

M. For Table of Contents:

Chapter 12. USING EXISTING SYSTEMATIC REVIEWS TO REPLACE *DE NOVO* PROCESSES IN COMPARATIVE EFFECTIVENESS REVIEWS

Introduction and Rationale

Locating Existing Systematic Reviews

Assessing the Relevance of Existing Systematic Reviews

Assessing the Quality of Existing Systematic Reviews

Determining How to Use Existing Systematic Reviews

Reporting Methods and Results

Discussion: Reiterate Justification for Using Existing Systematic Reviews

Future Directions

12. USING EXISTING SYSTEMATIC REVIEWS TO REPLACE *DE NOVO* PROCESSES IN COMPARATIVE EFFECTIVENESS REVIEWS

Box 12-1. Key Points

Using existing systematic reviews has potential benefits and risks. EPCs and the relevant Task Order Officer should discuss these points either at the kick-off call or during the project itself.

This chapter does not focus on the use of existing systematic reviews for obtaining background information, providing background or discussion context, or cross-checking of references. Rather, it concerns the use of existing systematic reviews to replace a *de novo* process. It also does not consider the processes used to create a separate product called “umbrella” reviews or reviews of reviews.

We propose a five-step process to standardize the approach that EPCs can use to decide whether existing systematic reviews might provide value (Figure 12-1).

Transparency is a priority; users of a comparative effectiveness review should be able to determine what was done (Figure 12-2).

Two independent reviewers using a modified AMSTAR instrument, as described herein, should assess the quality of relevant reviews (Table 12-1).

EPCs should incorporate existing systematic reviews (i.e., use them to replace all or part of a *de novo* process) only if they are fully relevant and of high quality. Partly relevant or suboptimal quality reviews should not be incorporated, although they may be useful for cross-checking references and for providing background.

Once EPCs identify relevant, high-quality systematic reviews, they may opt to use them in the following ways: adapting or adopting the search strategy, using the summarized evidence, or a combination of these.

EPCs can choose to replace a *de novo* process to answer a key question by selecting the best review or may choose to summarize all of the relevant and high-quality reviews.

EPCs should routinely review reference lists of such systematic reviews to identify relevant studies.

If EPCs do a *de novo* synthesis, they should routinely compare results with those of relevant, high-quality systematic reviews and formally address consistency or potential reasons for discrepancies in the discussion of the report.

Introduction and Rationale

Approximately 2,500 new systematic reviews (SRs) and meta-analyses were published in 2006, and more than 3,300 are expected to be published in 2008.¹ Given this extensive body of existing SR and meta-analysis literature, questions have been raised about whether EPCs should use existing SRs in a comparative effectiveness review (CER) and if so, in what capacity? While a more comprehensive definition is given in chapter 1 of this guide, an SR uses an explicit methodology for systematically searching and synthesizing the literature, and for grading evidence.

An informal survey of eight non-EPC centers that conduct systematic reviews in the United Kingdom, Australia, and New Zealand confirmed that they are facing these same questions about the use of existing SRs without any commonly accepted approach.² In the summer of 2008, the “Existing SR Working Group” queried EPC directors about their experiences (including experience with both EPC and non-EPC projects) in this area. Overall, EPCs considered the use of an existing SR 50 percent of the time and used existing SRs slightly more than 30 percent of the time. The most commonly stated reason for using an existing SR was for completeness, but existing SRs were also often used when EPCs faced a topic of extensive breadth, size of literature, and limitations in timeframe or budget. Some EPCs identified using the existing SR while updating the SR.

When queried about how they were using existing SRs, EPCs indicated that they used existing SRs predominantly (74 percent of the time) for background information or to ensure completeness of the literature search. EPCs sometimes used results of existing SRs to answer key questions in the new SR, but in more than two-thirds of these cases at least a sample of the original trials or studies included in the existing SR were verified to ensure the quality of original data extraction.

When EPCs considered using existing SRs in a new SR, the most common reason given *not* to use one was that the identified reviews were not relevant to the specific questions being asked in the new SR. Other frequent reasons not to use existing SRs included: no time savings associated with using the existing SR versus using *de novo* methods to answer the key question, poor quality of existing SRs after detailed assessment, outdated existing SRs, and uncertainty about how to include them in a new SR.

As a result of our queries and subsequent discussion within the Working Group, we identified six possible benefits associated with using existing SRs in CERs:

- Allows a cross check to assure that relevant trials and studies are captured in a new CER.
- Allows EPCs to compare and contrast findings directly between this CER and previous SRs that may be relevant to health care decisionmakers.
- May save EPCs time, effort, and resources to answer key questions.
- May allow EPCs to anticipate and plan for context-specific methodological issues.
- May help avoid unnecessary redundancy among SRs.

- May provide analyses that are not readily available from other sources (e.g., subgroup analyses from a meta-analysis of individual patient data not available in constituent studies or published reports).

Conversely, five main risks are associated with using existing SRs in CERs that do not arise in a purely *de novo* process.

- If EPCs find numerous existing SRs, the time and resources required to evaluate them may be wasted because earlier reports may not be recent enough, not relevant enough to answer the key questions posed, or not of acceptable quality.
- Incorporating the results of existing SRs into a CER could propagate errors arising from errors in data abstraction, selection of studies, and qualitative or quantitative synthesis. Propagating errors can reduce credibility for the CER and the EPC program among stakeholders and users.
- Using an existing SR to answer key questions might create a perception that EPCs are not performing due diligence in conducting its CER. This perception might reduce credibility for the CER and the EPC program among stakeholders and users.
- If the existing SR does not provide evidence from primary studies and analyses in sufficient detail, the methodological process of the CER may be perceived to lack transparency.
- Ambiguity about how to compare multiple existing SRs on the same subject remains an important challenge. Lack of clear methodological guidance on selecting the most appropriate SRs could introduce reviewer bias, which is especially true if existing SRs have discordant results.

The use of existing SRs to substitute for purely *de novo* CER methods may provide benefits and risks. Of course, examining existing SRs to provide background information or other useful references for a CER is a common practice in EPC work, and we do not discuss this procedure further in this chapter. Ultimately, EPCs need to work with those who commission the work (i.e., their Task Order Officers at AHRQ and decisionmakers who nominated the topic) to determine whether the potential benefits associated with the incorporation of existing SRs are worth the risks to a CER's comprehensiveness and transparency or of introducing bias. Given that a decision has been made to incorporate the use of existing SRs in answering one or more key questions in lieu of using a purely *de novo* process, we recommend that EPCs apply the following approaches.

Figure 12-1 is a flow diagram adapted from a methods article by Whitlock and colleagues.² It will help guide EPCs as they move through the process of identification, assessment, and use of existing SRs. To ensure transparency, EPCs can include the graphic in Figure 12-2 in a CER report that will allow users to identify the number of original citations identified after a SR search, the number of articles that are excluded, and how the existing SRs are being used.

Locating Existing Systematic Reviews

Although EPCs can apply many possible approaches to identify existing SRs for a CER, we recommend two procedures. One strategy is to use a targeted search of higher-yield databases.² Because SRs are a secondary literature source, identifying relevant, high-

quality SRs is probably more important than identifying all SRs because redundancy of primary studies across SRs is likely. Higher yield databases include the Evidence-based Practice Center program, MEDLINE's Top 120 Index Medicus Journals, Health Technology Assessment, Cochrane Database of Systematic Reviews, and Database of Abstracts of Reviews of Effects; EPCs can add other databases depending on the topic. Alternatively, EPCs can identify SRs during their title and abstract searches while conducting a broad *de novo* literature search for trials and studies, as long as the searches are structured not to exclude reviews. The EPC medical librarian is a valuable resource when making these decisions and developing the search strategy.

Assessing the Relevance of Existing Systematic Reviews

EPCs considering the inclusion of prior SRs in a CER should begin with a fundamental presumption – that the intent is to answer one or more key question(s) or a specific portion of a key question with an existing SR in lieu of a completely *de novo* process. Relevance requires consideration of the populations addressed, interventions and comparators, outcomes, settings, and study designs. Those SRs not completely relevant to the current review (partially relevant) may still be useful for background material or for cross-checking of references. Some existing SRs will not be relevant at all and should be eliminated from any further consideration at this stage.

Initial Screening for Relevance:

As depicted in Figure 12-1, after EPCs conduct a literature search for existing SRs (Step 1), they need to screen identified citations for relevance (Step 2). Citations that are not SRs (primary research, narrative reviews) or duplicate citations can be readily excluded.

Many factors that determine whether an existing SR is relevant or not are addressed in the SRs methods section. Timeliness of the existing SR is critical. Timeliness refers not to the publication date of the review, but to how recently the literature search had been conducted. When considering issues of timeliness, reviewers should be aware that SRs can become outdated quickly.³ Whether a SR is outdated depends primarily on the topic because some areas may not be as intensely researched and newer studies added only rarely. We generally recommend bridging any search date for a SR that ends a year or earlier than the present date. Given their clinical expertise, expert team members may be helpful in deciding acceptable date parameters; ideally they should make this decision *a priori*.

If EPCs regard an earlier SR to be outdated, they can still consider using the search results (obtaining data from the evidence tables) and then updating from 1 year before the date of the original literature search to the present time with a *de novo* process. By going back 1 year before the existing SR's search date, the lag time between the publication of an article and its inclusion into standardized literature retrieval databases ought not to be a major factor. Using the search results from these existing SR would require only that the earliest date for which studies could be included (e.g. 1960) is in line with the date the EPCs have set for their CER.

Focusing on Population, Intervention, Comparator, Outcomes, and the Timing of their Measurement, Setting and Study Design (PICOTS-SD) to Assess Relevance:

For existing SRs that make it to this stage, EPCs should compare the PICOTS-SD in the earlier SRs with these elements in the new CER protocol.⁴ Determining similarity will depend on how well the existing SR describes these elements. Poor reporting will make it impossible for an EPC to consider inclusion of an existing SR. Poor reporting, however, is an element of quality appraisal as well, so a poorly reported SR would not be eligible for incorporation for both relevance and quality reasons. Appreciating the subtle differences that may exist between an existing SR and the current CER is vital; this generally requires EPCs to give careful consideration of these elements.

Population: The need for the population in an existing SR to “match” completely the intended population in a new CER will depend to some degree on the clinical condition of interest and the questions being addressed. On the one hand, for example, a CER that is attempting to review interventions for hemorrhagic stroke may not be well served by including an existing SR with studies of patients with any kind of stroke, unless results clearly separate the subgroup of studies relevant to hemorrhagic stroke patients. On the other hand, a CER that is examining any kind of stroke might be able to incorporate a relevant, high-quality prior SR addressing hemorrhagic stroke only. Similarly, an existing SR restricted to adults will be of limited utility if the new key questions include young children. Other CERs, however, may require less rigidity, and modest differences in age range or geographic range (e.g., United States vs. North America) may be less important.

Intervention: To ensure that existing SRs evaluated the same intervention as intended for the new CER, the team should look carefully at criteria for inclusion used in the older review. Particularly important is making sure that issues such as dosing and mode of delivery match as closely as possible. When the existing SR was either more or less inclusive than the CER is intended to be, the experts on the team need to determine that this factor will not fundamentally change the conclusions. This may become an issue when dosing regimens change over time, as has been the case with use of higher-dose statins in recent years, or for example, in the evolution of cardiac devices such as pacemakers to newer, dual chamber versions.

Comparator: EPCs should consider whether they are interested in the effect of the intervention of interest as it compares with usual practice or another intervention and ensure that the existing SR matches this criterion. EPCs should note, when comparing treatments with usual care, whether usual practice has changed significantly since the timeframe of the earlier SR; this would make older studies – and perhaps a review of those studies – not applicable to the current concern. Such evolution of usual practice has been a significant issue, for instance, in “medical treatment” after acute coronary syndrome; older versions of medical treatment are no longer comparable with current practice. In surgical reviews, it may be important to know what supportive treatments were used in the past compared to those associated with interventions being reviewed – for example, if patients previously spent longer in post-operative care in bed rather than in active rehabilitation, those older studies may not reflect current practice. For issues of

this type, the input of clinical experts can be particularly useful to determine changes in usual care over time.

Outcomes: The outcomes assessed in existing SRs should be the same or similar to the outcomes envisioned for the CER. The usual caveats regarding use of intermediate or nonpatient-oriented outcomes apply for existing SRs just as they apply to inclusion criteria for constituent studies.

Timing of Outcome Measurement: Some SRs are restricted to studies with relatively short periods of followup. The period of appropriate followup of course depends on the condition, intervention under consideration, and the outcome being assessed. The rationale for such restriction may be the lack of availability of longer-term followup; when such studies become available, the relevance of the older SR is reduced. Often, short periods of followup involve surrogate outcome measures; both factors (length of followup; surrogate or proxy outcomes) decrease a SR's relevance. Timing is not the same as timeliness (how recent the existing SRs), which EPCs should examine early in the relevancy assessment.

Setting: Older SRs can address interventions in a broad or narrow range of settings, such as interventions to reduce falls in inpatient settings, in nursing homes, and in the home and other community settings. Although some of these distinctions will be clear by examining the populations addressed, a previous SR that covers a range of settings may not be relevant to a more narrowly scoped CER unless results of the former are stratified by setting.

Study Design: SRs can differ appreciably in the types of study designs that they consider acceptable. EPCs may find that surveying inclusion criteria related to study design is a useful early step in a relevance evaluation. If EPCs plan to include randomized and controlled clinical trials and high-quality comparative cohort studies as evidence in their CERs, but an existing SR covers only randomized controlled trials, then the latter is only partially relevant to the current effort.

Once EPCs have established relevance for an existing SR, they should assess and rate quality using the approach described below. Quality assessments (Figure 12-1, Step 3) are time intensive and should be conducted only on existing SRs found to be relevant.

Assessing the Quality of Relevant Systematic Reviews

EPCs should strive to include in their CERs only existing SRs adhering to high methodological standards; they should avoid routinely including all existing SRs in an attempt to be comprehensive. Note that this admonition is in contrast to another effort, a review of reviews, in which reviewers are asked to summarize the available evidence at the level of the systematic review.

Several instruments designed to rate quality of SRs are available.⁵ Regardless of the specific instrument that is chosen for this purpose, the instrument should address all the

following factors: methods used to identify, select, appraise, and synthesize studies; the possibility of publication bias; and potential conflicts of interest.⁶

Commonly Used SR Quality Instruments:

In assessing the quality (i.e., assessing the risk of bias) of existing SRs, EPCs should address both the methods used by the earlier systematic reviewers to minimize bias and the transparency and completeness with which they reported their methods, individual study details, and results. Checklists for improving reporting of SRs (e.g., QUOROM [recently renamed PRISMA], MOOSE) have been used as surrogate tools for quality assessment, although they were designed to improve transparency and consistency of reporting SR methods, not directly to assess methodological quality.⁷⁻⁹ For example, the QUOROM checklist requires detailed descriptions of the literature search strategy terms and sources searched, but it does not provide criteria for distinguishing adequate from inadequate searches.⁷ In addition, inadequate reporting of SR methods does not necessarily mean that the SR was conducted poorly. Nonetheless, rating the quality of an SR without understanding how it was conducted is difficult. Several items related to quality of reporting have been incorporated into instruments such as the ones from Oxman and Guyatt and AMSTAR.^{6,10}

The Oxman and Guyatt instrument was one of the early widely used standardized quality rating indices for evaluating the scientific quality of a review article; unlike other quality rating instruments specifically developed for SRs, some empiric evidence supports its use.¹⁰ Reviews with lower quality ratings on the Oxman and Guyatt instrument are more likely to show treatment benefit.^{11,12} However, methods for evaluating SRs have evolved since the Oxman and Guyatt instrument was developed, and it does not address several methodological domains now thought to be important.¹³

The newer Assessment of Multiple Systematic Reviews (AMSTAR) tool includes additional criteria, such as whether study selection and data extraction were conducted in duplicate, whether publication bias was assessed, and whether conflicts of interest were reported.⁶ Although more data are needed to determine its reliability and validity, AMSTAR has been proposed as the preferred instrument for assessing the quality of SRs by the World Health Organization and by the Canadian Optimal Medication Prescribing and Utilization Service (COMPUS), among others.^{14,15} One domain that is not included in AMSTAR pertains to nonbiased application of inclusion and exclusion criteria, although EPCs can adapt the AMSTAR instrument to include such an item (see recommendation below).

Limitations in Quality Rating Scales:

As much as possible, CER investigators should apply objective and reproducible criteria when using quality assessment instruments such as Oxman and Guyatt or AMSTAR.^{6,10} For example, a “comprehensive” literature search could be defined as requiring searches on at least two electronic databases, reference list searching, and expert queries. Although EPCs could use this definition in most instances, they may need to tailor criteria for specific topics. For example, for assessing the quality of SRs that evaluate

acupuncture, fully meeting the literature search criteria could require searching Asian language databases.

For some criteria included in quality rating instruments, delineating objective definitions is difficult; EPCs then must apply subjective judgments. For example, AMSTAR includes the items “Was the scientific quality of the included studies used appropriately in formulating conclusions?” and “Were the methods used to combine the findings of studies appropriate?”⁶ Assessing and rating quality using discrete categorical choices can make quality judgments appear more clear-cut and objective than they really are. Operationalizing subjective qualifiers such as “appropriate” at the outset of each assessment, taking into consideration factors relevant to the specific topic at hand, could help. Having at least two independent reviewers from an EPC assess quality and reporting methods for resolving discrepancies is desirable.

Another limitation in applying quality rating instruments is that they are not designed to detect inconsistencies in application of inclusion criteria or errors in data abstraction. For example, an SR¹⁶ of antidepressants for low back pain specified randomization as an inclusion criterion but included a nonrandomized clinical trial.¹⁷ Among the included studies, this trial reported the highest estimate of benefit and may have affected the SR’s conclusions.¹⁶ Checking data from SRs against primary studies can reveal important discrepancies.^{18,19}

Numerical summary scores (e.g., adding up the number of criteria that are adequately met) have been used to summarize the overall quality of SRs. Such scores can be misleading because reviews with different flaws may receive the same summary score. A summary score could not dissect the nature of the bias in the individual review. For example, an SR could meet nearly all methodological criteria and receive a near-perfect summary score, but one serious methodological shortcoming could invalidate its results; a summary score may well not reflect that important shortcoming.

We suggest that CER authors describe the implications of individual methodological flaws rather than rely on numerical summary scores. For example, exclusion of “grey literature” or non-English language citations may or may not have important effects on estimates of benefits or harms.^{20,21} If EPCs find no clear indication of publication bias in an SR and if stable and precise estimates are available for the outcome(s) of interest, excluding these types of literature is not likely to be a serious shortcoming. However, excluding “grey literature” or non-English language trials would be a serious shortcoming in an SR if large numbers of trials or important trials are known or suspected to exist in these literature types. As cases in point: medical device evaluations may rely on “grey literature,”²² and alternative and complementary medicine evaluations may rely on foreign language literature.²³

Assigning categorical quality scores (such as “good,” “fair,” or “poor”) may be appropriate after taking into account the number and seriousness of methodological shortcomings.²⁴ In general, good quality SRs should be defined as those that have few or no methodological shortcomings and a low risk of bias. Fair quality SRs have some

methodological flaws but the EPC conducting the CER determined that the flaws will not seriously bias or invalidate the results. Poor quality SRs contain a serious flaw (or flaws) that, in the judgment of the EPC conducting the CER, is (are) highly likely to bias or invalidate the results.

CER Quality Assessment Recommendations:

When EPCs assess the quality of an existing SR for a CER project, we recommend:

- At least two independent reviewers should assess SRs for quality.
- EPCs should report methods for resolving discrepancies between reviewers.
- EPCs should confirm the reproducibility of application for inclusion criteria and the accuracy of data abstraction in at least a sample of the studies.
- To have a common starting point, EPCs should use AMSTAR for quality evaluation for two reasons: (1) it was developed based on an SR of quality rating instruments and has undergone some construct and validity testing; and (2) it is becoming more widely used internationally.

AMSTAR assesses 11 criteria for quality and the choices are (Yes, No, Can't Answer, and Not Applicable).⁶ We suggest supplementing the AMSTAR questions as deemed appropriate for the particular project or topic at hand. Table 12-1 summarizes the criteria with some additional considerations that EPCs may have for their CERs.

Determining How to Use Existing Systematic Reviews

At this point in the process, we assume that EPCs have identified that one or more existing SRs are relevant for the CER and that they are of adequate quality. Now EPCs must determine the appropriate way to incorporate them into the CER (Figure 12-1, Step 4). Several possibilities are available (Figures 12-1 and 12-2), and they are not mutually exclusive:

- Incorporate already summarized evidence from existing SRs into the CER.
- Incorporate summarized evidence from existing SRs into the CER but conduct *de novo* sensitivity analyses. In essence, using an existing SR to answer a key question but then conducting additional analyses using data from the original studies. For example, using a SR to answer a key question in a CER about whether or not to use coenzyme Q10 in heart failure but then conducting *de novo* sensitivity analyses to determine the impact of publication date on the results.
- Utilize and SRs search strategy in lieu of a *de novo* process but then use *de novo* methods for analysis and synthesis. This would be possible if the search strategy was consistent with Chapter five of the EPC Methods Guide entitled "Finding Evidence", but the quality of other processes were inadequate or could not be determined.
- Build on existing SRs by updating meta-analyses or qualitative syntheses.
- Address conflicting results of existing SRs with a *de novo* analysis.
- Use at least part of the comprehensive literature search strategy to identify trials or other studies for the CER.

EPCs can incorporate existing SRs into a CER in their entirety if those SRs have research questions that are very similar to the CER's key question(s). They can also include SRs

in part if only a portion is either of interest or relevant to a key question or questions within the CER. This may include incorporating summarized evidence within a specific population or for a specific intervention. In these cases, the methods used in the SR would have to be consistent with the EPC Methods Guide for Comparative Effectiveness Reviews (available online at:

<http://www.effectivehealthcare.ahrq.gov/healthInfo.cfm?infotype=rr&ProcessID=60>) including "Finding Evidence", "Assessing Quality", "Grading the Strength of a Body of Evidence", and "Principles," including issues of scientific independence, and avoiding conflicts of interest.

Previous SRs are unlikely to be wholly sufficient to substitute for a CER because CER questions are identified by a process that assesses the redundancy of a topic with previously published SRs.²⁵ Moreover, other factors reduce the possibility that existing SRs will be able to answer all the key questions in a CER: the comprehensive and broad nature of many CERs; the need to evaluate efficacy, effectiveness, and harms; the inclusion of high-quality observational studies (often excluded in other SRs) in many CERs; and evaluations based on factors such as sex or gender, race, and/or ethnicity.

Reporting deficiencies are an important factor limiting the use of existing SRs. For instance, SRs may not present results of individual trials at all, or they may not give information that EPCs need in order to build the existing SR into the current CER. Because EPCs cannot determine the accuracy or validity of the results of earlier SRs, using summary findings without complete reporting may compromise transparency in the CER. Little is gained from incorporating such SRs into a CER because EPCs could not update the meta-analyses in the existing SR with more recent trials or studies without obtaining the primary articles and repeating the data abstraction.

EPCs may find that several recent, relevant, and high-quality SRs are appropriate for a given CER; they then need to determine how best to proceed. One approach is to incorporate the single "best" existing SR (most relevant and least biased) into their own reports.² However, selecting a single review may pose the risk of introducing selection bias; EPCs must ensure transparency in their criteria for eligibility. Another approach is to conduct a meta-review (also known as an "umbrella review"), whereby they select all relevant, high-quality SRs that meet an *a priori* publication date threshold and then assess the consistency among them.^{26,27} When using this approach, EPCs should provide summary tables with information about all the included SRs, so as to maximize transparency. If the selected relevant, high-quality SRs have discordant findings, EPCs should explore the reasons for these disagreements. If EPCs cannot readily give reasons for the discordant findings, then they can regard this as an indication that they need to adopt a *de novo* approach to answer that key question.

Reporting Methods and Results

This chapter of the CER methods guide for the AHRQ Effective Health Care program provides the recommended approach to use when locating existing SRs and assessing their relevance and quality, and it offers a strategy for dealing with multiple existing SRs that EPCs can use to replace a *de novo* process. We emphasize the need for both

reproducibility and transparency when using an existing SR (Figure 12-1, Step 5). By specifying the targeted search databases and terms used to locate existing SRs and employing a flow diagram to demonstrate the disposition of the citations identified (Figure 12-2), EPCs can ensure that readers of the CER will be able to assess the process and, if desired, reproduce it. If EPCs decide to search for only previous SRs within a specific date range, or to exclude citations based solely on the dates of the existing SR's literature search, then they should specify the rationale for using this cut-off date.

Providing a summary table that specifies the details of included existing SRs used to replace a *de novo* process is important.^{28,29} Summary tables of existing SRs should document the volume, type, and quality of the primary research included. In comparing these previous SRs, ideally the table should address the overlap (or lack of overlap) in primary research in these SRs: e.g., what studies or types of studies were included in one review versus another. Documenting these points will help readers in assessing such factors and the magnitude of net benefits; it will also clarify how EPCs have graded the strength of a body of evidence.²

Discussion: Reiterate Justification for Using Existing Systematic Reviews

In the discussion section of a CER report, EPCs should restate the initial justification for using one or more earlier SRs instead of following a *de novo* process. They should discuss clearly any limitations arising from the use of existing SR(s). Authors should comment on advantages and disadvantages identified through the process or creating the specific CER to help the conduct of future CERs.

Although not the focus of this paper, comparing findings from the CER with the findings from existing SRs is important because it helps health care decisionmakers understand how the CER in question relates to the existing SR literature. Authors can present similarities and differences and discuss potential reasons for any congruities or discrepancies that they have identified.

Future Directions

Many areas require further research to help determine how best to incorporate existing SRs into CERs. These include:

- Determining whether the targeted SR search strategy that has been proposed in this chapter consistently helps to identify the highest quality reviews with less resource allocation than a more broadly conducted search.
- Examining whether applying different relevance or quality criteria markedly changes the SRs that EPCs ultimately include in their CERs or the results derived from these SRs.
- In a situation involving several existing SRs with sufficient relevance and quality, investigating whether the conduct of a meta-review or selecting the best SR approach is the better strategy.
- Documenting savings or increases in time or resources (if any) that comes from using an existing SR approach in place of a *de novo* process.

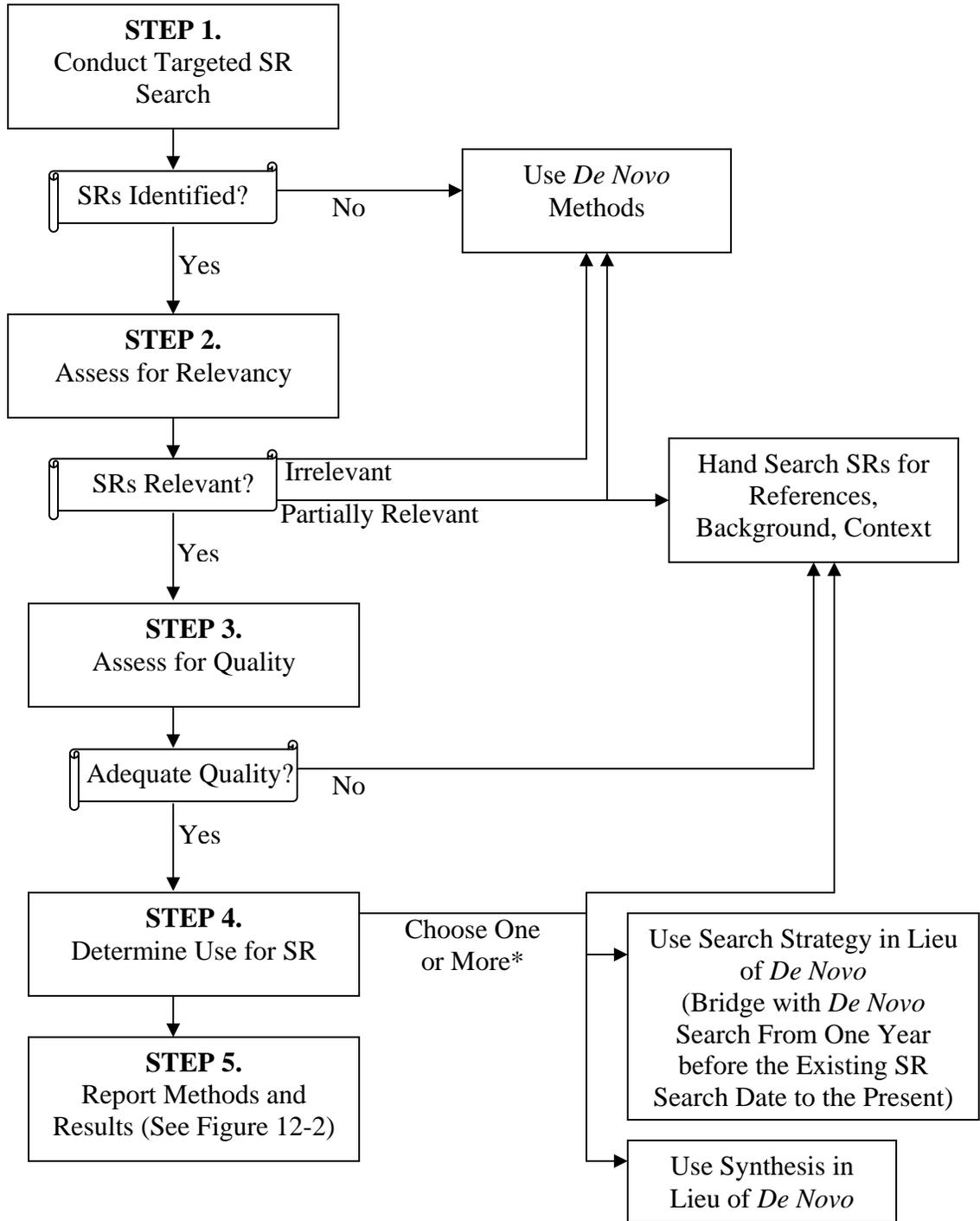
- Documenting the additional time or resources used in searching for and evaluating existing SRs when they are ultimately not used to replace a *de novo* process.
- Determining whether it is more efficient to search for SR as part of the overall search strategy for a topic, or as a first step before searching for primary literature.

REFERENCES

1. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. *PLoS Med* 2007;4:e78.
2. Whitlock EP, Lin JS, Shekelle P, Robinson KA. Using existing systematic reviews in complex systematic reviews. *Ann Intern Med* 2008;148:776-82.
3. Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. *Ann Intern Med* 2007;147:273-4.
4. Rothwell PM. External validity of randomized controlled trials: "to whom do the results of this trial apply?". *LANCET* 2005;365:13-4.
5. West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, Lux L. Systems to rate the strength of scientific evidence. Evidence Report/Technology Assessment No. 47. Research Triangle Institute-University of North Carolina Evidence-based Practice Center. AHRQ Publication No. 02-E016. 2002. Rockville, MD. Agency for Healthcare Research and Quality.
6. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007;7:10:10.
7. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999;354:1896-900.
8. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;283:2008-12.
9. Shea BJ, Dube C, Moher D. Assessing the quality of reports of systematic reviews: the QUOROM statement compared to other tools. In: Egger M, Smith GD, Altman DG, eds. *Systematic Reviews in Health Care: Meta-Analysis in Context*. 2nd Edition. London: BMJ Publishing Group; 2001.
10. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991;44:1271-8.
11. Jadad AR, McQuay HJ. Meta-analyses to evaluate analgesic interventions: a systematic qualitative review of their methodology. *J Clin Epidemiol* 1996;49:235-43.
12. Assendelft WJ, Koes BW, Knipschild PG, Bouter LM. The relationship between methodological quality and conclusions in reviews of spinal manipulation. *JAMA* 1995;274:1942-8.
13. Shea B, Dube C, Moher D. Assessing the quality of reports of systematic reviews: the QUORUM statement compared to other tools. In: Egger M, Smith GD, Altman DG, eds. *Systematic reviews in health care: meta-analysis in context*. London, UK: BMJ Publishing Group, 2001:122-39.
14. Oxman AD, Schunemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 8. Synthesis and presentation of evidence. *Health Research Policy and Systems* 2006;4:20.

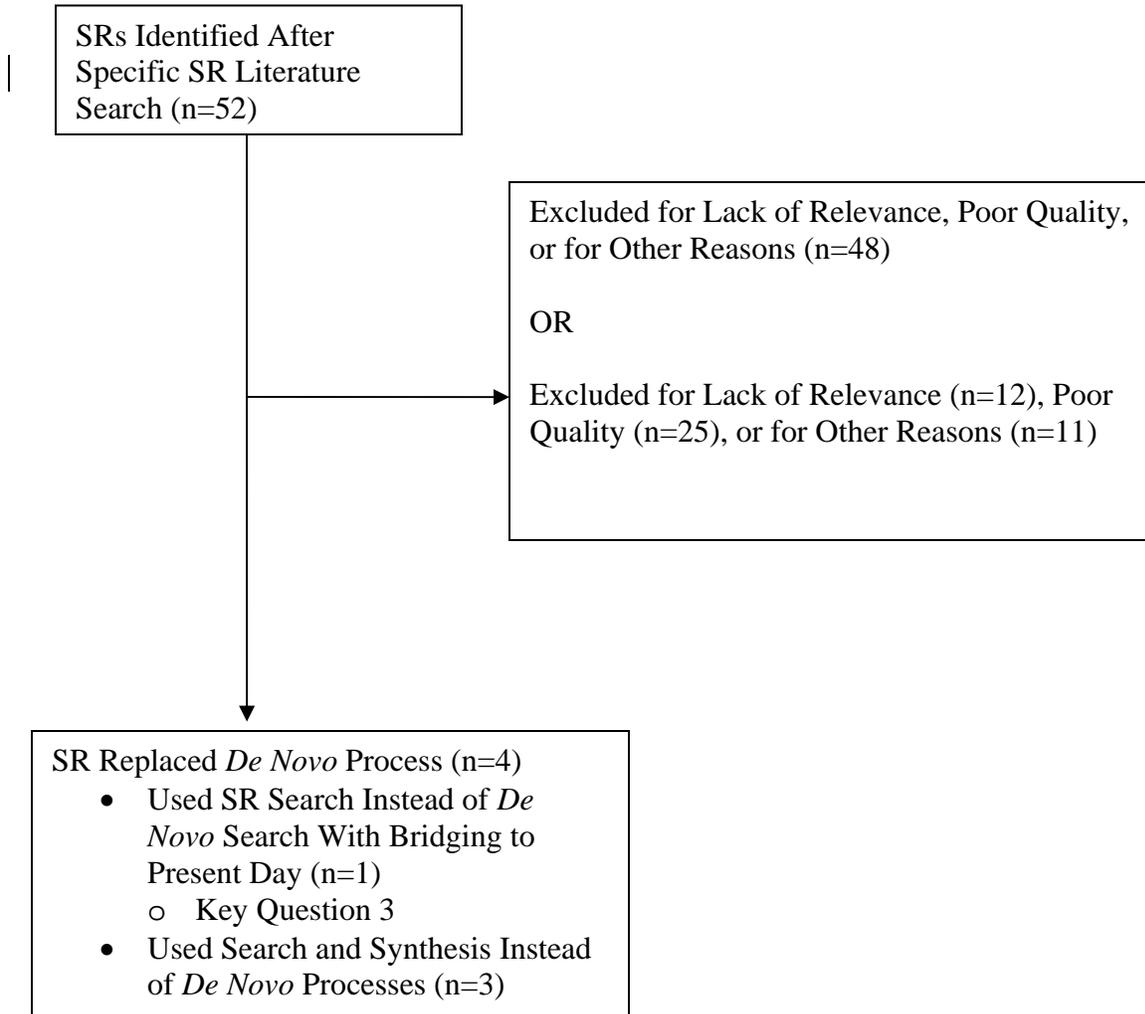
15. COMPUS Procedure. Evidence-based best practice recommendations. Available at: http://www.cadth.ca/media/compus/pdf/COMPUS_%20procedure_e.pdf (accessed 10/29/08).
16. Salerno SM, Browning R, Jackson JL. The effect of antidepressant treatment on chronic back pain. *Arch Intern Med* 2002;162:19-24.
17. Ward NG. Tricyclic antidepressants for chronic low-back pain. *Spine* 1986;11:661-5.
18. Gotzsche PC, Hrobjartsson A, Marie K, Tendal B. Data extraction errors in meta-analyses that use standardized mean differences. *JAMA* 2007;298:430-7.
19. Jones AP, Remington T, Williamson PR, Ashby D, Smyth RL. High prevalence but low impact of data extraction and reporting errors were found in Cochrane systematic reviews. *J Clin Epidemiol* 2005;58:741-2.
20. Egger M, Juni P, Bartlett C, Holenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technol Assess*. 2003;7:1-76.
21. Moher D et al. What contributions do languages other than English make on the results of meta-analyses. *J Clin Epidemiol* 2000;53:964-72.
22. Hartling L, McAlister FA, Rowe BH, Ezekowitz J, Friesen C, Klassen TP. Challenges in Systematic Reviews of Therapeutic Devices and Procedures. *Ann Intern Med*, Jun 2005; 142: 1100-11.
23. Shekelle PG, Morton SC, Suttorp MJ, Buscemi N, Friesen C. Challenges in Systematic Reviews of Complementary and Alternative Medicine Topics. *Ann Intern Med*, Jun 2005; 142: 1042-7.
24. Drug Effectiveness Review Project. Quality assessment methods for drug class reviews for the drug effectiveness review project. <http://www.ohsu.edu/ohsuedu/research/policycenter/DERP/about/upload/QualityAssessmentDERP-2.pdf>. Accessed October 23rd, 2008.
25. Whitlock EP, Lopez SA, Chang S, Helfand M, Floyd N. Identifying, selecting, and refining topics for research reviews: AHRQ and the Effective Health Care Program). *JCE*, Submitted.
26. Ruddy R, House A. Meta-review of high-quality systematic reviews of interventions in key areas of liaison psychiatry. *Br J Psych* 2005;187:109-20.
27. Moe RH, Haavardsholm EA, Christie A, et al. Effectiveness of nonpharmacological and nonsurgical interventions for hip osteoarthritis: an umbrella review of high-quality systematic reviews. *Phys Ther* 2007;87:1716-27.
28. Chou R, Huffman LH. American Pain Society. American College of Physicians. Nonpharmacologic therapies for acute and chronic low back pain: a review of the evidence for an American Pain Society/American College of Physicians clinical practice guideline. *Ann Intern Med* 2007;147:492-504.
29. Lorenz KA, Lynn J, Dy SM, et al. Evidence for improving palliative care at the end of life: a systematic review. *Ann Intern Med* 2008;148:147-59.

Figure 12-1. Systematic Process for Identifying, Assessing, and Using Existing Systematic Reviews (SRs).



Legend: Adapted from Whitlock et al. 2008.² * denotes that a *de novo* process is preferred if several relevant, high-quality SRs come to discordant findings; in that case, the existing SRs should be used solely for hand-searching and background context.

Figure 12-2. Illustrative Existing Systematic Review (SR) Diagram



Appendix. An Example of Quality Assessment of SRs Conducted by the Stanford-UCSF EPC.

Before AMSTAR was available, the Stanford-UCSF EPC conducted a comprehensive umbrella review (meta-review) for care coordination. The EPC assessed the quality of the SRs by using criteria previously used by the drug effectiveness review project of the Oregon EPC.²³ These criteria map fairly well to AMSTAR.

The SRs assessed did well on following:

1. Purpose stated
2. Selection criteria stated
3. Search strategy reported (similar to AMSTAR 3)
4. Time frame reported (similar to AMSTAR 3)
5. Article retrieval effort made (trying to retrieve all articles)
6. Appropriate synthesis (although most of these were narratives) (similar to AMSTAR 9)
7. Sufficient details of included articles presented (AMSTAR 6)

The SRs were inadequate in the following areas:

1. Dual abstract review (AMSTAR 2)
2. Dual data abstraction (AMSTAR 2)
3. Disagreements resolution reasonable/described
4. Team included personnel one would expect in a well-conducted SR (librarian, methods expert)
5. Validity assessment of included articles (AMSTAR 8)

In summary, this umbrella review found that most SRs uniformly met some of the criteria but that performance of SRs on other criteria was much less uniform. SR quality varied substantially and needed to be considered carefully.

Table 12-1. AMSTAR Quality Criteria With Considerations for Comparative Effectiveness Reviews

| Number | Criterion | Considerations for Comparative Effectiveness Reviews |
|--------|---|--|
| 1 | Was an ‘ <i>a priori</i> ’ design provided? | |
| 2 | Was there duplicate study selection and data extraction? | <ul style="list-style-type: none"> ▪ Was there dual review for study selection and data extraction? ▪ After checking a sample of original studies: <ul style="list-style-type: none"> ▪ Was the application of inclusion/exclusion criteria unbiased? ▪ Were any discrepancies between data from primary papers and the published systematic review identified? |
| 3 | Was a comprehensive literature search performed? | <ul style="list-style-type: none"> ▪ Was the search strategy appropriate for the posed key questions? This should be consistent with the EPC Methods Guide for Comparative Effectiveness Reviews entitled “Finding Evidence”. |
| 4 | Was the status of publication (i.e., grey literature) used as an inclusion criterion? | <ul style="list-style-type: none"> ▪ Some reviews do not restrict inclusion based on whether studies were peer reviewed or not. EPCs should state their criteria for inclusion/exclusion and justifications for the criteria (e.g., reasons for restriction to English language, excluding letters and abstracts, etc.) |
| 5 | Was a list of studies (included and excluded) provided? | |
| 6 | Were the characteristics of the included studies provided? | |
| 7 | Was the scientific quality of the included studies rated and documented? | <ul style="list-style-type: none"> ▪ Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate, etc.) assessed? This should be consistent with the EPC Methods Guide for Comparative Effectiveness Reviews entitled “Assessing the Quality”. ▪ Did the systematic review include high- quality primary studies (<i>N.B.</i>, no matter how well-conducted a systematic review, its findings are limited by the quality of included primary studies)? |
| 8 | Was the scientific quality of the included studies used appropriately in formulating conclusions? | <ul style="list-style-type: none"> ▪ This item applies only if EPCs use the conclusions from the prior systematic review(s) in their CERs. Often EPCs will use only the results and formulate conclusions based on the data and analysis presented. This should be consistent with the EPC |

| | | |
|----|---|--|
| | | Methods Guide for Comparative Effectiveness Reviews entitled “Grading the Strength of a Body of Evidence”. |
| 9 | Were the methods used to combine the findings of studies appropriate? | |
| 10 | Was the likelihood of publication bias assessed? | |
| 11 | Was the conflict of interest stated? | Do the authors have declared or known conflicts of interest disclosed. Examples include funding source for the project, consulting fees, and stock ownership. family members |