Methods Guide
for Comparative Effectiveness Reviews

Integrating Bodies of Evidence: Existing Systematic Reviews and Primary Studies

AHRQ
Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

This report is based on research conducted by the Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Centers' 2014 Methods Workgroup 3. The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information (i.e., in the context of available resources and circumstances presented by individual patients).

AHRQ or U.S. Department of Health and Human Services endorsement of any derivative products that may be developed from this report, such as clinical practice guidelines, other quality enhancement tools, or reimbursement or coverage policies may not be stated or implied .

This research was funded through contracts from the Agency for Healthcare Research and Quality to the following Evidence-based Practice Centers: The Johns Hopkins University (290-2012-00007-I), Pacific Northwest (290-2012-00014-I), University of Alberta (290-2012-00013-I), University of Minnesota (290-2012-00016-I), RAND Corporation (290-2012-0006-I), RTI International and the University of North Carolina (290-2012-00008-I), and the Scientific Resource Center for the EPC Program (290-2012-00004-C).

Persons using assistive technology may not be able to fully access information in this report. For assistance, contact EffectiveHealthCare@ahrq.hhs.gov

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

Suggested citation: Robinson KA, Chou R, Berkman ND, Newberry SJ, Fu R, Hartling L, Dryden D, Butler M, Foisy M, Anderson J, Motu'apuaka ML, Relevo R, Guise JM, Chang S. Integrating Bodies of Evidence: Existing Systematic Reviews and Primary Studies. Methods Guide for Comparative Effectiveness Reviews (Prepared by the Scientific Resource Center under Contract No. 290-2012-00004-C). AHRQ Publication No. 15-EHC007-EF. Rockville, MD: Agency for Healthcare Research and Quality. February 2015. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Prepared by:
Scientific Resource Center
Portland, OR

Investigators:
Karen A. Robinson, Ph.D.[1]
Roger Chou, M.D.[2]
Nancy D. Berkman, Ph.D.[3]
Sydne J. Newberry, Ph.D.[4]
Rongwei Fu, Ph.D.[5]
Lisa Hartling, B.Sc.P.T., M.Sc., Ph.D.[6]
Donna Dryden, Ph.D.[6]
Mary Butler, Ph.D.[7]
Michelle Foisy, B.A.-Hon, M.A.[6]
Johanna Anderson, M.P.H.[5]
Makalapua Motu'apuaka, B.S.[5]
Rose Relevo, M.L.I.S., M.S.[5]
Jeanne-Marie Guise, M.D., M.P.H.[5]
Stephanie Chang, M.D., M.P.H.[8]

[1]Johns Hopkins University Evidence-based Practice Center, Baltimore, MD
[2]Pacific Northwest Evidence-based Practice Center, Portland, OR
[3]RTI-University of North Carolina Evidence-based Practice Center, Research Triangle Park, NC
[4]Southern California Evidence-based Practice Center, RAND, Santa Monica, CA
[5]Scientific Resource Center for the AHRQ Effective Health Care Program, Portland VA Research Foundation, VA Portland Health Care Systems, Portland, OR
[6]Department of Pediatrics, University of Alberta, Edmonton, Canada
[7]Minnesota Evidence-based Practice Center, Minneapolis, MN
[8]Center for Evidence and Practice Improvement, Agency for Healthcare Research and Quality, Rockville, MD

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies and strategies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

Strong methodological approaches to systematic review improve the transparency, consistency, and scientific rigor of these reports. Through a collaborative effort of the Effective Health Care (EHC) Program, the Agency for Healthcare Research and Quality (AHRQ), the EHC Program Scientific Resource Center, and the AHRQ Evidence-based Practice Centers have developed a Methods Guide for Comparative Effectiveness Reviews. This Guide presents issues key to the development of Systematic Reviews and describes recommended approaches for addressing difficult, frequently encountered methodological issues.

The Methods Guide for Comparative Effectiveness Reviews is a living document, and will be updated as further empiric evidence develops and our understanding of better methods improves. We welcome comments on this Methods Guide paper. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by email to epc@ahrq.hhs.gov.

Richard Kronick, Ph.D.
Director
Agency for Healthcare Research and Quality

David Meyers, M.D.
Acting Director
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Director and Task Order Officer
Evidence-based Practice Program
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Key Informants

In designing the study questions, the EPC consulted several Key Informants who represent the end-users of research. The EPC sought the Key Informant input on the priority areas for research and synthesis. Key Informants are not involved in the analysis of the evidence or the writing of the report. Therefore, in the end, study questions, design, methodological approaches, and/or conclusions do not necessarily represent the views of individual Key Informants.

Key Informants must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any conflicts of interest.

The list of Key Informants who participated in developing this report follows:

Nadera Ahmadzai, M.D., M.P.H., M.Sc.
Senior Research Associate
OHRI Knowledge Synthesis Group

Samantha Barker, B.S.
Senior Project Manager, Evidence Synthesis and Nutritionist
Research Funding Program
Institute for Safety, Compensation and Recovery Research

Lorne Becker, M.D.
Emeritus Professor
Department of Family Medicine
SUNY Upstate Medical University

Kathryn M. Curtis, Ph.D.
Health Scientist
Division of Reproductive Health
Centers for Disease Control and Prevention

Chantelle Garritty, B.A., D.C.S., M.Sc.
Senior Operations Manager
Ottawa Methods Centre, Ottawa Hospital

David Hopkins, M.D., M.P.H.
Medical Officer
Community Guide Branch
Centers for Disease Control and Prevention

Susan Norris, M.D., M.Sc., M.P.H.
Guidelines Review Committee Secretariat
World Health Organization

Julie Obbagy, Ph.D., R.D.
Center for Nutrition Policy and Promotion
Nutrition Guidance and Analysis Division, USDA

Julie Polisena, M.Sc.
Scientific Advisor
Canadian Agency for Drugs and Technologies in Health

Joanne Spahn, M.S., R.D., F.A.D.A.
Director
USDA Nutrition Evidence Analysis Library Division

Lesley Stewart, Ph.D.
Director
University of York Centre for Reviews and Dissemination

Peer Reviewers

Prior to publication of the final evidence report, EPCs sought input from independent Peer Reviewers without financial conflicts of interest. However, the conclusions and synthesis of the scientific literature presented in this report does not necessarily represent the views of individual reviewers.

Peer Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential non-financial conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential non-financial conflicts of interest identified.

The list of Peer Reviewers follows:

Nadera Ahmadzai, M.D., M.P.H., M.Sc.
Senior Research Associate, OHRI
Knowledge Synthesis Group

Eric Bass, MD
Director, Johns Hopkins Evidence-based Practice Center
Professor of Medicine

Kathryn M. Curtis, Ph.D.
Health Scientist, Division of Reproductive Health
Centers for Disease Control and Prevention

Paul Glasziou, MRCGP, FRACGP, Ph.D., M.B.B.S.
Director (CREBP), Faculty of Health Sciences and Medicine
Bond University

David Hopkins, M.D., M.P.H.
Medical Officer, Community Guide Branch
Centers for Disease Control and Prevention

Bob Kane, M.D.
Co-Director, Minnesota Evidence-based Practice Center
School of Public Health
University of Minnesota

Tianjing Li, M.D., Ph.D., M.H.S.
Assistant Professor
Center for Clinical Trials
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health

Marian McDonagh, Pharm. D.
Associate Director, Pacific Northwest Evidence-based Practice Center
Professor of Medical Informatics & Clinical Epidemiology, Oregon Health & Science University

Melissa McPheeters, Ph.D., M.P.H.
Director, Vanderbilt Evidence-based Practice Center
Director, Epidemiology Track, M.P.H.
Research Associate Professor
Department of Health Policy
Vanderbilt University Medical Center

Dawid Pieper, M.P.H.
Universitaet Witten/Herdecke
Institute for Research in Operative Medicine

Julie Polisena, M.Sc.
Scientific Advisor, Canadian Agency for Drugs and Technologies in Health

Public Comment
American Physical Therapy Association

Paul Shekelle, Ph.D., M.D., M.P.H.
Director, Southern California Evidence-based Practice Center, RAND Corporation

Joanne Spahn, M.S., R.D., FADA
Director, USDA Nutrition Evidence Analysis Library Division

Jon Treadwell, Ph.D.
Associate Director, ECRI Institute Evidence-based Practice Center

Tom Trikalinos, M.D.
Director, Brown Evidence-based Practice Center
Director, Center for Evidence-based Medicine
Associate Professor, Health Services, Policy & Practice
Brown University

Meera Viswanathan, Ph.D.
Director, RTI-UNC Evidence-based Practice Center
RTI International

Integrating Bodies of Evidence: Existing Systematic Reviews and Primary Studies

Introduction

In 2008, recognizing the exponential growth in the number of systematic reviews being published, the need to update existing reviews, and the increasing time and money constraints, a group of researchers across the Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Centers (EPC) program developed preliminary guidance on the appropriate role of existing reviews in developing new reviews on related topics.[1] This work identified a series of issues to consider and developed guidance to address some of these issues, which subsequently became codified as a chapter in the AHRQ EPC Program Methods Guide.[2]

In 2012, an EPC Methods Workgroup sought to identify remaining challenges in integrating existing systematic reviews into new reviews. The Workgroup conducted targeted scans of published and grey literature to identify guidance used or offered by organizations dedicated to conducting systematic reviews or improving the systematic review process.[3,4] Discussions with EPC directors and staff were conducted to identify ways EPCs have found and used existing systematic reviews. The Workgroup identified eight areas where additional guidance is needed for reviewers considering integrating existing systematic reviews in new reviews:

- Criteria to identify when a new EPC review will add value to a field with many existing reviews;
- Organizing principles for integrating primary (study-level) and secondary (systematic review-level) evidence, (including templates for evidence tables);
- Guidelines for transparently reporting the methods used to identify, select, and decide how best to utilize existing systematic reviews;
- Methods for minimizing bias in selecting prior reviews to use or integrate when there are multiple existing reviews;
- Methods for minimizing bias in incorporating selected portions of an existing review;
- Qualitative and quantitative methods for summarizing bodies of evidence that include a systematic review as the only or as one source of evidence;
- Robust quality assessments for existing systematic reviews (beyond AMSTAR[5]), and;
- Methods to grade the strength of evidence for bodies of evidence that include a systematic review as the only or as one source of evidence.

We convened a new Workgroup in 2013 to develop recommendations on several of these areas that the EPCs had identified as most pressing. Based on these identified needs and preliminary discussions, our objective was to develop guidance on integrating bodies of evidence from systematic reviews and primary studies in new reviews. Different uses of existing systematic review, such as checking references and summarizing existing evidence in introduction or discussion, have been adequately described in prior work.[3,4] We specifically focused on methods for using reviews when there are multiple reviews, assessing risk of bias of primary studies in existing reviews, and summarizing and assessing the strength of bodies of evidence that include or are limited to existing systematic reviews.[3,4] The immediate intended audience is the EPC program, but we hope that this guidance may be useful to all systematic reviewers facing these issues.

Methods

Approach

We assembled a workgroup of EPC methodologists to develop recommendations on the integration of existing systematic reviews in new reviews, building on the work of the previous EPC Methods Workgroup. We sought information from an updated scan of the literature and interviews with leaders in the field to inform consensus recommendations developed through twice monthly conference calls.

Literature Search

The Scientific Resource Center (SRC) provides support for the AHRQ EPC Program for the advancement of scientific methods, strategic planning, peer review, topic nomination and education. As part of this work, the SRC curates a bibliographic database of nearly 10,000 citations on the methodology of systematic reviews and comparative effectiveness research, dating back to the 1950s.[6] We searched this database (22 April 2014) for publications that included any of the following terms in the title, abstract, or descriptor: Overview; Umbrella; Review of review; Use of secondary studies; Discordant review; Incorporating review; Multiple systematic review; Review of systematic review; Relevant review; Synthesis of systematic review; Secondary evidence; Synopsis of systematic; Synopsis of review. Citations were screened first by the SRC informationist to remove material clearly not relevant and then by at least one member of the Workgroup. We sought documents that provided guidance on the integration of existing systematic reviews in systematic reviews. We were seeking literature that could inform discussions and thus did not apply strict eligibility criteria. However, we did not collect examples of how existing systematic reviews have been used; rather we used Key Informant interviews to understand how reviews have been used. We also did not consider related but separate topics such as methods for updating reviews or the conduct of reviews of reviews (overviews).

Key Informant Interviews

We invited systematic reviewers, representatives from organizations that produce systematic reviews and methodologists to participate in 60-minute telephone interviews. Workgroup members interviewed 11 of these "key informants" (KIs) or thought leaders from organizations that conduct or use systematic reviews. Each KI completed a conflict of interest disclosure form prior to participation. Prior to initiating the interviews, we developed and piloted an interview guide to focus the interviews (Appendix A), which includes a brief introduction of the background of the workgroup, the purpose of the interview, and interview questions. The interview guide was sent to KIs prior to the call. The interview questions covered three general topics:

- Using multiple existing reviews
- Assessing risk of bias
- Summarizing and assessing bodies of evidence.

We developed five scenarios or case studies outlining alternative actions depicting a range of approaches to integrating an existing review into a new review to help frame the discussions.

These scenarios, which assume that at least one relevant existing review has been identified that is considered of acceptable quality, are not mutually exclusive:

- Scenario 1: Use review without modifying or adding new studies
- Scenario 2: Use review and add new studies
- Scenario 3: Use review with new or modified analysis
- Scenario 4: Use selected elements of review
- Scenario 5: Do not use review

Each interview was recorded and transcribed. One investigator analyzed transcripts of the interviews for key themes using NVivo10 software (QSR International) and these themes were reviewed by at least one other investigator. In the Results section, we present the themes by the three general topics described above. More detail about the issues focused on within each of the themes is provided in Appendix B, organized by both general topics and scenarios.

Development of Recommendations

All workgroup members reviewed the themes and examples from the KI interviews. Conference calls were held twice a month. We developed recommendations in an iterative manner until consensus was reached. As with all guidance for the conduct of systematic reviews, we found very little evidence on which to base our recommendations; therefore, we focused on providing guidance for areas in which the workgroup came to consensus on minimum standards. For areas in which there was less certainty that following the guidance would be worth the effort required, the workgroup provided suggested actions that review authors may elect not to follow due to resource or other constraints.

Results

Literature Search

After screening the 470 citations found in the methods research database, we identified no literature relevant to informing our discussions, other than the previous EPC methods work.[1-4]

Synthesis of Key Informant Interviews

We interviewed 11 KIs from various organizations that conduct systematic review. While one organization noted that it chooses not to include any existing systematic reviews in its reviews, most organizations described a process to evaluate and include existing reviews, though none of them has published guidance on this issue. One organization mentioned using the prior EPC methods work in this area.[1] Key themes from the interviews are organized by the three general topics in this section and more detail is presented in Appendix B.

Using Multiple Existing Reviews

KIs reported that it is common to identify multiple relevant existing systematic reviews and that they would typically use the “best” review rather than include all existing reviews. KIs cited several considerations in deciding which was the “best” among a group of reviews. In general, they noted that priority is given to reviews that most closely match the current review; scope (populations, interventions, and outcomes of interest), inclusion/exclusion criteria, and methods.

If an existing review matches only some characteristics of the current review, elements of the existing review might be incorporated or the existing review might only need to be supplemented with additional studies. However, "empty" reviews, that is, those with very little evidence, regardless of their relevance and quality, might not be used at all by some organizations.

In addition, KIs noted the importance of considering the quality and recency or search date of the existing systematic review(s) in selecting reviews and in deciding how to use existing reviews. KIs most often reported that the AMSTAR[5] tool was used to rate the quality of systematic reviews, though they recognized that it has some limitations. Some organizations pick the most recent review(s) among those with the highest quality, while others set an absolute threshold for the AMSTAR score, such as a score greater than eight. Some organizations would also consider whether the review was produced by a reputable source, and do not use reviews with perceived bias or conflict of interest, for example, industry funded reviews.

Reviews that are selected may be used in a variety of ways. The search date(s) of the existing review and, for some KIs, the likelihood that new studies might change the conclusions, helped to determine what elements or how much of the prior review was used or if the existing review was used at all.

KIs indicated that transparency and level of detail reported in the existing reviews was critical for evaluating whether and/or how to use an existing review. Adequate details need to be reported to effectively assess the fit, and quality of the review, as well as whether it is up-to-date. Details about how the statistical analyses were conducted in the existing review are important to KIs so they can assess whether an analysis was adequate and appropriate for the research questions or current standards. If an existing review does not provide sufficient details, KIs said they may not use the review at all or may not use the earlier analysis, and instead would newly conduct their own.

KIs noted that the presence of substantial or unexplained discordance among prior reviews was worrisome and could be interpreted as a signal to conduct a new review. Still, KIs expressed the belief that it is important to acknowledge and discuss discordant reviews in the new review, even if they are not formally "included" as evidence in the new review.

Assessing Risk of Bias

KIs noted that assessments of the risk of bias (RoB) of individual studies are among the key findings of a systematic review. Most KIs noted that the tools for assessing the quality of an existing systematic review were not adequate to determine whether the ROB assessment of the individual studies could be used in the current review. The two most important considerations for KIs in determining whether to use the RoB assessment from an existing review were the type of RoB tool used and the transparency of the description of study RoB. KIs also reported that they had more confidence in the RoB assessments in a review conducted by a source that they consider trustworthy (most frequently cited examples were The Cochrane Collaboration and the EPC Program). While KIs said that an existing review need not have used the same RoB tool that will be used in the current review, the existing review needs to have used a tool that is widely accepted and that the review team considers appropriate for the given study design. The importance of transparency was emphasized by all KIs; the RoB tool should have been described in the methods section and study level details that are provided should allow for the reassessment of RoB for a sample of studies. The combination of an acceptable RoB tool, sufficient details about the process of assessment, and agreement on RoB ratings from the sample of studies typically is sufficient for KIs to accept the ratings of the whole review. However, not trusting and

therefore needing to redo the RoB assessment would result in questioning whether the existing review could be used at all.

Summarizing and Assessing Bodies of Evidence

KIs reported a wide range of practices for using assessment of bodies of evidence from an existing review. As with RoB assessment, the importance of transparency (i.e., being able to understand the factors underlying the strength of evidence grading) was emphasized. This is especially important as there are different systems and different versions or adaptations of systems used to consider the strength of a body of evidence. Some organizations reported using the existing grading of a prior review if the assessment is described in sufficient detail. Other organizations always complete the assessment of bodies of evidence again using their own criteria and judgment, while others do not grade the body of evidence. KIs indicated that they may not choose to use a review with an inadequate conclusion and no clear indication of the strength of evidence.

Recommendations

Selecting Reviews

The incorporation of existing systematic review(s) into a current systematic review assumes the identification of relevant reviews of sufficient quality. We refer readers to search filters for identifying systematic reviews, such as those found on the InterTASC Information Specialists' Sub-Group Search Filter Resource site.[7] Our scope did not include the assessment of such filters.

It cannot be assumed that a report called a systematic review or meta-analysis is, in fact, a systematic review. Reports need to be screened in full text to identify systematic reviews. Although we found no validated set of criteria to conduct such screening, we would consider the following as minimum criteria based on standard definitions of systematic reviews:[8,9]
(i) presence of explicit and adequate search, (ii) applied pre-defined eligibility criteria, (iii) consideration of quality of included studies or RoB assessment, and (iv) synthesis or attempt to synthesize the findings, either quantitatively and/or qualitatively.

Recommendation: Existing reviews should be confirmed as systematic reviews through the application of a minimum set of eligibility criteria. We propose that the minimum eligibility criteria for systematic reviews include an explicit and adequate search, application of pre-defined eligibility criteria to select studies, risk of bias assessment for included studies, and synthesis of results.

The identification of multiple prior systematic reviews presents challenges. While it is important for systematic reviewers to consider all potentially relevant primary studies, it may not be the case that all potentially relevant prior systematic reviews need to be considered. It is more important to assess and include prior reviews that are most relevant and of high quality than to attempt to include all reviews. Several factors can be considered in assessing relevance, including the date(s) of the search (currency), and the review methods. Relevancy should be assessed using the PICOT (population, intervention, comparison, outcome, time) framework for the review question. It is also important to consider the study design(s) included in the existing review. Older reviews may be less useful to use if they use versions of RoB tools or methods that do not consider key sources of bias or are not compatible with methods of the current review.

Other factors to consider in the existing review(s) include whether details about the characteristics of included studies, RoB and study-level data are provided. The number of and reasons for excluded studies, such as in a PRISMA diagram,[10] could be an added criterion; a list of excluded studies with reasons for exclusion is ideal but it may not be reasonable to exclude a prior review for not including such a list. It may be worthwhile to review the search strategy, such as by using the PRESS (peer-review of search strategies) model[11] or by checking that key studies were identified. There are tradeoffs in selecting the most recent review, highest quality review or the most relevant review. Systematic reviewers should try to be transparent in how these selections were made.

Recommendation: Criteria to assess the relevance, in terms of question elements and currency, and quality of existing systematic reviews under consideration for inclusion in reviews should be predefined.

Several tools exist for assessing the methodological strengths and limitations of systematic reviews. The tool most often cited during our interviews, AMSTAR, is currently under revision (B. Shea, personal communication, March 2014). The Cochrane Collaboration is developing a tool to assess RoB for systematic reviews called ROBIS, which is in pilot testing (P. Whiting, personal communication, June 2014). Given the work of these and other groups, we have not assessed tools nor made specific recommendations about which tool(s) to use in assessing the quality of existing systematic reviews.

It is difficult to set a threshold of quality for when we would "trust the results" of an existing systematic review or have sufficient confidence to use an existing review. We suggest establishing a minimum set of criteria for good or high quality systematic review that would be applied to reviews judged to be relevant:

- Search that includes multiple data sources
- RoB assessment using a generally accepted tool appropriate for the design(s) of the included studies and a process to avoid bias (such as independent reviewers)
- Explicit system or method for considering the body of evidence, or sufficient information to assess the major domains of strength of evidence (SOE) such as RoB, directness, consistency, precision and reporting bias.

Recommendation: The quality of relevant existing systematic reviews should be assessed in an explicit manner with a minimum set of quality criteria that include search of multiple sources, use of a generally accepted tool for risk of bias assessment, and sufficient information to assess the strength of the body of evidence that includes the major domains of risk of bias, directness, consistency, precision and reporting bias.

Assessing Risk of Bias

It is important to remember that determining the quality of a systematic review tells us nothing about the RoB of the primary studies or strength of evidence of the body of evidence included in that review. Even when incorporating existing systematic reviews into a review, we need to consider the underlying primary studies in evaluating the body of evidence. As such, the question is the extent to which we can rely upon the work completed in the prior review. Whether the RoB assessments of the studies from the prior review can be used first assumes that the process used in conducting RoB assessment was clearly reported.

The second consideration is the approach used to assess RoB in the prior review. Appraising the approach used to assess RoB includes determining if the prior review used a generally accepted tool and explicit methods, as well as whether that approach is similar enough to that being used in the new review. Because the RoB is ultimately collapsed to study limitations domain categories of high, medium or low for the synthesis and grading of the SOE, a tool need not be exactly the same as the one being used in the current review. The main consideration is that the tool used in the prior review covers the key sources of potential bias, such as those outlined in the EPC Program Methods Guide,[1] so that the assessment of key sources of bias of the previously and newly identified primary studies could be reasonably synthesized together. If the prior review used an approach that is the same or similar to the approach used in the current systematic review, then we recommend that RoB assessment only needs to be conducted again on a sample of the primary studies. This step is suggested to confirm the concordance of those prior assessments with those of the current systematic review authors. This is really about determining the comfort of relying on the assessments from the prior review and is suggested to confirm the consistency of those assessments with those of the current systematic review authors. Small discrepancies in RoB assessments may not be important if the overall grades are consistent (i.e., low RoB studies are consistently identified as low RoB and high RoB studies as high RoB). If there are marked discrepancies, a review team should consider conducting RoB again, or may decide to not attempt to integrate the systematic review into their review. This step might also provide further information as to the ability to translate and use the prior assessments if a different tool is used.

Recommendation: The RoB assessments from the existing systematic review may be used when the review described an explicit process, including the use of a tool or method that is compatible with the approach of the current review and that assessed the key sources of potential bias.

Recommendation: We suggest that RoB assessment be repeated in a sample of studies from an existing review under consideration for inclusion in a new review to confirm concordance with current review team approach.

Qualitative and Quantitative Synthesis

One rationale often used for including existing systematic reviews in new reviews is to leverage the work completed by the prior systematic review authors. The more limited funding allocated to updates within the EPC program is also predicated on this assumption of being able to use elements of the prior work, such as data abstraction, evidence tables and synthesis. However, in all tables and syntheses, whether presenting evidence from the existing review(s) or new review, it should be clear that the synthesis is based on the evidence in the underlying primary studies.

For evidence tables, we suggest using separate tables or subheadings to make a clear distinction between the data abstracted by the current review authors and information from the existing review(s). We recognize that the review authors may want to display different data than were collected in the prior review or that the detailed tables of individual primary studies may not be available from prior reviews. In other cases, data from the primary studies in the prior reviews do not necessarily need to be reabstracted.

Recommendation: We recommend that at a minimum reviews should narratively describe findings of the prior review(s), including the number and types of studies included, and the overall findings.

Recommendation: We recommend that newly identified studies be clearly distinguished from studies in the existing review(s) when presented in the narrative and any tables (e.g., separate tables).

Summary tables of existing reviews should incorporate review characteristics or assessments that are tied to the SOE domains. These tables should summarize this information with sufficient detail to make the body of evidence clear. Information that should be presented includes the number of studies, the number of study participants, point estimates of effect measures and their confidence intervals. If multiple prior reviews are included, it is helpful to provide a matrix comparing which studies were included in which reviews. Study characteristics of newly identified individual primary studies may be added to those from the prior reviews but we suggest that these are clearly distinguished. In some cases it may be better to describe the overall evidence and briefly mention how many studies were new.

Recommendation: Summary tables should include sufficient information to support ratings for overall strength of evidence, including ratings for individual strength of evidence domains (study limitations, consistency, precision, directness, reporting bias). The strength of evidence ratings should be based on the underlying primary evidence, not the number or quality of existing systematic reviews.

No clear rules exist for when a new quantitative synthesis needs to be conducted or when a synthesis, qualitative or quantitative, may be used from a prior review; however, a group of EPCs has attempted to address this gap as it relates to updating reviews.[12-14] We suggest that review authors consider the SOE domains approach in synthesis. Using SOE domains as a framework, authors would consider if any new primary studies identified would change the judgments about the SOE domains (i.e., study limitations, consistency, precision, directness, and reporting bias). If the new studies would change the conclusions or the SOE judgments or the new studies are in some way different, it will be necessary to conduct a new quantitative synthesis. If the new studies are consistent with prior syntheses and likely will not to change the conclusion of the review, the reviewer authors may choose not to conduct an updated synthesis. Rather, the synthesis from the prior review could be presented along with an updated qualitative synthesis including the newly identified studies and an explanation of how they are consistent with the prior findings. However, review authors may wish to conduct a new quantitative synthesis (of all studies, from prior review and newly identified) regardless of any changes in conclusions expected in order to present a more precise or more up-do-date estimate. In addition, the development of new standards in the conduct of systematic reviews, such as the selection of the model used for quantitative synthesis,[15] may necessitate updating reviews that might not have otherwise been considered out-of-date.

Recommendation: Using strength of evidence domains as a framework (study limitations, consistency, precision, directness and reporting bias), review authors should consider how new evidence would change estimates of effect or ratings for strength of evidence. A new quantitative synthesis (i.e., pooled estimate) is needed if new studies would change conclusions or strength of evidence judgements, or to obtain a more precise or more up-to-date estimate.

Summarizing and Assessing Body of Evidence

The considerations of whether to use the SOE grading from an existing review are similar to those in determining whether to use the RoB assessments: Did the prior review use an acceptable grading system in an appropriate manner? We would consider an acceptable grading system to include the domains outlined in the AHRQ EPC Program Methods Guide: summary of the strengths and limitations of primary studies (study limitations), directness, consistency, precision, and reporting bias. Assessments that are compatible include the EPC SOE, GRADE, and USPTF tools.[16-18] However, no matter the approach used in the previous review, because SOE is a judgment about the body of evidence for a particular question and outcome, it may not be possible to use the prior grades.

When no new studies have been identified, the review team needs to consider if the prior SOE grading was conducted using acceptable criteria. If the existing systematic review used the same or similar grading system, we suggest that grading be conducted again on a sample of the questions and outcomes to check for consistency with the approach of the current review team. As with RoB, concerns with the SOE assessment may prompt a review team to conduct all SOE assessments again. It is also important to assess consistency to ensure that the questions with and without new studies are graded in a similar manner.

If new studies have been identified that address a particular key question, it may be desirable to identify thresholds or triggers for when grading needs to be repeated. However, the process for how to determine if there is sufficient evidence to change a prior grade is an open question. As described above, recent EPC work has addressed when to update a review[12-14] and an EPC Workgroup is currently seeking to determine the predictive validity of SOE grading. In general, the judgment is whether enough new evidence exists to change the conclusions or confidence in the conclusions. For example, if the prior review included 10 studies with low RoB and reviewers identify 3 new smaller studies with high RoB, it is unlikely that the conclusions will change.[12]

Recommendation: In cases where the existing systematic review(s) did not complete strength of evidence grading for a comparison and outcome of interest, the strength of evidence should be assessed for the body of evidence, considering primary studies from prior review(s) and any new studies identified.

Recommendation: In cases where no new studies are added to the body of evidence, the strength of evidence assessment from the existing systematic review may be used if conducted using an acceptable grading approach consistent with current review context. In these cases, we suggest that the overall strength of evidence assessment be reviewed, considering the strength of evidence domains, to confirm consistency with current review team assessments.

Recommendation: In cases where new studies are added to the body of evidence, the strength of evidence may need to be reassessed based on all studies/evidence.

Discussion

Summary of Recommendations

Existing reviews should be confirmed as systematic reviews through the application of a minimum set of eligibility criteria. We propose that the minimum eligibility criteria for systematic reviews include an explicit and adequate search, application of predefined eligibility criteria to select studies, risk of bias assessment for included studies, and synthesis of results.

Criteria to assess the relevance, in terms of question elements and currency, and quality of existing systematic reviews under consideration for inclusion in reviews should be predefined.

The quality of relevant existing systematic reviews should be assessed in an explicit manner with a minimum set of quality criteria that include search of multiple sources, use of a generally accepted tool for risk of bias assessment, and sufficient information to assess the strength of the body of evidence that includes the major domains of risk of bias, directness, consistency, precision and reporting bias.

The risk of bias assessments from the existing systematic review may be used when the review described an explicit process, including the use of a tool or method that is compatible with the approach of the current review and that assessed the key sources of potential bias.

We suggest that risk of bias assessment be repeated in a sample of studies from an existing review under consideration for inclusion in a new review to confirm concordance with current review team approach.

We recommend that at a minimum reviews should narratively describe findings of the prior review(s), including the number and types of studies included, and the overall findings.

We recommend that newly identified studies be clearly distinguished from studies in the existing review(s) when presented in the narrative and any tables (e.g., separate tables).

Summary tables should include sufficient information to support ratings for overall strength of evidence, including ratings for individual strength of evidence domains (study limitations, consistency, precision, directness, reporting bias). The strength of evidence ratings should be based on the underlying primary evidence, not the number or quality of existing systematic reviews.

Using strength of evidence domains as a framework (study limitations, consistency, precision, directness and reporting bias), review authors should consider how new evidence would change estimates of effect or ratings for strength of evidence. A new quantitative synthesis (i.e., pooled estimate) is needed if new studies would change conclusions or strength of evidence judgements, or to obtain a more precise or more up-to-date estimate.

In cases where the existing systematic review(s) did not complete strength of evidence grading for a comparison and outcome of interest, the strength of evidence should be assessed for the body of evidence, considering primary studies from prior review(s) and any new studies identified.

In cases where no new studies are added to the body of evidence, the strength of evidence assessment from the existing systematic review may be used if conducted using an acceptable grading approach consistent with current review context. In these cases, we suggest that the overall strength of evidence assessment be reviewed, considering the strength of evidence domains, to confirm consistency with current review team assessments.

In cases where new studies are added to the body of evidence, the strength of evidence may need to be reassessed based on all studies/evidence.

The motivation for this work was the concern expressed by members of EPCs about the lack of guidance on how to integrate existing systematic reviews into new reviews. We sought but did not find evidence in the literature to inform our recommendations. Therefore, our recommendations are based on expert opinion and this work should be considered a working document. There are tradeoffs in determining whether it is more efficient, and methodologically sound, process to rely on a prior review or to start from scratch.

Future Research

We envision additions and changes to these recommendations as more work in this area is conducted. We identified several areas for such future research:

- Specific to this document, there is a need for feedback from reviewers as they implement these recommendations. This will help to assess if the recommendations are helpful, to identify any pragmatic considerations or challenges in implementing the guidance, and to identify areas of remaining challenges. We suggest that the methodology committee of the EPC Program consider how to evaluate this and other guidance as part of their mission.
- The recommendations were developed to be generally applicable. Going forward, we need to consider if different recommendations may be needed for different types of reviews, such as network meta-analysis or individual patient data reviews.
- There is a need for empiric work on the time and money used in integrating existing systematic reviews into new reviews with comparison to standard methods for new reviews to help guide decisions about when to integrate existing reviews and when to start from scratch.
- Further work is also needed to determine if reviews that integrate existing prior reviews come to different conclusions than reviews conducted from scratch.
- Further research or consensus is needed on specific elements such as:
  - The definition of a systematic review, operationalized to aid in searching and selection.
  - Identifying existing systematic reviews:
    - Is it possible to produce accurate, unbiased and informative systematic reviews by selectively using prior systematic reviews rather than conducting a full comprehensive search for existing reviews? This could be through a sampling mechanism or by prioritizing reviews from particular sources.
    - What type of searching is necessary or optimal? For instance, it is not known if full comprehensive searches, as we conduct for primary studies, are needed to identify systematic reviews. The implications of identifying existing systematic reviews from only certain sources, such as those considered high quality like The Cochrane Collaboration, are also unknown.
  - Evaluating quality of existing systematic reviews, particularly if there are criteria for determining when a prior review may be included or excluded.
- Further work around decisions about synthesis including:
  - A methods study to empirically test approaches for combining new studies with the summary estimate from prior meta-analyses versus with estimates from the individual studies included in the prior meta-analyses.
  - If the review authors choose not to do an updated quantitative synthesis, determining when it is appropriate to use an estimate from a prior review if different meta-analyses methods were used. As new methods are developed or old methods questioned, when are the prior estimates no longer considered reliable? For instance, if the prior review used DerSimonian-Laird model,[19] do new summary estimates need to be obtained using better models? In cases of different models for meta-analysis were used, are there standards that can be established for when the prior estimates would be acceptable or thresholds for determining when a new estimate would be needed?

- Further assessment of how to most appropriately and informatively present reviews that integrate existing reviews needs to be conducted with end users of the reviews or those using our reviews to inform decisionmaking.

Conclusions

The increasing number of systematic reviews, along with the time and money required to undertake a review, has motivated a desire to incorporate existing systematic reviews in a new review. In considering the integration of existing systematic reviews into new reviews, there is a tradeoff between accepting the results of the prior review and needing to either complete again the selected elements of the review or the review in its entirety. The key is to find the right balance in terms of an efficient and unbiased approach to conducting and reporting the integration of existing systematic reviews into the new review. In this working document, we have provided preliminary guidance to help find that balance.

References

1. Whitlock E, Lin J, Chou R, et al. Using existing systematic reviews in complex systematic reviews. Ann Intern Med. 2008;148(10):776-82.
2. White CM, Ip S, McPheeters M, et al. Using Existing Systematic Reviews to Replace De Novo Processes in Conducting Comparative Effectiveness Reviews. In: Agency for Healthcare Research and Quality. Methods Guide for Comparative Effectiveness Reviews. Rockville, MD: Agency for Healthcare Research and Quality; September 2009. http://effectivehealthcare.ahrq.gov.
3. Robinson KA, Whitlock E, O'Neil M, et al. Integration of Existing Systematic Reviews. AHRQ Publication No. 14-EHC016-EF. Rockville, MD: Agency for Healthcare Research and Quality; June 23, 2014. http://effectivehealthcare.ahrq.gov.
4. Robinson KA, Whitlock EP, Oneil ME, et al. Integration of existing systematic reviews into new reviews: identification of guidance needs. Syst Rev. 2014;3(1):1-17.
5. Shea B, Grimshaw J, Wells G, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Medical Research Methodology. 2007;7(1):10.
6. Effective Health Care Scientific Resource Center. Methods Article Alert. Portland: Scientific Resource Center. www.epc-src.org/methods_library/index.cfm. Accessed May 2014.
7. Filters to identify systematic reviews. InterTASC Information Specialists' Sub-Group. https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/filters-to-identify-systematic-reviews. Accessed October 2014.
8. Systematic reviews: CRD's guidance for undertaking reviews in health care: Centre for Reviews and Dissemination; 2009.
9. Higgins J, Green S. Cochrane handbook for systematic reviews of interventions. The Cochrane Collaboration; 2011.
10. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151(4):264-9.
11. Relevo R, Paynter R. Peer review of search strategies. Methods Research Reports. AHRQ Publication No.: 12-EHC068-EF. Rockville, MD: Agency for Healthcare Research and Quality; June 2012.
12. Ahmadzai N, Newberry S, Maglione M, et al. A surveillance system to assess the need for updating systematic reviews. Syst Rev. 2013;2(1):1-16.
13. Chung M, Newberry S, Ansari M, et al. Two methods provide similar signals for the need to update systematic reviews. J Clin Epi. 2012;65(6):660-8.

14. Shekelle PG, Newberry SJ, Wu H, et al. Identifying Signals for Updating Systematic Reviews.: A Comparison of Two Methods. Methods Research Report. AHRQ Publication No. 11-EHC042-EF. Rockville, MD: Agency for Healthcare Research and Quality; June 2011. http://effectivehealthcare.ahrq.gov.

15. Cornell JE, Mulrow CD, Localio R, et al. Random-effects meta-analysis of inconsistent effects: a time for change. Ann Intern Med. 2014;160(4):267-70.

16. Berkman ND, Lohr K, Ansari MT, et al. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. AHRQ Publication No. 13(14)-EHC130-EF. Rockville, MD: Agency for Healthcare Research and Quality; November 2013. http://effectivehealthcare.ahrq.gov.

17. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. Am J Prev Med. 2001;20(3):21-35.

18. Oxman AD, Group GW. Grading quality of evidence and strength of recommendations. BMJ. 2004;328(19):1490-4.

19. DerSimonian R, Kacker R. Random-effects model for meta-analysis of clinical trials: an update. Contemporary clinical trials. 2007;28(2):105-14.

Appendix A. Interview Guide

Introduction

The overall mission of the Agency for Healthcare Research and Quality's (AHRQ) Effective Health Care (EHC) Program is to provide evidence-based information to health care stakeholders that is relevant to their needs, timely, objective, scientifically rigorous in construct, and developed and presented with transparency. In the production of systematic reviews, we aim to answer questions about effectiveness of interventions and average population effects. We are aware that for certain conditions and behavioral interventions, these questions may miss important issues.

AHRQ engages stakeholders in all facets of their research enterprise, including the producing of systematic reviews, with the goals of ensuring that research findings reflect the needs of diverse users, are relevant to their unique challenges, and are applicable in real-world situations.

Purpose of the Discussion Session

The goal of our project is to understand qualitative and quantitative methods for synthesis of evidence based on one or more existing systematic reviews.

We are very interested in learning from your experience.

There are not right or wrong answers, so please feel free to share your thoughts openly.

We would welcome any materials that you would like to share with us either before or after the discussion session. Please send any materials to johanna.anderson2@va.gov.

Ground Rules for Discussion Session

The discussions will be tape recorded, transcribed, and analyzed for overarching themes.

Although the report may list individuals who were interviewed, answers will not be identifiable to individuals or specific organizations.

You may refrain from answering any questions and are welcome to leave the discussion at any time.

Interview Guide

Introduction

There are several scenarios in which an existing systematic review or multiple reviews may be used in a new review. Questions to consider in each of these scenarios are presented below along with general questions to consider when using existing systematic reviews. These scenarios assume that existing reviews for consideration are on point (i.e., relevant PICO) and of "sufficient quality" (i.e., well conducted and well-reported). These scenarios are not mutually exclusive and any of these scenarios may arise alone or in combination in a single review for different review questions, outcomes, and/or comparators.

In this discussion, we will present you with one or more specific examples of reviews using existing systematic reviews. The goal of this discussion is to examine these examples within each scenario and understand how you would address the questions which arise in incorporating existing systematic reviews.

Scenarios

Scenario 1: Use existing review without modifying or adding new studies
Scenario 2: Use review and add new studies
Scenario 3: Use review with new or modified analysis
Scenario 4: Use selected elements of review
Scenario 5: Don't use review

1. Which of these scenarios do you have experience with?
2. Do you have any specific guidance you rely on in using existing systematic reviews?

General Considerations for This Interview

Risk of Bias

1. What factors make it possible to translate/use prior risk of bias (RoB) assessment? What level of detail is needed to help you make this decision?
2. Under what circumstances would you need to complete assessment again?

Strength of Evidence

1. What factors make it possible to translate/use prior strength of evidence (SOE) grading? What level of detail is needed to help you make this decision?
2. Under what circumstances would you need to complete grading again?

Multiple Existing Systematic Reviews: (Example 1)

1. Do you try to use all concordant reviews or are you more selective. If so, what factors do you select?
2. What factors do you use to resolve discrepancies between reviews?

Scenario-Specific Considerations

Scenario 1: Use existing review without modifying or adding new studies (Examples 2 & 5).

1. What factors allow you to use a review without modifications?
2. How do you integrate the existing SR synthesis into a new review?
3. How do you integrate existing SOE into a new review?
4. How do you use existing risk of bias?
5. When is it okay not to add new studies (or conduct a search for new studies)? Very recent review (within 1 year, 2 years, 3 years)? Well established body of evidence in which the findings are unlikely to change with addition of new studies? Lack of resources? Other reasons?

Scenario 2: Use review and add new studies (Examples 3 & 5).

1. How do you integrate the existing SR synthesis into a new review with new studies added?
2. How do you integrate existing SOE and SOE of added studies? Is there a need to complete judgments about strength of evidence again?
3. What is enough, in terms of studies/type of evidence, to prompt a change in grade?
4. How do you integrate existing risk of bias and risk of bias of added studies?

Scenario 3: Use review with new or modified analysis (Example 4)

1. What factors make you want to modify to redo parts of the analysis (different statistical methods, confirm risk of bias ratings or use different method, etc.)?
2. How do you use the existing systematic review synthesis in a modified or new analysis?
3. How do you use the existing strength of evidence in a modified or new analysis? Is there a need to complete judgments about strength of evidence again?
4. What is enough, in terms of studies/type of evidence, to prompt a change in grade?
5. How do you use existing risk of bias in a modified or new analysis?

Scenario 4: Use selected elements of review

1. What factors make you only use selected elements of the review (e.g., problems with the analysis or concerns they missed studies, ongoing controversy, etc.)?
2. What elements might you use (might one only use the included studies or reference lists or some elements of data abstraction)?

Scenario 5: Don’t use review

1. What reasons may cause you to not use a review at all (e.g., few studies, poor quality, etc.)?

Appendix B. Interview Themes

Appendix Table B1. Interview themes: Practices and opinions on using existing reviews

Overall Themes	
Multiple Existing Reviews	
	Use the 'best' review per question rather than including all systematic reviews
	• Choose the review that is the best match for scope and PICOTS and that is of the highest quality, includes the most recent studies and has no perceived bias due to conflict of interest.
	• Adequate details need to be reported to effectively rate the review
	• AMSTAR, though not perfect, has been used to assess the quality of existing reviews.
	Discordant reviews is a signal to conduct own review if discordance cannot easily be explained
	Discuss existing discordant reviews in discussion section
Risk of Bias for Individual Studies	
	An existing review must have completed some sort of risk of bias assessment of primary studies in order to be used.
	An existing review must have used an accepted and validated tool.
	Risk of bias assessment methods need to be transparent
	May confirm assessment by redoing a few studies. If confirmed, will accept the risk of bias assessment of all studies in an existing review.
	If the risk of bias assessment needs to be redone, the existing review will not be used
Grading Strength of Evidence	
	Practices range from not conducting grading strength of evidence, using the existing review's grading to always using own grading criteria
	If using existing review's grading, the methods need to be transparent
	Once one domain of the SOE is called into question, the whole SOE needed to be redone.
	Some will not use an existing review if the grading needs to be re-done
Scenario-Specific Themes	
Scenario 1	Use review without modifying or adding new studies
	Will use a review with no changes or new studies added if: • Current in the context of the research question and includes all relevant studies • Matches of PICOTS, scope, study designs • Meets quality standard—it helps if an existing review is from a trusted source. • Methods are transparent. • No conflict of interest
	Synthesis is qualitative with narrative summary of the review's results and a critique of the limitations and strengths of the review.
	Will use existing review's risk of bias assessment
	Will use existing review's grading
	Factors to consider when it is fine not to add new studies: recency, new studies not likely changing the results and lack of resources. • One organization chooses to do a rapid review.

Appendix Table B1. Interview themes: Practices and opinions on using existing reviews (continued)

Scenario 2	Use review and add new studies
	It is common to use an existing review with a bridge search.
	The way to integrate the existing SR synthesis into a new review with new studies may depend on the amount and quality of the new studies and the purpose of the review.
	Synthesis can range from a qualitative comparison of the existing review's results with analysis of new studies to adding the new studies and rerunning the existing analysis.
	Will use the review's risk of bias results as long as review used acceptable tool and has transparent methods
	If an organization does grade SOE, there is usually some kind of effort of re-do the SOE to incorporate the new studies. There is difference between updating one's own review vs. updating other's review. Grading practices range from grading all studies using own criteria to grading the new studies and comparing that to the grading in the existing review.
Scenario 3	Use review with new or modified analysis
	Will do a new analysis if the existing analysis does not meet the new standard; or has a more general scope.
	The data synthesis will need to consider the impact of the new analysis or new methods.
Scenario 4	Use selected elements of review
	May not use the review at all if only selected elements could be used.
	May use existing review's search strategy, reference list or summary if chose not to use it as a whole
	May use the data for the subgroup of interest
	Decision was made on a case-by-case basis.
	Only using certain parts of an existing review may introduce bias
Scenario 5	Do not use review
	Will not use an existing review if: • Different scope • Inadequate quality • Outdated • Funded by industry/ Conflict of interest • Lack of transparency • Only selected elements could be used. • "Empty review" with very little evidence.
	May discuss existing reviews that are not used in the discussion section