEBM. Additionally, where applicable, the synthesis may be complemented by a graphical representation for a more comprehensive overview.

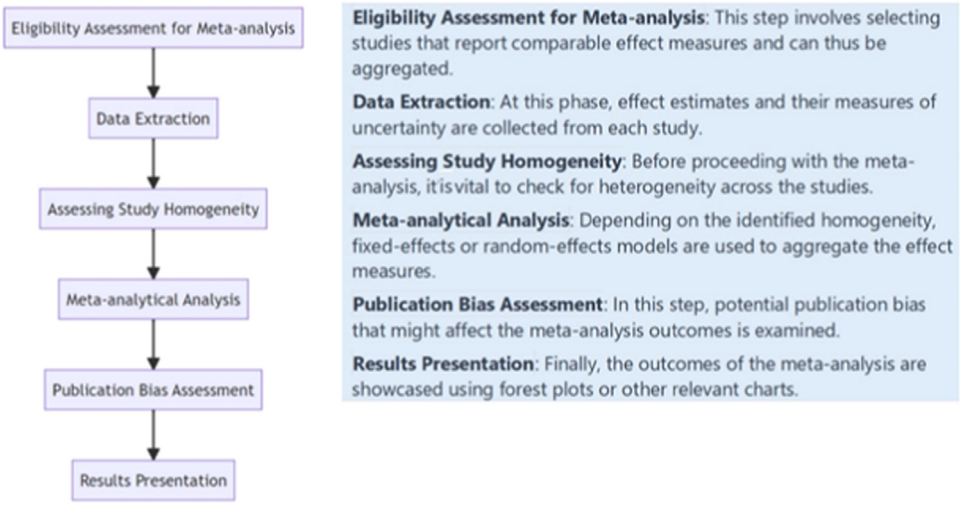

Fig. 6. Methodology for integrating a meta-analysis.
Legend: Methodology for integrating a meta-analysis into a systematic review.

Implications for research

1. **Bridging the Evidence Gap:** This novel approach highlights the importance of considering both CPGs and SRs in the review process, potentially filling evidence gaps and ensuring that the most recent and relevant data inform clinical recommendations.
2. **Promotion of Rigor and Transparency**: The rigorous evaluation tools like AGREE II and AMSTAR 2, combined with the dual-reviewer methodology, ensure that the selected CPGs and SRs meet high standards of quality and relevance. Such meticulousness promotes trust among end-users.
3. **Highlighting the Necessity for Continued Integration**: The dynamic nature of scientific research necessitates methodologies that can evolve and integrate new evidence sources. This method could serve as a foundation for future research methodologies that seek to combine diverse evidence forms.
4. **Encouraging Further Development**: While this method offers a robust approach to systematic reviews, there remains a scope for refining and expanding this methodology. Research communities should be encouraged to adopt, adapt, and enhance this method in line with emerging evidence and technological advancements.
5. **Potential for a Meta-analysis Approach**: Given the robustness of this integrated approach, it is poised to support a meta-analytic framework when multiple studies offer congruent data on a subject. This meta-analysis potential augments the method's versatility and its capability to provide a statistically robust overview of a topic.

Discussion

The landscape of scientific literature encompasses a diverse array of primary and secondary bibliographic records. Among these, Clinical Practice Guidelines (CPGs) and Systematic Reviews (SRs) play a pivotal role in providing synthesized evidence for clinical practice. Despite their invaluable contributions, relying exclusively on either source presents inherent limitations. This research methodology represents a natural progression in the evolution of systematic review methodologies that need to be integrated into summary documents.

By integrating these resources, our aim is to provide a broader, updated, and consequently, more clinically useful overview of available evidence. The proposed methodology strives to be rigorous, transparent, and reproducible, emphasizing robust evaluation methods such as AGREE II for CPGs and AMSTAR 2 for SRs. However, it is important to note that, despite their relevance, not all SRs and CPGs found in the literature comply with the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) methodology. In this context, we have chosen to standardize the assessment of evidence certainty using a methodology specifically outlined for healthcare professionals, namely the OECBM methodology. This choice has been motivated by the need to adopt a standardized and coherent approach in evidence evaluation, thereby contributing to the overall robustness and reliability of the obtained results. We believe that this strategy can enhance transparency and comparability in evidence assessments within the context of clinical research.

Ethics statements

No ethical approval was required for the purposes of this study, as it is unnecessary. The original protocol was registered in the International prospective register of systematic reviews (PROSPERO) of the National Institute of Health Research available at https://www.crd.york.ac.uk/prospero/ with protocol registration number: CRD42022372303.

Related research article

E. Aromataris, R. Fernandez, C.M. Godfrey et al. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. International journal of evidence-based healthcare, 13(3), (2015). 132–140. 10.1097/XEB.0000000000000055

V Smith, D. Devane, C.M Begley & M. Clarke, M. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. BMC medical research methodology, 11(1), (2011) 15. 10.1186/1471–2288–11–15

A. Grant, A. Booth. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Info Libr J. 2009 Jun;26(2):91–108. doi:10.1111/j.1471–1842.2009.00848.x

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Stefano Mancin: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Investigation, Visualization. **Marco Sguanci:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Investigation, Visualization, Formal analysis. **Desirèe Andreoli:** Writing – original draft, Writing – review & editing. **Fanny Soekeland:** Writing – original draft, Writing – review & editing. **Giuliano Anastasi:** Writing – original draft, Writing – review & editing. **Michela Piredda:** Writing – original draft, Writing – review & editing, Visualization, Supervision. **Maria Grazia De Marinis:** Writing – original draft, Writing – review & editing, Visualization, Supervision.

Data availability

Data will be made available on request.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.mex.2023.102532.

References

[1] W.M. Bramer, M.L. Rethlefsen, J. Kleijnen, et al., Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study, Syst. Rev. 6 (2017) 245, doi:10.1186/s13643-017-0644-y.
[2] D. Demner-Fushman, S. Hauser, G. Thoma, The role of title, metadata and abstract in identifying clinically relevant journal articles, AMIA Annu. Symp. Proc. (2005) 191–195.
[3] K. Godin, J. Stapleton, S.I. Kirkpatrick, et al., Applying systematic review search methods to the grey literature: a case study examining guidelines for school-based breakfast programs in Canada, Syst. Rev. 4 (2015) 138, doi:10.1186/s13643-015-0125-0.
[4] A. Paez, Grey literature: an important resource in systematic reviews, J. Evid. Based Med. 10 (3) (2017) 233–240, doi:10.1111/jebm.12266.
[5] Y.M. El-Khayat, C.S. Forbes, J.G. Coghill, Guideline.gov: a database of clinical specialty guidelines, Med. Ref. Serv. Q. 36 (1) (2017) 62–72, doi:10.1080/02763869.2017.1259914.
[6] E. Guerra-Farfan, Y. Garcia-Sanchez, M. Jornet-Gibert, et al., Clinical practice guidelines: the good, the bad, and the ugly, Injury 54 (2023), doi:10.1016/j.injury.2022.01.047.
[7] N. Jahan, S. Naveed, M. Zeshan, et al., How to Conduct a systematic review: a narrative literature review, Cureus 8 (11) (2016) e864, doi:10.7759/cureus.864.
[8] S. Kraus, M. Breier, W.M. Lim, et al., Literature reviews as independent studies: guidelines for academic practice, Rev. Manag. Sci. 16 (2022) 2577–2595, doi:10.1007/s11846-022-00588-8.
[9] K. Sriganesh, H. Shanthanna, J.W. Busse, A brief overview of systematic reviews and meta-analyses, Indian J. Anaesth 60 (9) (2016) 689–694, doi:10.4103/0019-5049.190628.
[10] A.C. Hardi, S.A. Fowler, Evidence-based medicine and systematic review services at becker medical library, Mo. Med. 111 (5) (2014) 416–418 PMID: 25438363; PMCID: PMC6172095.
[11] P. Tugwell, V. Welch, O. Magwood, et al., Protocol for the development of guidance for collaborator and partner engagement in health care evidence syntheses, Syst. Rev. 12 (1) (2023) 134, doi:10.1186/s13643-023-02279-1.
[12] T. Platz, Methods for the development of healthcare practice recommendations using systematic reviews and meta-analyses, Front. Neurol. (2021), doi:10.3389/fneur.2021.699968.
[13] Z. Munn, M.D.J. Peters, C. Stern, et al., Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach, BMC Med. Res. Methodol. 18 (2018) 143, doi:10.1186/s12874-018-0611-x.
[14] A. Grant, A. Booth, A typology of reviews: an analysis of 14 review types and associated methodologies, Health Info. Libr. J. 26 (2) (2009) 91–108 Jun, doi:10.1111/j.1471-1842.2009.00848.xV.
[15] Smith, D. Devane, C.M. Begley, et all., Methodology in conducting a systematic review of systematic reviews of healthcare interventions, BMC Med. Res. Methodol. (2011), doi:10.1186/1471-2288-11-15.
[16] C. Lunny, D. Pieper, P. Thabet, et al., Managing overlap of primary study results across systematic reviews: practical considerations for authors of overviews of reviews, BMC Med. Res. Methodol. 21 (2021) 140, doi:10.1186/s12874-021-01269-y.
[17] E. Hasanpoor, J. Hallajzadeh, Y. Siraneh, et al., Using the methodology of systematic review of reviews for evidence-based medicine, Ethiop. J. Health Sci. 29 (6) (2019) 775–778, doi:10.4314/ejhs.v29i6.15.
[18] A. Johnston, S.E. Kelly, S.C. Hsieh, et al., Systematic reviews of clinical practice guidelines: a methodological guide, J. Clin. Epidemiol. 108 (2019) 64–76, doi:10.1016/j.jclinepi.2018.11.030.
[19] C. Lunny, C. Ramasubbu, L. Puil, et al., Over half of clinical practice guidelines use non-systematic methods to inform recommendations: a methods study, PLoS One 16 (4) (2021) e0250356, doi:10.1371/journal.pone.0250356.

[20] C. O'Connor, E. O'Connor, S. Leitao, et al., Clinical practice guidelines for the antenatal management of dichorionic diamniotic twin pregnancies: a systematic review, BMC Pregnancy Childbirth 23 (1) (2023) 347, doi:10.1186/s12884-023-05652-z.
[21] A.P. Siddaway, A.M. Wood, L.V. Hedges, How to do a systematic review: a best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses, Annu. Rev. Psychol. 70 (2019) 747–770, doi:10.1146/annurev-psych-010418-102803.
[22] C. Lunny, C. Ramasubbu, S. Gerrish, et al., Impact and use of reviews and 'overviews of reviews' to inform clinical practice guideline recommendations: protocol for a methods study, BMJ Open 10 (1) (2020) e031442, doi:10.1136/bmjopen-2019-031442.
[23] E. Aromataris, R. Fernandez, C.M. Godfrey, et al., Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach, Int J Evid Based Healthc. 13 (3) (2015) 132–140, doi:10.1097/XEB.0000000000000055.
[24] M. Hennessy, R. Dennehy, S. Meaney, A protocol for a systematic review of clinical practice guidelines for recurrent miscarriage, HRB Open Res. 3 (2020) 12, doi:10.12688/hrbopenres.13024.3.
[25] M.B. Eriksen, T.F. Frandsen, The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: a systematic review, J. Med. Libr. Assoc. 106 (4) (2018) 420–431, doi:10.5195/jmla.2018.345.
[26] A.M. Methley, S. Campbell, C. Chew-Graham, et al., PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews, BMC Health Serv. Res. 14 (2014) 579, doi:10.1186/s12913-014-0579-0.
[27] A. Pollock, E. Berge, How to do a systematic review. International journal of stroke, Off. J. Int. Stroke Soc. 13 (2) (2018) 138–156, doi:10.1177/1747493017743796.
[28] M.C. Brouwers, M.E. Kho, G.P. Browman, J.S. Burgers, F. Cluzeau, G. Feder, et al., AGREE next steps consortium. AGREE II: advancing guideline development, reporting and evaluation in health care, CMAJ 182 (18) (2010) E839–E842 Dec 14, doi:10.1503/cmaj.090449.
[29] R. Andrade, R. Pereira, R. Van Cingel, et al., How should clinicians rehabilitate patients after ACL reconstruction? A systematic review of clinical practice guidelines (CPGs) with a focus on quality appraisal (AGREE II), British J. Sports Med. 54 (9) (2020) 512–519, doi:10.1136/bjsports-2018-100310.
[30] B.J. Shea, B.C. Reeves, G. Wells, et al., AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both, BMJ 21 (2017) 358 j4008, doi:10.1136/bmj.j4008.
[31] K. Matthias, O. Rissling, D. Pieper, et al., The methodological quality of systematic reviews on the treatment of adult major depression needs improvement according to AMSTAR 2: a cross-sectional study, Heliyon 6 (9) (2020), doi:10.1016/j.heliyon.2020.e04776.
[32] J. Howick, I. Chalmers, P. Glasziou, OCEBM Levels of Evidence Working Group 'The Oxford 2011 Levels of Evidence', Oxford Centre for Evidence-Based Medicine, Oxford, 2011.
[33] D. Ogilvie, D. Fayter, M. Petticrew, A. Sowden, S. Thomas, M. Whitehead, et al., The harvest plot: a method for synthesising evidence about the differential effects of interventions, BMC Med. Res. Methodol. 8 (1) (2008) 8, doi:10.1186/1471-2288-8-8.
[34] J.E. McKenzie, S.E. Brennan, R.E. Ryan, H.J. Thomson, & R.V. Johnston. Summarizing study characteristics and preparing for synthesis. In cochrane handbook for systematic reviews of interventions. 2019. 229–240. doi:10.1002/9781119536604.ch9