

EPC Response to IOM Standards for Systematic Reviews



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

EPC Response to IOM Standards for Systematic Reviews

Prepared for:

Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. 290-2007-10055-I

Prepared by:

Tufts Evidence-based Practice Center, Tufts Medical Center
Boston, MA

Investigators:

Joseph Lau, M.D.
Stephanie Chang, M.D., M.P.H.
Nancy Berkman, Ph.D.
Sara J. Ratichek, M.A.
Howard Balshem, M.S.
Michelle Brasure, Ph.D.
David Moher, Ph.D.

This report is based on research conducted by the Tufts Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2007-10055-I). 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.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

This document is in the public domain and may be used and reprinted without permission except those copyrighted materials that are clearly noted in the document. Further reproduction of those copyrighted materials is prohibited without the specific permission of copyright holders.

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: Lau J, Chang S, Berkman N, Ratichek SJ, Balshem H, Brasure M, Moher D. EPC Response to IOM Standards for Systematic Reviews. Research White Paper (Prepared by the Tufts Evidence-based Practice Center, Tufts Medical Center under Contract No. 290-2007-10055-I). AHRQ Publication No. 13-EHC006-EF. Rockville, MD: Agency for Healthcare Research and Quality; April 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

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.

To improve the scientific rigor of these evidence reports, AHRQ supports empiric research by the EPCs to help understand or improve complex methodologic issues in systematic reviews. These methods research projects are intended to contribute to the research base in and be used to improve the science of systematic reviews. They are not intended to be guidance to the EPC program, although may be considered by EPCs along with other scientific research when determining EPC program methods guidance.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality. The reports undergo peer review prior to their release as a final report.

We welcome comments on this Methods Research Project. 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.

Carolyn M. Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

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

Acknowledgments

The authors gratefully acknowledge the contributions provided to this report by all participants in the EPC Workgroups on IOM Standards (see Appendix A), and all EPC investigators who provided the initial EPC response to the practice of the IOM Standards (see Appendix B).

EPC Response to IOM Standards for Systematic Review

Structured Abstract

Background. The Institute of Medicine appointed an independent committee of experts to assess and to recommend a set of methodological standards that would assure objective, transparent, and scientifically valid systematic reviews of comparative effectiveness research. Following the release of these standards in March 2011, the EPC program established a collaborative process to comparatively examine the standards with respect to general EPC practice and guidance.

Purpose. The purpose of this process was to assess which elements of the IOM standards should be adopted into EPC methods guidance and how to best implement these changes, and which elements require further empirical evidence.

Methods. A two-phase approach is adopted, where in phase one, 13 EPC directors, in consultation with their respective staff, identified areas where general agreement exist and where further deliberation was necessary, and in phase two, workgroups, consisting of EPC investigators, were tasked to further deliberate and provide a disposition of each of those elements. These elements were categorized into one of four topic groups: program policies or procedures, protocol elements, searching/screening/reporting biases, and synthesis of evidence. Based on current practices and through discussions, four workgroups determined whether there was “agreement,” “agreement with modifications,” or “disagreement” for each element. Where there were modifications recommended, each workgroup provided a description of the differences between EPC practice or methods guidance and IOM guidelines, summarized the deliberative discussion, and made recommendations for further action.

Results. EPC directors identified 34 elements across the 21 standards that required remediation and assignment to one of the four workgroups. Workgroups described general agreement with the majority of these 34 elements in principle. There were three elements with which the EPCs were in disagreement, and were not recommended for routine practice as currently stated. Discussion on the remaining 31 elements pointed out inconsistency of practice and need for clearer guidance, the need for more empiric evidence in some cases, the difficulty in balancing benefits and the required resources for implementation of elements, and in some cases, specific suggestions on how to implement particular elements.

Conclusions/Recommendations. This process engaged the EPCs in a productive, collaborative evaluation and response to the work of the IOM committee for systematic review standards. Recommendations for further research or development, for updating of EPC guidance, and specific recommendations for practical implementation were itemized. Principally, this process will result in improvements in EPC practice.

Contents

Introduction	1
Methods	2
Results	3
Workgroup on Program Policies and Procedures (“Program Workgroup”).....	8
Description of the Differences Between EPC Practice/EPC Methods Guidance and IOM Guidelines.....	8
Summary of Deliberative Discussions.....	9
Recommendations/Actions/Changes to EPC Procedures and Policies.....	13
Workgroup on Protocol Elements (“Protocol Workgroup”)	13
Description of the Differences Between EPC Practice/EPC Methods Guidance and IOM Guidelines.....	13
Summary of Deliberative Discussions.....	13
Recommendations/Actions/Changes to EPC Procedures and Policies.....	14
Workgroup on Searching, Screening and Reporting Bias (“Searching Workgroup”).....	15
Description of the Differences Between EPC Practice/EPC Methods Guidance and IOM Guidelines.....	15
Summary of Deliberative Discussions.....	15
Recommendations/Actions/Changes to EPC Procedures and Policies.....	18
Workgroup on Synthesizing Evidence (“Synthesis Workgroup”).....	19
Description of the Differences Between EPC Practice/EPC Methods Guidance and IOM Guidelines.....	19
Summary of Deliberative Discussions.....	19
Recommendations/Actions/Changes to EPC Procedures and Policies.....	21
Discussion	22
Conclusion and Recommendations.....	24
References	25

Tables

Table 1. Summary of phase 2 discussion by workgroup (N=34)	3
Table 2. EHC approach to identifying and managing bias in conducting SRs.....	11

Appendixes

Appendix A. Workgroups

Appendix B. EPC Directors’ Summary of Responses to IOM Standards for Systematic Reviews

Appendix C. Summary of Decisions and Protocol for Resolution in the Event of Difference
of Opinion Between EPC Report Authors and AHRQ Officer

Introduction

The Institute of Medicine (IOM), the health arm of the National Academy of Sciences, was established to provide independent, objective science-based advice to policymakers, health professionals, the private sector, and the public with regards to health care. At the request of Congress, in an effort to address the lack of universally accepted standards and the variations in addressing conflicts of interest, appraisals of evidence, and the overall rigor of evaluations, an independent committee of 16 experts was appointed by the IOM and charged to assess potential methodological standards that would assure objective, transparent, and scientifically valid systematic reviews (SRs) of comparative effectiveness research (CER), and to recommend a set of methodological standards for developing and reporting such SRs. The resulting report on standards for SR was subsequently released in March 2011.

The committee defined a standard as “a process, action, or procedure that is deemed essential to producing scientifically valid, transparent, and reproducible results.”¹ The IOM report includes 21 standards, with 82 distinct elements of performance (defined as essential components of the standards.) While acknowledging that the recommended standards and elements are provisional, pending better empirical evidence, they are based on current evidence, expert guidance, and thoughtful reasoning using an a priori set of criteria for assessment, and are therefore considered to be current “best practices.” In particular, the committee reviewed the published methods manuals of leading SR groups including the Agency for Healthcare Research and Quality (AHRQ), the Centre for Reviews and Dissemination (University of York, UK), and the Cochrane Collaboration. They additionally conferred with experts at the Drug Effectiveness Review Project, the Emergency Care Research Institute, the National Institute for Health and Clinical Excellence (UK), and several AHRQ Evidence-based Practice Centers (EPCs), to finalize the list of essential steps and considerations in the SR process. Members of several of the EPCs also served on the report committee as members, consultants or peer reviewers.

Following the release of the IOM standards, the EPC program established a process to comparatively examine the standards, and, more specifically, each of their elements of performance, with respect to general EPC practice and guidance. Many elements of performance in the IOM report are already intrinsic to the EPC program. Other elements lack supporting empirical data but possess an intuitive validity. Others are difficult to implement practically. The overarching agenda of this examination was to assess which elements should be adopted into EPC methods guidance; determine the best methods for implementing these changes, including dissemination and training; and lay out a process of evaluating adherence. Where standard recommendations are determined to require further empirical evidence, methods research should be conducted for validation and improvement. The present report describes and summarizes the collaborative process undertaken by the EPCs in evaluating the IOM standards, elements of performance, and the resulting conclusions.

Methods

The IOM report organized the standards into four chapters:

- Standards for Initiating an SR (consisting of 8 standards with 28 elements)
- Standards for Finding and Assessing Individual Studies (consisting of 6 standards with 30 elements)
- Standards for Synthesizing the Body of Evidence (consisting of 4 standards with 13 elements)
- Standards for Reporting an SR (consisting of 3 standards with 11 elements)

The following methods were employed to identify and resolve discrepancies between these 82 elements of the standards for SRs and the EPC Methods Guide and current EPC practice, acknowledging the latter two entities may be distinct. During an initial phase (phase 1), all EPC directors, in consultation with their staff, were requested to categorize each standard and element into one of three categories:

1. Currently practiced by EPC with little/no variation; no further discussion needed
2. Common EPC practice with some variations; some discussion needed before adoption
3. Uncommon EPC practice or with substantial variations; much discussion and/or more evidence needed before adoption.

The feedback from all EPCs was compiled and presented in May 2011 at the EPC directors' meeting, along with an outline of the remaining steps in the process. The rationale for both adhering to the IOM standards and for a considered response prior to automatic adoption were also outlined in this presentation. Adherence would result in reducing variability of practice, increasing transparency, improving quality, and clarifying an authoritative source in response to criticisms. Reasons for a considered response were because standards were developed based on structured consensus with reference to empirical data only when available, standards may be infeasible to adapt, and methods may be changing as technologies evolve.

The second phase (phase 2) of the process focused on the standards or elements in either category 2 or 3 (above) where less than 80 percent of EPCs reported the element as common practice currently with little or no variation. Items in these categories were grouped into four topics loosely corresponding to the four chapters of the IOM standards. Work groups to address the following topics were formed, with representative participation across EPCs.

1. program policies or procedures,
2. protocol elements,
3. searching/screening/reporting biases, and
4. synthesis of evidence

The list of workgroup participants can be found in Appendix A. Each workgroup was tasked with evaluating these category 2 or 3 items according to current EPC methods guidance or procedural protocol, and determining an appropriate disposition. Teleconference meetings held over a period of 4 months were the primary forums of discussion on each assigned item. For each, the relevant current EPC guidance was identified, and the disposition collectively determined as either "agreement," "agreement with modifications," or "disagreement." In the cases where modification was recommended, updates to the EPC methods guide, or improved adherence to the methods guidance were suggested.

Results

The results of phase 1 of this exercise are summarized in the EPC Directors’ Summary of Responses to IOM standards in Appendix B. A total of 34 elements (41%), spread across the 21 standards, were determined to require remediation and subsequently disseminated to one of the four topic work groups of phase 2. Table 1 provides a summary of the workgroup discussion and dispensation for each of these 34 elements. For example, the Program policies and procedures workgroup selected “Agree with Modification” as a dispensation for Element 2.2.3 because, while they agree in principle with the importance of excluding review team members who have professional or intellectual bias, they raised the discussion point that intellectual bias overlaps with expertise, which is needed in conducting an evidence review. They suggest both that intellectual bias can be addressed by balancing intellectual expertise/perspectives, and also by establishing methods to assist others in identifying their own biases.

Additionally, in some instances where the EPC workgroup agreed with the element, added discussion points were noted (in the “Key Discussion Point” column of the table.) The details of these discussions, including workgroup recommendations, follow below Table 1.

Table 1. Summary of phase 2 discussion by workgroup (N=34*)

Elements Requiring Discussion for Possible Remediation	Workgroup Section	Dispensation	Key Discussion Point
STANDARD 2.2 Manage bias and conflict of interest (COI) of the team conducting the SR 2.2.3: Exclude individuals whose <i>professional or intellectual</i> bias would diminish the credibility of the review in the eyes of the intended users.	1. Program policies & procedures	Agree with Modification	There is no measure for bias and intellectual bias may also be content expertise. In some cases it may be more appropriate to balance rather than exclude. Further work needs to be done to help individuals identify their own biases.
STANDARD 2.3 Ensure user and stakeholder input as the review is designed and conducted 2.3.1: Protect the independence of the review team to make the final decisions about the design, analysis, and reporting of the review.	1. Program policies & procedures	Agree	EPC reports are conducted under contract. AHRQ officers are responsible to ensure adherence to the contract requirements.
STANDARD 2.4 Manage bias and COI for individuals providing input into the SR 2.4.2: Exclude input from individuals whose COI or bias would diminish the credibility of the review in the eyes of the intended users.	1. Program policies & procedures	Agree with Modification	Allow public input during formulation of key questions and on draft report. Minimize or balance financial and nonfinancial COI for solicited input on key questions and protocol. Solicit input from individuals without financial COI on draft report. Ensure that any nonfinancial COI are appropriately balanced.

Table 1. Summary of phase 2 discussion by workgroup (N=34*) (continued)

Elements Requiring Discussion for Possible Remediation	Workgroup Section	Dispensation	Key Discussion Point
STANDARD 2.5 Formulate the topic for the SR 2.5.1: Confirm the need for a new review.	1. Program policies & procedures	Agree	Best efforts to do this are undermined by lack of publication or registration of protocols
2.5.3: Use a standard format to articulate each clinical question of interest.	1. Program policies & procedures	Agree	Reports other than for treatment effectiveness may not have an established standard format. For improved presentation, the EPC may elect to summarize clinical questions with delineation of the PICOTS separately.
2.5.4: State the rationale for each clinical question.	1. Program policies & procedures	Agree	This refers to the clinical rationale for the question, which is developed through a structured process with public and solicited input.
STANDARD 2.6 Develop a SR protocol 2.6.1: Describe the context and rationale for the review from both a decisionmaking and research perspective.	2. Protocol Elements	Agree	“From a research perspective” refers to existing systematic reviews on a given topic.
2.6.9: Describe the method for evaluating the body of evidence, including the quantitative and qualitative synthesis strategies.	2. Protocol Elements	Agree	The amount of detail required to describe methods is unclear. Methods may need to be adjusted based on the type of research and data uncovered in the search and data extraction.
2.6.10: Describe and justify any planned analyses of differential treatment effects according to patient subgroups, how an intervention is delivered, or how an outcome is measured.	2. Protocol Elements	Agree with Modification	“Justification” may include citing biological plausibility, interest by stakeholders, or previous data to support heterogeneity in treatment effect. Potentially meaningful subgroup analyses that arise once the review is underway should be described as post-hoc in the review, as per IOM element 5.1.6.
STANDARD 2.7 Submit the protocol for peer review Standard 2.7.1 Provide a public comment period for the protocol and publicly report on disposition of comments.	1. Program policies & procedures	Disagree	The EPC program provides public comment periods on the separate elements that comprise the protocol—the specific Key Questions, and the Methods Guide chapters that drive most protocol elements.
Standard 2.8: Make the final protocol publicly available, and add any amendments to the protocol in a timely fashion.	1. Program policies & procedures	Agree	Major changes such as changes in scope should be documented as a protocol amendment. Minor changes such as a change in search strategy may not require amending the publicly posted protocol and could be documented in the report.

Table 1. Summary of phase 2 discussion by workgroup (N=34*) (continued)

Elements Requiring Discussion for Possible Remediation	Workgroup Section	Dispensation	Key Discussion Point
<p>STANDARD 3.1 Conduct a comprehensive systematic search for evidence 3.1.3 Use an independent librarian or other information specialist to peer review the search strategy.</p>	<p>3. Searching, Screening & reporting bias</p>	<p>Disagree</p>	<p>Peer review of search strategies by an experienced reviewer within the EPC should be considered sufficient until there is empirical evidence that review by an independent librarian or information specialist would be more effective.</p>
<p>3.1.5 Search citation indexes.</p>	<p>3. Searching, Screening & reporting bias</p>	<p>Agree with Modification</p>	<p>Citation searching may be warranted in cases of landmark trials or when updating reviews, but there is insufficient evidence that as a general practice it results in additional included studies.</p>
<p>3.1.9 Search regional bibliographic databases if other databases are unlikely to provide all relevant evidence.</p>	<p>3. Searching, Screening & reporting bias</p>	<p>Agree</p>	<p>There are certain circumstances where searches of regional databases may be warranted, such as when abstracts from certain geographical regions consistently show results different from those found in major, mainstream journals. Consultation with experts could help determine when searches of regional databases will be useful.</p>
<p>Standard 3.2 Take action to address potentially biased reporting of research results 3.2.1 Search grey literature databases, clinical trials registries, and other sources of unpublished information about studies.</p>	<p>3. Searching, Screening & reporting bias</p>	<p>Agree with Modification</p>	<p>Certain sources of grey literature, like ClinicalTrials.gov, FDA documents, and EPC evidence reviews, will likely be of use in most reviews. However, other sources should be included at the discretion of the review team.</p>
<p>3.2.2 Invite researchers to clarify information about study eligibility, study characteristics, and risk of bias.</p>	<p>3. Searching, Screening & reporting bias</p>	<p>Agree with Modification</p>	<p>Contacting authors may be beneficial when clarification regarding study eligibility, design, or conduct may affect review conclusions.</p>
<p>3.2.3 Invite all study sponsors and researchers to submit unpublished data, including unreported outcomes, for possible inclusion in the SR.</p>	<p>3. Searching, Screening & reporting bias</p>	<p>Agree with Modification</p>	<p>The Scientific Resource Center invites industry to submit unpublished data on behalf of the EPCs. Expanding invitations to all study sponsors and researchers may create significant additional work with little or unknown value.</p>
<p>3.2.4 Handsearch selected journals and conference abstracts.</p>	<p>3. Searching, Screening & reporting bias</p>	<p>Agree with Modification</p>	<p>Hand searches can be important when it has been determined that studies are likely to be found in a journal not indexed in bibliographic databases.</p>
<p>3.2.5 Conduct a Web search.</p>	<p>3. Searching, Screening & reporting bias</p>	<p>Disagree</p>	<p>This standard is too vague to understand its intent.</p>
<p>3.2.6 Search for studies reported in languages other than English if appropriate.</p>	<p>3. Searching, Screening & reporting bias</p>	<p>Agree</p>	<p>Because there is insufficient empirical evidence on when it is appropriate to search for studies in languages other than English, this will be a decision left to the judgment of the review team.</p>

Table 1. Summary of phase 2 discussion by workgroup (N=34*) (continued)

Elements Requiring Discussion for Possible Remediation	Workgroup Section	Dispensation	Key Discussion Point
<p>STANDARD 3.3 Screen and select studies 3.3.3 Use two or more members of the review team, working independently, to screen and select studies.</p>	3. Searching, Screening & reporting bias	Agree with Modification	Dual screening is ideal; however it is important to focus on assuring quality control in the screening and selection process, especially when dual review is not feasible.
3.3.4 Train screeners using written documentation; test and retest screeners to improve accuracy and consistency.	3. Searching, Screening & reporting bias	Agree	Implementation of methods to achieve accuracy and consistency is difficult.
3.3.6 Taking account of the risk of bias, consider using observational studies to address gaps in the evidence from randomized clinical trials on the benefits of interventions.	3. Searching, Screening & reporting bias	Agree	Evidence from observational studies should be included when evidence from RCTs does not fully answer key questions of the review.
<p>STANDARD 3.5 Manage data collection 3.5.1 At a minimum, use two or more researchers, working independently, to extract quantitative and other critical data from each study. For other types of data, one individual could extract the data while the second individual independently checks for accuracy and completeness. Establish a fair procedure for resolving discrepancies—do not simply give final decisionmaking power to the senior reviewer.</p>	3. Searching, Screening & reporting bias	Agree with Modification	Dual abstraction is ideal, however when dual abstraction is not possible because of resource limitations, fact checking, having one researcher extract the data and a second independent researcher check the extracted data against the full text article is sufficient. A process for resolving differences is essential, but resolution by a senior investigator alone is not sufficient.
<p>STANDARD 3.6 Critically appraise each study 3.6.3 Assess the fidelity of the implementation of interventions.</p>	4. Synthesizing Evidence	Agree	Fidelity may be difficult to evaluate because there is no standard in the field or when the information that would be needed to determine fidelity is inconsistently defined and/or inadequately described across studies.
<p>STANDARD 4.1 Use a prespecified method to evaluate the body of evidence 4.1.1 For each outcome, systematically assess the following characteristics of the body of evidence:</p> <ul style="list-style-type: none"> • Risk of bias • Consistency • Precision • Directness • Reporting bias 	4. Synthesizing Evidence	Agree with Modification	EPCs should identify a priori which “major” outcomes are considered important enough to warrant formal grading of the strength of the evidence, depending on the key questions, the clinical or policy context, and the purpose of the report. Reporting bias refers to both publication bias and selective outcome reporting. There are no established methods for assessing this, but reviewers should consider this domain particularly when there is high and moderate strength bodies of evidence.

Table 1. Summary of phase 2 discussion by workgroup (N=34*) (continued)

Elements Requiring Discussion for Possible Remediation	Workgroup Section	Dispensation	Key Discussion Point
4.1.2 For bodies of evidence that include observational research, also systematically assess the following characteristics for each outcome: <ul style="list-style-type: none"> • Dose-response association • Plausible confounding that would change the observed effect • Strength of association 	4. Synthesizing Evidence	Agree	Each of these concerns may not always be relevant to the body of literature.
4.1.3 For each outcome specified in the protocol, use consistent language to characterize the level of confidence in the estimates of the effect of an intervention.	4. Synthesizing Evidence	Agree with Modification	The protocol will specify all outcomes to be reviewed as well as the outcomes to be graded. The EPC will characterize the level of confidence in the estimate of effect only for those outcomes specified in the protocol as most critical to be graded.
STANDARD 4.3 Decide if, in addition to a qualitative analysis, the SR will include a quantitative analysis (meta-analysis) 4.3.1 Explain why a pooled estimate might be useful to decision makers.	4. Synthesizing Evidence	Agree	EPC guidance is presented in relation to ensuring that meta-analysis produces a meaningful and interpretable result based on available data. Because review authors are not working directly with guideline panels and policy makers, it is not always clear what “might be useful to decision makers.”
STANDARD 4.4 If conducting a meta-analysis, then do the following: 4.4.1 Use expert methodologists to develop, execute, and peer review the meta-analyses.	4. Synthesizing Evidence	Agree	Although it is important that at least one of the peer reviewers should be an expert methodologist when one or more meta-analyses are included in a review, it may be beyond the control of the review authors to obtain agreement from an expert methodologist to review the report.
4.4.4 Assess the sensitivity of conclusions to changes in the protocol, assumptions, and study selection (sensitivity analysis).	4. Synthesizing Evidence	Agree	Better and more detailed guidance may improve consistency in implementation.
STANDARD 5.1 Prepare final report using a structured format 5.1.4: Include a summary written for the lay public.	1. Program policies & procedures	Agree	The John M. Eisenberg Center for Clinical Decisions and Communication Science develops summary guides for different audiences (clinician, consumer, and policy-maker) for many EHC reports as appropriate for the topic. Resource may constrain development for all products and may require partnership with other outside organizations.

Table 1. Summary of phase 2 discussion by workgroup (N=34*) (continued)

Elements Requiring Discussion for Possible Remediation	Workgroup Section	Dispensation	Key Discussion Point
STANDARD 5.2 Peer review the draft report 5.2.1: Use a third party to manage the peer review process.	1. Program policies & procedures	Agree	AHRQ officers evaluate and invite peer reviewers without conflicts of interest.
5.2.2: Provide a public comment period for the report and publicly report on disposition of comments.	1. Program policies & procedures	Agree	Public posting is standard practice for the EHC program. Building in time and resources for a public comment period is a priority and currently under development for other programs for which EPCs conduct reviews.

AHRQ = Agency for Healthcare Research and Quality; COI = conflict of interest; EHC = ; EPC = Evidence-based Practice Center; FDA = U.S. Food and Drug Administration; IOM = Institute of Medicine; PICOTS = populations, interventions, comparators, outcomes, timing, and setting; SR = systematic review

*34 Standards identified during phase 1 of this study (see Methods).

Each workgroup provided a summary (below) of their discussions according to the following outline:

- A description of the difference between EPC practice/EPC methods guidance and IOM guidelines
- A summary of the deliberative discussion
- Recommendations/Actions

Workgroup on Program Policies and Procedures (“Program Workgroup”)

Description of the Differences Between EPC Practice/EPC Methods Guidance and IOM Guidelines

EPCs conduct SRs under contract from AHRQ with support from AHRQ and other AHRQ contractors, including the Scientific Resource Center and the John M Eisenberg Center for Clinical Decisions and Communications Science. These reviews are commissioned on behalf of various AHRQ initiatives, such as the Effective Health Care (EHC) Program, the Technology Assessment Program, the U.S. Preventive Services Task Force, as well as on behalf of other Federal agencies. Individual EPCs operate within the constraints of each of these AHRQ initiatives’ established policies and procedures. The EHC program in particular is one of the largest funders of SRs (description of EHC program can be found at: www.effectivehealthcare.ahrq.gov). Specific policies (some of which are based on Federal regulations) are defined within the contract signed by an individual EPC and AHRQ. This workgroup focused on the agreement between the IOM standards and the EPC program conduct within the EHC program.

The EHC Program is driven by the principles of relevance, timeliness, objectivity, and scientific rigor, and by incorporating public participation and transparency², similar to the IOM described principles of Acceptability, Applicability, Efficiency, Patient-Centeredness, Scientific rigor, Timeliness, and Transparency.³

During phase 1, EPC directors identified 11 elements from the IOM standards relating to program policies or procedures as described above. Of these, only one was not yet an established

standard practice by the EHC program: Standard 2.7.1 that asks that SR teams provide a public comment period for the protocol and publicly report on disposition of comments.

Two of the 11, which related to managing or mitigating bias in the composition of review teams and their conduct, were determined to require some clarification and modification for adoption with EHC program philosophy and practice: Standard 2.2.3: Exclude individuals whose professional or intellectual bias would diminish the credibility of the review in the eyes of the intended users; and Standard 2.4.2: Exclude input from individuals whose COI or bias would diminish the credibility of the review in the eyes of the intended users. Discussion and clarification of these elements resulted in a recommendation for an update to and clarification of EHC processes for these standards.

The remaining eight elements required further discussion or minor clarifications to IOM standards or EHC processes: Standard 2.3.1: Protect the independence of the review team to make the final decisions about the design, analysis, and reporting of the review; Standard 2.5.1: Confirm the need for a new review; Standard 2.5.3: Use a standard format to articulate each clinical question of interest; 2.5.4) State the rationale for each clinical question; Standard 2.8: Make the final protocol publicly available, and add any amendments to the protocol in a timely fashion; Standard 5.1.4: Include a summary (of the final report) written for the lay public; Standard 5.2.1: Use a third party to manage the peer review process; 5.2.2) Provide a public comment period for the report and publicly report on disposition of comments. A summary of the discussions concerning these elements, and the recommendations for clarifications or changes in EHC procedures and policies is provided below.

Summary of Deliberative Discussions

IOM Elements in Which There Is Disagreement

2.7.1: Provide a public comment period for the protocol and publicly report on disposition of comments.

Program workgroup members discussed the purpose for public comment on the protocol, which is presumably to increase the transparency and usability of a review. While these are worthwhile goals, the time and effort may be prohibitive, especially when they can be accomplished with alternative methods.

The EHC program engages in a transparent process and engages experts and the public for developing each part of the protocol—both the standard elements and the individual questions—and provides a public comment period for each these elements. The EHC Methods Guide² describes the recommended EPC approach for conducting a SR. Each Methods Guide chapter is peer reviewed and posted for public comment prior to being adopted as final guidance. Key questions specific to each review define the scope of the reviews and drive any slight nuances or differences from the standard in the methodologic approach. These individual key questions are developed with input from key informants and posted for public comment. Also posted is the final protocol, which documents changes to the key questions (based on public comments) and specific plans for adherence to the methods guide, as well as any particular nuances or deviations that derive naturally from the key questions.

IOM Elements Requiring Modification or Clarification

2.2.3: Exclude individuals whose professional or intellectual bias would diminish the credibility of the review in the eyes of the intended users.

Program workgroup members agreed that because individual judgment is an essential element of SRs, any bias of the review team or those providing input can introduce bias into the conclusions of the report. However, in reality, bias is difficult to measure. Individuals may be biased because of a COI,⁴ such as those due to professional experiences or affiliations, but COI does not automatically implicate a biased perspective. While financial COI may be generally accepted proxies for increased risk of bias, the association between non-financial COI and risk of bias is less well established.

Furthermore, eliminating such conflicts may inappropriately restrict the ability to meet another IOM standard—the need for appropriate clinical expertise (Standard 2.1). Appropriate clinical understanding and context expertise requires time and dedicated professional effort and experience. Simply having an affiliation or an established scientific or policy position does not preclude the ability to reform or take an alternate stance in the face of evidence.

Instead, Program workgroup members assert that it is more important to carefully consider nonfinancial conflicts of interest or other indications of preconceived beliefs in terms of the ability to take a fresh unbiased look at the evidence. In particular, because nonfinancial conflicts of interest are imperfect measures of potential bias, it is more important to manage and balance nonfinancial conflicts of interest rather than exclude participation outright.

Program workgroup members agreed that previous publication of an opinion piece might suggest such entrenched views that a particular individual be excluded from the core team if his or her perspective cannot be appropriately counterbalanced with other views. Program workgroup members also agreed that if SR teams included individuals whose primary studies were eligible for inclusion (assuming their participation in the review team is appropriately counterbalanced), that these individuals should not be involved in grading their own study, either at the individual-study level or at the body-of-evidence level.

2.4.2: Exclude input from individuals whose COI or bias would diminish the credibility of the review in the eyes of the intended users.

Program workgroup members disagreed with the blanket application of this element, as it conflicts with three other IOM Elements—2.5.5, which asks for user and stakeholder input on the key questions; 2.7.1, which asks for public comment on the protocol; and 5.2.2, which asks for public comment on the draft report—all of which highlight the importance of broad input and feedback without restriction for potential biases. Program workgroup members felt that solicited input, as well as public input, is beneficial at each stage, and that EPC authors should be aware of these influences and the potential biases of those providing input.

AHRQ requires disclosure of financial COI as defined and directed by the Code of Federal Regulations.⁵ Financial conflicts of interest would prohibit participation in the core team as authors. However, when soliciting input into the key questions and protocol, because of the limited expertise in some areas, financial conflicts should be limited or managed as appropriate. For example, the input may be balanced such that those with financial interest in all of the interventions under study are represented. Invited peer reviewers, however, should have no direct financial COI.

As described above and illustrated in Table 2, the EPC workgroup felt that many nonfinancial conflicts of interest due to professional affiliations could be balanced. Participation

in a study or trial should not be considered a COI, since the role of those providing solicited input is not to assess studies but to provide input on the clinical context and the approach of the review.

Table 2. EHC approach to identifying and managing bias in conducting SRs

Proxies for Bias	Core Team in Authoring Review	Solicited Input for Key Questions	Solicited Input for Protocol	Solicited Input for Draft Report
Financial COI	Exclude if identified	Exclude or balance if necessary*	Exclude or balance if necessary*	Exclude if identified
Nonfinancial COI (professional affiliations)	Balance if identified	Balance if identified	Balance if identified	Balance if identified
Established beliefs	Avoid or balance if identified	Balance if identified	Balance if identified	Balance if identified
Involvement in study included in the review	Exclude or restrict role to refrain from grading own study at individual-study or body-of-evidence level	Balance if identified	Balance if identified	Include and balance if identified

*In some cases such as rare diseases, it may be difficult to find appropriate experts without financial affiliations. Input is limited to protocol development and not results analysis.

Essential Agreement With IOM Elements of Performance

2.3.1: Protect the independence of the review team to make the final decisions about the design, analysis, and reporting of the review.

With the emphasis from other IOM standards on broad input, program workgroup members felt it important to emphasize the responsibility and independence of the review team as authors of the report. While input from a broad range of stakeholders can provide authors with a range of perspectives (and biases), ultimately the authors of the report alone are responsible for the conduct and conclusions of the report and should be able to defend the design and analysis contained within. Public and solicited input is only input, and neither subvert or obviate the authors' responsibility to or authority over the review. Other possible influences on authors may include sponsors or editors.

Under the EHC program, EPCs are commissioned to conduct research under contract. The contract mechanism means that the scientific work is conducted under mutually agreed upon parameters, and AHRQ officers are responsible for ensuring that the conditions and contract requirements are fulfilled, and that the quality of the report meets the standards of the EPC program. Specifically, AHRQ officers ensure that the report is internally consistent, in compliance with EPC program methods (as outlined in the EPC Methods Guide), and responsive to public and solicited input.

2.5.1: Confirm the need for a new review.

2.5.3: Use a standard format to articulate each clinical question of interest.

A standard format helps readers understand the question and ensures that key parts of the question are not overlooked. SRs besides those for clinical treatment effectiveness (i.e., technical briefs, prevalence questions, or medical tests) may not have an established format, but the IOM standards are meant only to relate to SRs of treatment interventions. Program workgroup members agree that key questions should include standard information that clearly defines the

elements of PICOTS: Population(s), Intervention(s), Comparator(s), Outcome(s), Timing and Setting(s). For greater clarity, the EPC may also define a clinical question more broadly, with detailed questions including the PICOTS described subsequently.

2.5.4: State the rationale for each clinical question.

Program workgroup members agreed with these recommendations, but discussed challenges in their efforts to implement them. EPCs work with key informants and conduct preliminary literature searches to determine the need for a review and document this rationale in the background. EPCs can identify recently published reports, but lack of public accessibility of protocols in general mean that new reviews may still be published during the course of a review. To avoid this problem, the EHC program supports IOM recommendation 2.8 that recommends making the final protocol publicly available.

2.8: Make the final protocol publicly available, and add any amendments to the protocol in a timely fashion.

The purpose of a public protocol is to reduce reporting bias by documenting methodological approaches a priori, to describe the scope and process of the review for other systematic reviewers to consider as they embark on their own projects. Having a protocol document available also helps to establish a common understanding of the project with the involved entities, including funders or partners (AHRQ, other Federal agencies, or those translating or reformulating the report for dissemination). However, the IOM standard does not clearly define what would constitute the need for a protocol amendment. Program workgroup members suggested that changes in the protocol that require amendment include modifications to the scope of the questions, inclusion or exclusion criteria, or methodology that will be used to assess, analyze, or grade the literature. Minor changes, such as those to the search strategy, or small clarifications of definitions of particular terminology, data extraction items, sub-analyses, etc. may be documented in the draft report.

5.2.1: Use a third party to manage the peer review process.

Presumably, using a third party ensures that the final decision for selecting peer reviewers and assessing the adequacy of response to comments is independent from the authors. Within the EHC program, the decision for selecting and inviting peer reviewers does not remain with the EPC authors, but with AHRQ staff, who are responsible for identifying peer reviewers without conflicts of interest as described above. The administrative tasks of collecting and collating the comments may be done through the Scientific Resource Center for the EHC Program.

5.1.4: Include a summary written for the lay public.

The lay summary is intended to improve uptake and use of evidence by the public. The EHC program encourages evidence-based decisions by identifying and engaging partners in using and translating EPC reports for use by their constituents. In many cases, the EHC program prepares unique lay summaries for consumers, clinicians, or policy-makers via the John M. Eisenberg Center for Clinical Decisions and Communications Science as appropriate for each clinical topic.

5.2.2: Provide a public comment period for the report and publicly report on disposition of comments.

This is standard practice in the EHC program, although EPC reviews may be conducted for other programs which have not yet integrated this process.

Recommendations/Actions/Changes to EPC Procedures and Policies

As described above, the IOM recommendations lack clarity with regards to several standards. Ongoing work may help augment and define some of the recommendations and associated terminology. The EPC program has convened a workgroup to define nonfinancial conflicts of interest and suggest frameworks for how they can be used to identify and reduce risk of bias. Outside groups such as the PRISMA-P group (Preferred Reporting Items for SRs and Meta-Analyses—for Protocols) are working on defining the major required protocol elements (and thus the level of detail required in a protocol) and those elements for which a protocol amendment would be required; their efforts may be helpful in further elucidating and clearly defining standards for the field of SR.

The EHC program has updated many of its internal process documents to clarify roles and responsibilities of, and lines of communications between the review authors, those providing solicited input, and the AHRQ officer to more effectively identify and manage conflicts of interest (See Appendix C).

Workgroup on Protocol Elements (“Protocol Workgroup”)

Description of the Differences Between EPC Practice/EPC Methods Guidance and IOM Guidelines

The EHC program maintains a template to define the essential elements in a SR protocol. The 10 protocol elements deemed essential by the IOM committee are all included in the EPC program SR protocol template. However, EPC directors flagged three of the IOM protocol elements (phase 1) that warrant further discussion in relation to their implementation within the EHC program: Standards: 2.6.1: Describe the context and rationale for the review from both a decision-making and research perspective; 2.6.9: Describe the method for evaluating the body of evidence, including the quantitative and qualitative synthesis strategies; and 2.6.10: Describe and justify any planned analyses of differential treatment effects according to patient subgroups, how an intervention is delivered, or how an outcome is measured.

Summary of Deliberative Discussions

2.6.1: Describe the context and rationale for the review from both a decisionmaking and research perspective.

The current EPC guidance related to this item reads, “describe topic background, including information about the topic nomination. Include the clinical context, decisional dilemmas, current relevant practices, or other information to provide context for the SR.” Protocol workgroup members found the wording of the IOM element to be unclear, particularly the phrasing, “from a research perspective.” Rather, the wording should emphasize the rationale for the review, in the context of what is already known on the topic. For instance, if there are other SRs available on the topic, this should be stated.

2.6.9: Describe the method for evaluating the body of evidence, including the quantitative and qualitative synthesis strategies.

Current EPC guidance is in agreement with this IOM element. However, protocol workgroup members found the wording allows for considerable variation in the level of detail and content of

methods to be provided. Specific guidance for describing whether meta-analysis is appropriate in light of included data and planned statistical measures and methods for combining the expected type of data, if appropriate, are lacking. The EPC workgroup agrees with the IOM report (Element 4.4.2) that a description of methods for exploring statistical heterogeneity should be included, as is currently addressed in the EPC guidance.

2.6.10: Describe and justify any planned analyses of differential treatment effects according to patient subgroups, how an intervention is delivered, or how an outcome is measured.

The EPC workgroup agrees that all planned subgroups to explore clinical heterogeneity should be described a priori; EPC's should make every effort to identify subgroups at the protocol stage. This may be done through consultation with content experts and other stakeholders to identify particular subgroups that may react differently due to biologic plausibility. Subgroup selection may be based on previous empirical evidence of heterogeneity of treatment effect based on certain factors. It should be noted that there is no clear "gold standard" approach for identifying all major subgroups for separate analysis. As such, it is possible that potentially meaningful subgroup analyses may arise once the review is underway or once reviewers have had a look at included studies. Such subgroup analyses should be described as post-hoc in the review, as per IOM element 5.1.6, which asks reviewers to describe which analyses were pre-specified. Excessive post-hoc subgroup analysis should be limited to reduce the potential for spurious findings, and moreover, should be explicitly stated as "hypothesis generating" rather than "hypothesis testing" which is the purpose of a priori subgroup analyses.

Recommendations/Actions/Changes to EPC Procedures and Policies

Current EPC protocol guidance addresses all items of the IOM protocol reporting recommendations. The Protocol workgroup suggested clarifying current EPC protocol instructions to "describe the rationale for conducting a SR and what this review would add to the existing evidentiary base, if relevant" and to include a statement on how evidence will be summarized "in a clinically relevant manner".

There are areas where the IOM recommendations do not provide sufficient detail regarding the level of specificity required. Further work is needed to determine the appropriate level of detail required at the protocol level, which methods may be specified at a later date and what level of methods development or change would require a protocol amendment. A set of minimum items that should be reported in protocols—the Preferred Reporting Items for SR and Meta-Analyses Protocols (PRISMA-P) guidelines—are currently being developed, supported in part by AHRQ. The guidance will be accompanied by an explanatory document, providing rationale and evidence for each item; this initiative is intended to improve clarity and transparency of protocols. Ongoing work by a workgroup examining SR methods for considering heterogeneity of treatment effect may provide greater guidance for how and when to select subgroup factors.

Workgroup on Searching, Screening and Reporting Bias ("Searching Workgroup")

Description of the Differences Between EPC Practice/EPC Methods Guidance and IOM Guidelines

EPC methods guidance addresses the use of citation indexes, searching regional databases, searching grey literature, contacting researchers for clarifications, requests of scientific information packets, hand searching, inclusion of non-English studies, quality control of the screening process, pilot testing of screening process, and the inclusion of observational studies. While this guidance is generally consistent with IOM standards, certain elements of the IOM standards include details not specifically addressed. Other elements of the standards were vague making context difficult to assess therefore limiting our ability to sufficiently evaluate consistency with EPC guidance. Standards not directly addressed in EPC guidance included standard 3.1.3 regarding the use of a peer librarian to review search strategies and standard 3.2.5 regarding searching the Web. Despite the lack of formal EPC guidance, EPCs frequently internally peer review electronic search strategies and conduct Web searches.

Summary of Deliberative Discussions

IOM Elements in Which There Is Disagreement

3.1.3: Use an independent librarian or other information specialist to peer review the search strategy.

The Searching workgroup members agreed that peer review was appropriate, but emphasized the lack of evidence linking independent peer review by a librarian or information specialist to changes in the included literature or the conclusions of the review, and thus maintained that peer review by experienced reviewers within the EPC was sufficient.

3.2.5: Conduct a Web search.

The Searching workgroup members agreed that this standard is vague, resulting in a lack of clarity regarding its intent.

IOM Elements Requiring Modification or Clarification

3.1.5: Search citation indexes.

The Searching workgroup members acknowledged that citation searching is not typically conducted as part of the usual search process; however, in some situations it is warranted, such as in the case of landmark trials or studies, or when updating a review. The Searching workgroup members also noted that adding citation searching to usual practice may not be an efficient use of resources. (The IOM report acknowledges that the search elements may be time and resource intensive.) Anecdotal evidence suggests that these searches can be labor intensive and rarely result in additional included studies beyond bibliographic database searching and backward citation searching of previous SRs.

3.2.1: Search grey literature databases, clinical trials registries, and other sources of unpublished information about studies.

The Searching workgroup members discussed certain sources of grey literature recognized as important to most review topics, and recommended that ClinicalTrials.gov, FDA documents, and EPC evidence reports, technology assessments, and comparative effectiveness reviews should always be included in the literature search. Other trial registries were also considered important, but the workgroup suggested that they be included in the search only when considered appropriate by the review team or when discussions with the TEP suggest that additional sources should be included. The Searching workgroup members recommended that conference or meeting abstracts should be searched as an indicator of unpublished studies, and that the decision to include or exclude abstracts be made a priori by the EPCs. Other specialty sources may be searched when the review team considers them to be important sources of information for the specific topic of the review. In all cases it is important to prespecify what will be searched and why.

3.2.2: Invite researchers to clarify information about study eligibility, study characteristics, and risk of bias.

Contacting study authors is suggested in EPC guidance, but not commonly practiced among EPCs. The Searching workgroup members agreed that this practice may be beneficial in some circumstances, such as when clarification regarding study eligibility, study design, or other aspects of the study conduct may affect review conclusions. An example of when contacting an author might be appropriate is the case where disaggregated data may be available, and is needed to evaluate benefits and harms in subpopulations included in the aggregate data. However, it was agreed that decisions about whether, and under which circumstances, to contact researchers should be discussed in the Methods section of the review.

3.2.3: Invite all study sponsors and researchers to submit unpublished data, including unreported outcomes, for possible inclusion in the SR.

The Searching workgroup members recommended continuance of the current practice in which the AHRQ Scientific Resource Center requests Scientific Information Packets on behalf of the EPCs. However, expanding invitations to all study sponsors and researchers may not be feasible.

3.2.4: Hand search selected journals and conference abstracts.

The Searching workgroup members discussed the excess burden of hand searches, and agreed with EPC guidance that hand searches were necessary only when a highly relevant journal not indexed in bibliographic databases was identified. Hand searching conference abstracts to identify ongoing studies can be conducted as deemed appropriate by the review team.

3.3.3: Use two or more members of the review team, working independently, to screen and select studies.

The Searching workgroup members noted that dual screening may improve review quality; however, they agreed that this element should focus on quality control processes that ensure that studies are excluded for consistent and appropriate reasons rather than prescribing dual screening. The Searching workgroup members recommended that EPCs use a systematic approach to ensuring a reliable screening and selection process but recommended against

requiring that it be fully redundant, dual review. The process could include fully redundant dual review, but could also include dual review of a sample of retrieved abstracts and articles. The protocol for the review should specify whether full dual review, sampling, or some other mechanism will be used to ensure the quality and reliability of the screening and selection processes. When sampling is used, a level of agreement considered acceptable should be documented in the protocol.

3.5.1: At a minimum, use two or more researchers, working independently, to extract quantitative and other critical data from each study. For other types of data, one individual could extract the data while the second individual independently checks for accuracy and completeness. Establish a fair procedure for resolving discrepancies—do not simply give final decision-making power to the senior reviewer.

Searching workgroup members agreed that dual abstraction of quantitative data was ideal; however, resource limitations make it infeasible all of the time. In these situations, fact checking by a second independent reviewer may be sufficient. The workgroup recommended that risk of bias assessments be conducted by two independent investigators. Resolution of disagreements in abstracted data by a senior investigator is not sufficient to resolve such disagreements and an acceptable method of resolution should be stated a priori.

IOM Elements in Which There Is Agreement

3.1.9: Search regional bibliographic databases if other databases are unlikely to provide all relevant evidence.

Current EPC guidance recommends searching databases with stronger international coverage of languages of interest when research on a particular topic is conducted primarily outside of the United States. The Searching workgroup members acknowledged that this practice would often be unnecessary, but agreed that certain circumstances might warrant searches of regional databases (e. g., abstracts from certain geographic regions demonstrated results inconsistent with the mainstream medical journals). The Searching workgroup members concluded that the Technical Expert Panel (TEP) could help determine whether searching regional databases might be beneficial for a particular topic.

3.2.6: Search for studies reported in languages other than English if appropriate.

The Searching workgroup members noted that there is currently insufficient evidence that inclusion of foreign language studies changes review conclusions. Given the significant resources necessary to identify and translate non-English studies, the workgroup recommended that such searches only be conducted when considered appropriate in the judgment of the review team, and that reasons for conducting an a priori search of the non-English language literature should be explained in the search protocol. A priori decisions not to include non-English studies may be changed post hoc when there are consistent differences in study results found in English versus non-English abstracts, when most relevant studies are found to be reported in languages other than English, or when most of the relevant studies are found to have been conducted in non-English speaking regions of the world. The Searching workgroup members also recommended against using English-only filters in searching bibliographic databases so these ad hoc decisions can be made based upon appropriate data. All ad hoc decisions should be fully documented in the Methods section of the report.

3.3.4: Train screeners using written documentation; test and retest screeners to improve accuracy and consistency.

EPC guidance, which recommends the use of pilot-testing, is consistent with this standard. The Searching workgroup members recommended that all EPCs should implement methods, including training and testing, for improving the accuracy and consistency of the study selection process. These methods should include feedback mechanisms that identify additional required training or modification to forms used to guide decision-making.

3.3.6: Taking account of the risk of bias, consider using observational studies to address gaps in the evidence from randomized clinical trials on the benefits of interventions.

The Searching workgroup members agreed with this standard and emphasized that observational studies be included when evidence from RCTs does not fully answer review key questions.

Recommendations/Actions/Changes to EPC Procedures and Policies

Workgroup consensus on these standards suggested several necessary actions. First, the Searching workgroup recommended new methods research to establish the impact of several standards. If future methods research suggests that independent peer review of search strategies, citation searching, and foreign language studies (Standards 3.1.3, 3.1.5, 3.1.9) can potentially impact review conclusions, guidance should be updated accordingly.

Secondly, the workgroup recommended EPC guidance be updated in the following cases: to recommend Technical Expert Panel consultation in cases where a significant portion of the literature may only be available in a regional database (Standard 3.1.9); to provide more explicit guidance in identifying and searching for specific types of grey literature, and to describe the value of each grey literature source to specific types of reviews (Standard 3.2.1, 3.2.3, 3.2.5); to suggest hand searching as necessary only in circumstances where an identified highly relevant journal is not indexed in a bibliographic database (3.2.4); to suggest that EPCs should document a priori decisions about whether searching of non-English studies will be conducted and why (3.2.6); to suggest contacting researchers in certain special circumstances (3.2.2); and to address dual screening and abstraction (3.3.3, 3.5.1). With regard to dual abstraction, updated guidance should specifically emphasize that when dual, independent abstraction is not conducted, an independent reviewer should check abstracted data against study full text. Resolution of disagreements in abstracted data by a senior investigator is not sufficient to resolve such disagreements and an acceptable method of resolution should be stated a priori.

Lastly, implementation of the other standards should be enhanced across EPCs. EPCs should train screeners with written documentation and test screeners for accuracy. EPCs should also improve rates of inclusion of observational studies in cases where evidence from RCTs is insufficient to address a key question.

Workgroup on Synthesizing Evidence (“Synthesis Workgroup”)

Description of the Differences Between EPC Practice/EPC Methods Guidance and IOM Guidelines

Guided by the Methods Guide², EPCs synthesize evidence through a three-step process of: (1) evaluating of the risk of bias of each included individual study, (2) evaluating the strength of the body of evidence in relation to each of the prespecified research questions (Key Questions), and (3) determining the applicability of the findings to populations of interest in “real world” setting. Strength of evidence grading requires assessing four required domains (risk of bias, consistency, directness, and precision) and four additional domains if they are considered relevant to the evidence in the review (dose-response association, plausible confounding, strength of association, and publication bias).

Relevant IOM standards to this discussion on synthesizing evidence state that assessing individual studies requires that the planned approach be clearly explained and documented in the final report. Synthesizing the body of evidence should use prespecified, analytic methods and be based on agreed-on concepts of study quality. A meta-analysis is desirable because it can offer valuable insights into the pattern of results across studies. Lastly, assessing the sensitivity of conclusions to choices about outcome metrics and statistical models is good practice.

In phase 1, EPC directors identified five elements from the IOM standards related to synthesizing evidence as described above. Of the five elements, all were either standard practice, had been recently revised to be standard practice or were under consideration by a currently convened methods guidance development group. Several that were standard practice did not differ from the IOM in relation to their goal but the Synthesis workgroup thought that further clarification was needed in implementing the standard.

Summary of Deliberative Discussions

3.6.3: Assess the fidelity of the implementation of interventions.

Synthesis workgroup members agreed in principal with this standard and clarified that fidelity in the implementation of an intervention is evaluated through one or both of the following:

1. Identifying a protocol and determining if the protocol was followed (fidelity of the intervention to the protocol) AND/OR
2. Identifying a well-established standard in the field for implementing an intervention and determining if the study has been correctly implemented, consistent with the standard (fidelity of the intervention to standards).

Synthesis workgroup members recommended that EPC reviews should comment on whether fidelity can be evaluated, even if it is only to say, “not enough information to assess fidelity.” In some cases assessing fidelity may not be applicable because it is impossible to evaluate. These cases include:

1. When there is no standard in the field for the implementation of an intervention. This may be compounded by heterogeneity across protocols/interventions OR
2. When the information that would be needed to determine fidelity is inconsistently defined and/or inadequately described across studies.

In addition, the EPC Methods Guide on Assessing the Risk of Bias⁶ specifically refers to this issue and notes that failure of the intervention to maintain fidelity to the protocol can influence performance bias; it is, therefore, a component assessment of risk of bias. However, the interpretation of fidelity may differ by clinical topic. When interventions implement protocols that have minimal concordance with what can be adopted in practice, this would be considered an issue of applicability.

4.1.1: For each outcome, systematically assess the following characteristics of the body of evidence: risk of bias, consistency, precision, directness, and reporting bias.

The Synthesis workgroup members emphasized that EPCs should assess strength of evidence for “major” outcomes. Outcomes considered important enough to warrant formal grading of the strength of the evidence will depend on the key questions, the clinical or policy context, and the purpose of the report, and will be identified a priori in the protocol. Synthesis workgroup members interpreted the term reporting bias to include both publication bias and selective outcome reporting.

The Synthesis workgroup agrees that reporting bias is a significant concern, but recommends additional work to establish methods for identifying and assessing reporting bias prior to becoming standard practice.

4.1.2: For bodies of evidence that include observational research, also systematically assess the following characteristics for each outcome: dose-response association, plausible confounding that would change the observed effect, and strength of association.

Synthesis workgroup members agreed with this standard in principal, although noted that each of these concerns may not be relevant to all bodies of evidence.

4.1.3: For each outcome specified in the protocol, use consistent language to characterize the level of confidence in the estimates of the effect of an intervention.

The working group agreed with this standard in principal. However, in practice, not all outcomes included in the protocol are graded. Because of the large volume of potential outcomes, those considered the most critical are graded. For those outcomes that are graded, consistent language to characterize the level of confidence in the estimate is used. This is expressed through our four strength of evidence grading categories (high, moderate, low and insufficient), which are standardized across reviews.

4.3.1: Explain why a pooled estimate might be useful to decision makers.

Synthesis workgroup members clarified that EPC guidance is stated in relation to insuring that meta-analysis produces a meaningful and interpretable result. Synthesis workgroup members noted that it is difficult to determine what “might be useful to decision makers.” However, Synthesis workgroup members believed the goals and results were the same.

4.4.1: Use expert methodologists to develop, execute, and peer review the meta-analyses.

While there may be some difference of opinion in the determination of an expert methodologist, the working group agreed with this standard in principal.

4.4.4: Assess the sensitivity of conclusions to changes in the protocol, assumptions, and study selection (sensitivity analysis).

Synthesis workgroup agreed with this standard, noting that current guidance requires a plan for conducting sensitivity analysis, although note that better and more detailed guidance may improve implementation of this standard.

Recommendations/Actions/Changes to EPC Procedures and Policies

The Synthesis workgroup recommended that the update of previous guidance on grading the strength of evidence⁷ include greater clarity to improve consistent implementation among EPCs when assessing the strength of evidence. In particular, greater attention to best methods for assessing reporting bias is needed. Overall, the Synthesis workgroup believes that EPC and IOM standards are in agreement.

Discussion

While the standards and elements of performance laid out by the IOM committee on standards for SR are considered to be current “best practices”, the committee itself described them as “provisional, pending better empirical evidence about their scientific validity, feasibility, efficiency, and ultimate usefulness in healthcare decisionmaking.”¹ This acknowledgement by the committee emphasizes the need for empiric testing of these recommendations. Many of the recommendations were based on theoretical principles and variable levels of empiric evidence without overall evaluation of the balance of benefits and required resources when implementing all recommendations, a task beyond the scope of the committee.

In general, this paper describes the high level of agreement between the EPC program and the IOM recommendations. As previously stated, of the 82 elements of performance within the IOM standards, 34 (41%) were identified for potential disagreement and further discussion. Of these 34, EPC workgroups generally agreed with the majority of these recommendations in principle. However, EPC workgroups identified three elements that they did not recommend routinely for the EPC program. One of these rare disagreements (IOM element 3.1.3: Use an independent librarian or other information specialist to peer review the search strategy) was due to concerns about whether the empiric evidence to support the recommendations outweighed concerns about the feasibility or burden. In this case, the EPC workgroup recommends further evaluation prior to routine implementation of the recommendation. Although the EPC workgroup disagreed with the recommended approach for implementing IOM element 2.7.1 (provide a public comment period for the protocol and publicly report on disposition of comments), they agreed on the underlying principle and suggest an alternative method for implementation. The final IOM element that the EPC workgroup did not recommend for routine adoption (IOM standard 3.2.5: Conduct a Web search) was because of lack of clarity and justification in the IOM standards.

The remaining 31 elements identified through phase 1 of this process were generally in agreement with current EPC guidance. Discussion by the workgroups identified areas where EPCs have tried to implement the recommended elements with variable success. Appendix B shows the directors’ responses to the initial survey, revealing a variation of practice among EPCs on a number of items. This lack of consistency highlights areas where greater clarity and specific instructions would be helpful to improve adherence to generally agreed upon principles. This report acknowledges the need for further research, where possible, in areas of uncertainty to inform best practice. The EPCs will use this evidence to determine appropriate trade-offs between empiric evidence and feasibility (limited resources) for all standards and elements for each particular task.¹

Workgroup discussions focused on cohesion of recommended standards and the effects of implementing all recommended standards. As one example of this, the Program workgroup struggled with how to manage bias and conflicts of interest among those providing input into the review (IOM Standard: 2.4.2). However, this needs to be balanced with the need to ensure appropriate clinical expertise and to ensure opportunities for public input and comment.

This process identified areas for improved practice within the EPC program. It identified standards that are recommended in EPC methods guidance but not routinely practiced and require clarifications and greater efforts at ensuring consistent implementation through dissemination and training. Workgroups also identified standards which are already stated practice across EPCs but where the IOM standards highlight reasonable suggestions which could improve on existing practice and procedures. For example, discussion of Standard 2.8, which

states that the protocol should be publicly available with the inclusion of any amendments, led to a more clearly defined criteria for protocol changes which qualify for an amendment. A revised protocol template inclusive of a table documenting protocol amendments will improve transparency of adherence to this standard.

As a summary of this exercise, across all workgroups, two major categories for further action were identified which would result in improved practice across EPCs. There were areas that workgroups determined would require additional research or development, and there were specific areas of EPC guidance that require updating.

Further research or development needed before adoption:

- Defining nonfinancial conflicts of interest with suggested frameworks for how they can be used to identify and reduce risk of bias
- Defining the major required protocol elements (or level of required detail) and those items requiring a protocol amendment
- Examining SR methods for considering heterogeneity of treatment effect for guiding the selection of subgroup factors (see conclusions of Workgroup 2 recommendations for action)
- Establishing evidentiary basis for Independent peer review, by librarian or information specialist, of search strategies
- Establishing best methods for identifying and assessing reporting bias
- Validating the benefit of searching citation indexes
- Validating that inclusion of non-English studies changes study conclusions

Update EPC guidance to:

- Clarify current EPC protocol instructions mandating that they:
 - describe the rationale for conducting each particular SR,
 - describe its potential contribution to the existing evidentiary base,
 - state how evidence will be summarized in a clinically relevant manner
- Include a minimum set of items to report in protocols (per the in-process PRISMA-P guidelines, see Workgroup 2 recommendations for action.)
- Use TEP to advise in cases where a significant portion of the literature may only be available in a regional database
- Document a priori decisions on the use of non-English studies, while allowing for ad hoc decisions which are fully documented in the Methods section
- Decide when to contact study authors (researchers) and how best to report this practice in the methods section of the review
- Conduct systematic approaches to ensure a reliable screening process when dual screening is not used
- Select an alternative practice (e.g., an independent reviewer checking abstracted data against full text) when dual, independent data abstraction is not done
- Implement strength of evidence assessment consistently, across EPCs, and in particular best methods for assessing reporting bias
- Develop a more detailed plan for conducting sensitivity analysis
- Identify and search for specific types of grey literature and to describe the value of each grey literature source for specific types of reviews
- Explain the circumstances requiring hand searching of journals and conference abstracts

Although not within the intended scope of the workgroup discussions, several workgroups made recommendations for the implementation of some elements. These are included here to inform future discussion:

- Implementation of newly defined standards for reducing risk of bias through nonfinancial conflicts of interest among the study team
- Review of role and responsibilities, and the establishment of a fair procedure for resolving discrepancies among review authors, those providing solicited input, and the AHRQ officer
- Implementation of methods for improving the accuracy/consistency of study selection
- Training and testing of screeners for accuracy
- Clarification in Methods section of circumstances under which researchers should be contacted

Conclusion and Recommendations

This project has engaged the EPCs in a productive process of evaluation and response to the work of the IOM committee on SR standards. This paper describes the EPC experiences in implementing many of the IOM recommendations and the limitations they have experienced regarding lack of clear instructions or other difficulties in consistent implementation across many groups of systematic review investigators. This paper outlines resulting recommendations to conduct research to further validate certain of the IOM standards and elements, to update the EPC guidance where it was found lacking in detail or clarity, to conduct further work where the IOM standards lack sufficient detail, and to improve training of those conducting evidence reviews. This will provide a better framework for determining when the variation of practice among EPCs is a desirable thing, and when standards should be more consistently applied. Principally, this process has, and will, result in improvements in EPC practice, and has identified where further work is needed to provide scientific justification for these practices.

References

1. Institute of Medicine. Finding What Works in Health Care: Standards for Systematic Reviews (Report Brief). 2011. www.iom.edu/Reports/2011/Finding-What-Works-in-Health-Care-Standards-for-Systematic-Reviews/Report-Brief.aspx. Accessed July 2012.
2. Slutsky J, Atkins D, Chang S, et al. AHRQ series paper 1: Comparing medical interventions: AHRQ and the effective health-care program. *J Clin Epidemiol*. 2010 May;63(5):481-3.
3. Institute of Medicine. Finding What Works in Health Care: Standards for Systematic Reviews. 2011. www.iom.edu/Reports/2011/Finding-What-Works-in-Health-Care-Standards-for-Systematic-Reviews/Standards.aspx. Accessed July, 2012.
4. Institute of Medicine Committee on Conflict of Interest in Medical Research. Conflict of interest in medical research, education, and practice. Washington, DC: National Academies Press; 2009.6.
5. “Responsible Prospective Contractors” Title 45 Code of Federal Regulations, Pt. 94 October 1, 2011. U.S. Government Printing Office. www.gpo.gov/fdsys/. Accessed July 2012.
6. Viswanathan M, Ansari MT, Berkman ND, Chang S, Hartling L, McPheeters LM, Santaguida PL, Shamliyan T, Singh K, Tsertsvadze A, Treadwell JR. Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions. Agency for Healthcare Research and Quality Methods Guide for Comparative Effectiveness Reviews. March 2012. AHRQ Publication No. 12-EHC047-EF. <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=998&pageaction=displayproduct>. Accessed November 2012.
7. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: Grading the strength of a body of evidence when comparing medical interventions—Agency for Healthcare Research and Quality and the Effective Health Care Program. *J Clin Epidemiol*. 2010 May;63(5):513-23.

Appendix A. Workgroups

EPC Workgroups on IOM Standards for SR:

Workgroup 1: Program

Leader / TOO: Stephanie Chang

Participants:

- Eric Bass (JHU)
- Stanley Ip (Tufts)
- Melissa McPheeters (Vanderbilt)
- Sydney Newberry (RAND)
- Susan Norris (Oregon)
- Margaret Piper (BCBS)
- Paul Shekelle (RAND)
- Meera Viswanathan (RTI)
- Evelyn Whitlock (Oregon)
- Michael White (UConn)
- Renee Wilson (JHU)

Workgroup 2: Protocol

Leader: David Moher and Stephanie Chang

TOO: Christine Chang

Participants:

- Nancy Berkman (RTI)
- Mark Helfand (Oregon)
- Joseph Lau (Tufts)
- Kathleen Lohr (RTI)
- Melissa McPheeters (Vanderbilt)
- Susan Norris (Oregon)
- Jodi Peters (Ottawa)
- Lisa Sarsany (BCBS)
- Jodi Segal (JHU)
- Larissa Shamseer (Ottawa)
- Paul Shekelle (RAND)
- Tom Trikalinos (Tufts)
- Evelyn Whitlock (Oregon)

Workgroup 3: Searching

Leader: Howard Balshem

TOO: Kim Wittenberg

Participants:

- Naomi Aronson (BCBS)
- Michelle Brasure (MN)
- Donna Dryden (Alberta)

- Bob Kane (Minnesota)
- Rose Relevo (OHSU)
- Gillian Sanders Schmidler (Duke)
- Paul Shekelle (RAND)
- Tom Trikalinos (Tufts)
- Michael White (CT)

Workgroup 4: Synthesis

Leader: Nancy Berkman

TOO: Bill Lawrence

Participants:

- Mohammed Ansari (Ottawa)
- Ethan Balk (Tufts)
- Brittany Burda (Oregon)
- Gerald Gartlehner (RTI)
- Lisa Hartling (Alberta)
- Bob Kane (Minnesota)
- Susan Norris (Oregon)
- Parminder Raina (McMaster)
- James Reston (ECRI)
- Gillian Sanders Schmidler (Duke)
- Sonal Singh (JHU)

Appendix B. EPC Directors' Summary of Responses* to IOM Standards for Systematic Reviews

STANDARDS MARCH 2011

Finding What Works in Health Care Standards for Systematic Reviews

These standards are for systematic reviews of comparative effectiveness research of therapeutic medical or surgical interventions
Standards for Initiating a Systematic Review

Standard	Currently Practiced by EPC With Little/No Variation; No Further Discussion Needed	Common EPC Practice With Some Variations; Some Discussion Needed Before Adoption	Uncommon EPC Practice or With Substantial Variations; Much Discussion and/or More Data Needed Before Adoption
Standards for Initiating a Systematic Review			
STANDARD 2.1 Establish a team with appropriate expertise and experience to conduct the systematic review			
2.1.1 Include expertise in the pertinent clinical content areas	(11) 85%	(2) 15%	
2.1.2 Include expertise in systematic review methods	100%		
2.1.3 Include expertise in searching for relevant evidence	100%		
2.1.4 Include expertise in quantitative methods	(12) 92%	(1) 8%	
2.1.5 Include other expertise as appropriate	(12) 92%	(1) 8%	
STANDARD 2.2 Manage bias and conflict of interest (COI) of the team conducting the systematic review			
2.2.1 Require each team member to disclose potential COI and professional or intellectual bias	92%		8%
2.2.2 Exclude individuals with a clear financial conflict	85%	15%	

2.2.3 Exclude individuals whose professional or intellectual bias would diminish the credibility of the review in the eyes of the intended users [‡]	(8) 61%	(4) 31%	(1) 8%
STANDARD 2.3 Ensure user and stakeholder input as the review is designed and conducted			
2.3.1 Protect the independence of the review team to make the final decisions about the design, analysis, and reporting of the review [‡]	(8) 61%	(4) 31%	(1) 8%
STANDARD 2.4 Manage bias and COI for individuals providing input into the systematic review			
2.4.1 Require individuals to disclose potential COI and professional or intellectual bias	(11) 85%	(2) 15%	
2.4.2 Exclude input from individuals whose COI or bias would diminish the credibility of the review in the eyes of the intended users [‡]	(7) 54%	(5) 38%	(1) 8%
STANDARD 2.5 Formulate the topic for the systematic review			
2.5.1 Confirm the need for a new review [‡]	(10) 77%	(2) 15%	(1) 8%
2.5.2 Develop an analytic framework that clearly lays out the chain of logic that links the health intervention to the outcomes of interest and defines the key clinical questions to be addressed by the systematic review	100%		
2.5.3 Use a standard format to articulate each clinical question of interest [‡]	(8) 62%	(5) 38%	
2.5.4 State the rationale for each clinical question [‡]	10) 77%	(3) 23%	
2.5.5 Refine each question based on user and stakeholder input	(12) 92%	(1) 8%	
STANDARD 2.6 Develop a systematic review protocol			
2.6.1 Describe the context and rationale for the review from both a decisionmaking and research perspective [‡]	(10) 77%	(3) 23%	
2.6.2 Describe the study screening and selection criteria (inclusion/exclusion criteria)	100%		

2.6.3 Describe precisely which outcome measures, time points, interventions, and comparison groups will be addressed	(12) 92%	(1) 8%	
2.6.4 Describe the search strategy for identifying relevant evidence	100%		
2.6.5 Describe the procedures for study selection	100%		
2.6.6 Describe the data extraction strategy	(12) 92%	(1) 8%	
2.6.7 Describe the process for identifying and resolving disagreement between researchers in study selection and data extraction decisions	100%		
2.6.8 Describe the approach to critically appraising individual studies	100%		
2.6.9 Describe the method for evaluating the body of evidence, including the quantitative and qualitative synthesis strategies [‡]	(10) 77%	(3) 23%	
2.6.10 Describe and justify any planned analyses of differential treatment effects according to patient subgroups, how an intervention is delivered, or how an outcome is measured [‡]	(10) 77%	(3) 23%	
2.6.11 Describe the proposed timetable for conducting the review	85%	(2) 15%	
STANDARD 2.7 Submit the protocol for peer review			
2.7.1 Provide a public comment period for the protocol and publicly report on disposition of comments [‡]	(10) 77%	(2) 15%	(1) 8%
STANDARD 2.8 Make the final protocol publicly available, and add any amendments to the protocol in a timely fashion[‡]	(8/11) 73%	(1/11) 9%	(2/11) 18%
Standards for Finding and Assessing Individual Studies			
STANDARD 3.1 Conduct a comprehensive systematic search for evidence			
3.1.1 Work with a librarian or other information specialist trained in performing systematic reviews to plan the search strategy	(11) 85%	(2) 15%	
3.1.2 Design the search strategy to address each key research question	100%		
3.1.3 Use an independent librarian or other information specialist to peer review the search strategy [‡]	(3) 23%	(4) 31%	(6) 46%

3.1.4 Search bibliographic databases	(12) 92%	(1) 8%	
3.1.5 Search citation indexes [‡]	(9) 69%	(3) 23%	(1) 8%
3.1.6 Search literature cited by eligible studies	(11) 85%	(2) 15%	
3.1.7 Update the search at intervals appropriate to the pace of generation of new information for the research question being addressed	(11) 85%	(2) 15%	
3.1.8 Search subject-specific databases if other databases are unlikely to provide all relevant evidence	100%		
3.1.9 Search regional bibliographic databases if other databases are unlikely to provide all relevant evidence ^{†‡}	(5) 38%	(5) 38%	(3) 23%
STANDARD 3.2 Take action to address potentially biased reporting of research results			
3.2.1 Search grey literature databases, clinical trial registries, and other sources of unpublished information about studies [‡]	(8) 61%	(4) 31%	(1) 8%
3.2.2 Invite researchers to clarify information about study eligibility, study characteristics, and risk of bias [‡]	(3) 23%	(4) 31%	(6) 46%
3.2.3 Invite all study sponsors and researchers to submit unpublished data, including unreported outcomes, for possible inclusion in the systematic review ^{†‡}	(2) 15%	(6) 46%	(5) 38%
3.2.4 Hand search selected journals and conference abstracts [‡]	(7) 54%	(2) 15%	(4) 31%
3.2.5 Conduct a Web search [‡]	(4) 31%	(4) 31%	(5) 38%
3.2.6 Search for studies reported in languages other than English if appropriate [‡]	(2) 15%	(7) 54%	(4) 31%
STANDARD 3.3 Screen and select studies			
3.3.1 Include or exclude studies based on the protocol's prespecified criteria	(12) 92%		(1) 8%
3.3.2 Use observational studies in addition to randomized clinical trials to evaluate harms of interventions	(11) 85%	(2) 15%	
3.3.3 Use two or more members of the review team, working independently, to screen and select studies [‡]	(9) 69%	(3) 23%	(1) 8%
3.3.4 Train screeners using written documentation; test and retest screeners to improve accuracy and consistency [‡]	(6) 46%	(6) 46%	(1) 8%

3.3.5 Use one of two strategies to select studies: (1) read all full-text articles identified in the search or (2) screen titles and abstracts of all articles and then read the full text of articles identified in initial screening	100%		
3.3.6 Taking account of the risk of bias, consider using observational studies to address gaps in the evidence from randomized clinical trials on the benefits of interventions [‡]	(8) 62%	(5) 38%	
STANDARD 3.4 Document the search			
3.4.1 Provide a line-by-line description of the search strategy, including the date of every search for each database, web browser, etc. [†]	(11) 85%	(1) 8%	(1) 8%
3.4.2 Document the disposition of each report identified including reasons for their exclusion if appropriate	(12) 92%	(1) 8%	
STANDARD 3.5 Manage data collection			
3.5.1 At a minimum, use two or more researchers, working independently, to extract quantitative and other critical data from each study. For other types of data, one individual could extract the data while the second individual independently checks for accuracy and completeness. Establish a fair procedure for resolving discrepancies—do not simply give final decisionmaking power to the senior reviewer [‡]	(3) 23%	(6) 46%	(4) 31%
3.5.2 Link publications from the same study to avoid including data from the same study more than once	100%		
3.5.3 Use standard data extraction forms developed for the specific systematic review	100%		
3.5.4 Pilot-test the data extraction forms and process	(12) 92%	(1) 8%	
STANDARD 3.6 Critically appraise each study			
3.6.1 Systematically assess the risk of bias, using predefined criteria	100%		
3.6.2 Assess the relevance of the study's populations, interventions, and outcome measures	(11) 85%	(2) 15%	
3.6.3 Assess the fidelity of the implementation of interventions [‡]	(6) 46%	(3) 23%	(4) 31%

Standards for Synthesizing the Body of Evidence

NOTE: The order of the standards does not indicate the sequence in which they are carried out.

STANDARD 4.1 Use a prespecified method to evaluate the body of evidence

4.1.1 For each outcome, systematically assess the following characteristics of the body of evidence: [‡] <ul style="list-style-type: none"> • Risk of bias • Consistency • Precision • Directness • Reporting bias 	(9) 69%	(4) 31%	
4.1.2 For bodies of evidence that include observational research, also systematically assess the following characteristics for each outcome: [‡] <ul style="list-style-type: none"> • Dose-response association • Plausible confounding that would change the observed effect • Strength of association 	(6) 46%	(7) 54%	
4.1.3 For each outcome specified in the protocol, use consistent language to characterize the level of confidence in the estimates of the effect of an intervention [‡]	(9) 69%	(3) 23%	(1) 8%

STANDARD 4.2 Conduct a qualitative synthesis

4.2.1 Describe the clinical and methodological characteristics of the included studies, including their size, inclusion or exclusion of important subgroups, timeliness, and other relevant factors	(11) 85%	(2) 15%	
4.2.2 Describe the strengths and limitations of individual studies and patterns across studies	(11) 85%	(2) 15%	
4.2.3 Describe, in plain terms, how flaws in the design or execution of the study (or groups of studies) could bias the results, explaining the reasoning behind these judgments	(11) 85%	(2) 15%	
4.2.4 Describe the relationships between the characteristics of the individual studies and their reported findings and patterns across studies	(11) 85%	(2) 15%	
4.2.5 Discuss the relevance of individual studies to the populations, comparisons, cointerventions, settings, and outcomes or measures of interest	(12) 92%	(1) 8%	

STANDARD 4.3 Decide if, in addition to a qualitative analysis, the systematic review will include a quantitative analysis (meta-analysis)			
4.3.1 Explain why a pooled estimate might be useful to decision makers ^{†‡}	(6) 46%	(5) 38%	(2) 15%
STANDARD 4.4 If conducting a meta-analysis, then do the following:			
4.4.1 Use expert methodologists to develop, execute, and peer review the meta-analyses [†]	(10) 77%	(3) 23%	
4.4.2 Address the heterogeneity among study effects	100%		
4.4.3 Accompany all estimates with measures of statistical uncertainty	100%		
4.4.4 Assess the sensitivity of conclusions to changes in the protocol, assumptions, and study selection (sensitivity analysis) [†]	(9) 69%	(4) 31%	
Standards for Reporting Systematic Reviews			
STANDARD 5.1 Prepare final report using a structured format			
5.1.1 Include a report title	100%		
5.1.2 Include an abstract	100%		
5.1.3 Include an executive summary	(12) 92%		(1) 8% (not required)
5.1.4 Include a summary written for the lay public ^{†‡}	(6) 46%	(2) 15%	(5) 38%
5.1.5 Include an introduction (rationale and objectives)	100%		
5.1.6 Include a methods section. Describe the following: <ul style="list-style-type: none"> • Research protocol • Eligibility criteria (criteria for including and excluding studies in the systematic review) • Analytic framework and key questions • Databases and other information sources used to identify relevant studies • Search strategy • Study selection process • Data extraction process • Methods for handling missing information • Information to be extracted from included studies • Methods to appraise the quality of individual studies • Summary measures of effect size (e.g., risk ratio, difference in means) • Rationale for pooling (or not pooling) results of included studies • Methods of synthesizing the evidence (qualitative and meta-analysis) • Additional analyses, if done, indicating which were prespecified 	100%		

5.1.7 Include a results section. Organize the presentation of results around key questions. Describe the following (repeat for each key question): <ul style="list-style-type: none"> • Study selection process • List of excluded studies and reasons for their exclusion • Appraisal of individual studies' quality • Qualitative synthesis • Meta-analysis of results, if performed (explain rationale for doing one) • Additional analyses, if done, indicating which were prespecified • Tables and figures 	(12) 92%	(1) 8%	
5.1.8 Include a discussion section. Include the following: <ul style="list-style-type: none"> • Summary of the evidence • Strengths and limitations of the systematic review • Conclusions for each Key Question • Gaps in evidence • Future research needs 	100%		
5.1.9 Include a section describing funding sources and COI	(10/12) 83%	(2/12) 17%	
STANDARD 5.2 Peer review the draft report			
5.2.1 Use a third party to manage the peer review process [‡]	(8) 62%	(5) 38%	
5.2.2 Provide a public comment period for the report and publicly report on disposition of comments [‡]	(10) 77%	(2) 15%	(1) 8%
STANDARD 5.3 Publish the final report in a manner that ensures free public access[†]	(11) 85%	(1) 8%	(1) 8%

* Includes responses from 13 out of 14 possible EPC directors in all cases except where marked.

[†] In these instances, numbers are rounded to the nearest percentage point, making it slightly greater or less than 100.

[‡] <80 percent report current practice with little/no variation; standards of greatest controversy

Appendix C. Summary of Decisions and Protocol for Resolution in the Event of Difference of Opinion Between EPC Report Authors and AHRQ Officer

This summary has been incorporated into the EPC procedure guide.

1. EPCs are the authors of the report and should not write anything that they do not agree with. They should respond to all comments from the AE, TOO, reviewers, or AHRQ and if changes were not made, justify why.
2. Justifiable reasons for nonacceptance of a report by the AHRQ officer would include inconsistencies within a report, nonadherence to accepted Methods Guidance, inadequate justification or response to comments.
3. A common practice by AHRQ officers is to repeat comments that they feel particularly strongly about 2 or 3 times. To save time in the back-and-forth, the AHRQ officer should stratify comments by those that are critical to be changed from those that are helpful suggestions. Likewise, to save time, the EPC should provide a clear justification for why any changes are not made and perhaps schedule a call to discuss if necessary.
4. If there is continued disagreement about how to address the comments of the AE or TOO, the EPC should try to resolve the disagreement or uncertainty in a conference call with the AE and TOO.
5. If disagreement cannot be resolved with a conference call, the EPC should prepare a letter explaining the disagreement and send to the AE and TOO, who will review the letter before the TOO forwards it to the EPC Program Director.
6. The EPC Program may then request input from an unconflicted external content, statistical, or methods expert prior to making a determination. If no resolution can be made with third-party review, AHRQ will not prohibit publication and the EPC may publish with the appropriate disclaimer that is contained in Section H.1.b.2.B of the EPC contract.

If there is significant disagreement on the revisions requested, the EPC should schedule a comments call with the TOO and AE to discuss the requested changes within 1 week of receipt of the decision letter. If disagreement cannot be resolved with a conference call, the EPC should prepare a letter explaining the disagreement and send to the AE and TOO, who will review the letter before the TOO forwards it to the EPC Program Director. The EPC Program may then request input from an unconflicted external content, statistical, or methods expert prior to making a determination.