Evidence-based decision making has become a benchmark of best practice. Sources of evidence are systematic reviews and meta-analyses. To support their decision making, health care managers and supervisors need to be able to critically read and interpret systematic reviews and meta-analyses. They also need to be able to determine the applicability of the evidence to their problems and settings. This 2-part series of articles aims to equip health care managers and supervisors with these skills. This article, part 1, explains the types of systematic reviews, defines key terms, and outlines the process of systematic reviews and meta-analyses. Part 2 focuses on the additional procedures associated with meta-analyses, describes the potential shortcomings of both systematic reviews and meta-analyses, and finally, provides a way to appraise the applicability of their results. Key words: evidence-based practice, health services administration, meta-analysis, systematic review
and synthesizing evidence is often difficult for decision makers. Reasons for this difficulty include the following:

- Overwhelming volumes of data exist.
- Data are scattered through many individual studies.
- Individual studies may be methodologically flawed, producing flawed data.
- Individual studies may be unique to a time and setting (not generalizable).
- Data may be misinterpreted or misrepresented.
- Conclusion validity may be lacking.
- Conclusions may conflict across individual studies.

By integrating and synthesizing information, systematic reviews assist decision makers in determining which results are reliable. Moreover, the aggregated result of multiple studies provides decision makers with a more precise estimate of an intervention’s potential effects than a single study does. Decision makers, therefore, can have confidence in the information upon which they base practice and policy.

OVERVIEW

A systematic review is a comprehensive review of the evidence on a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant published and unpublished research studies; to extract and analyze data from the studies that are included in the review; and to present integrated and synthesized information. Avoidance of bias is essential to systematic reviews. Moreover, as a scientific procedure, systematic reviews have a specialized terminology (Table 1). This scientific procedure results in an assessment of the strength of the body of evidence pertinent to a question or problem. Thus, systematic reviews make information available to decision makers in a usable format.

COMPARISON WITH TRADITIONAL NARRATIVE LITERATURE REVIEWS

A traditional narrative literature review is a critical appraisal, conducted by summarizing, classifying, and comparing the published knowledge available about a topic from sources such as research and theoretical articles, annual reviews, books, and dissertations. Traditional narrative literature reviews are being replaced by systematic reviews because they are susceptible to bias. Traditional narrative literature reviews do have purposes in today’s health sector, and a need for them still exists. First, they can trace the history of policies, principles, theories, and other concepts. Second, they can support practitioners’ continuing education by keeping them abreast of current trends. Finally, though, traditional literature reviews have less of a role in a field’s body of knowledge because they are less able to support evidence-based decision making than systematic reviews are.

Systematic reviews differ from traditional narrative literature reviews. Systematic reviews must meet strict standards to avoid bias and to assess the quality of the articles. Therefore, systematic reviews establish whether scientific findings are consistent and, therefore, can be generalized across populations, settings, and variations in interventions or exposures. Consequently, systematic reviews are becoming prevalent. Table 2 provides a comparison of traditional narrative literature reviews and systematic reviews.

TYPES OF SYSTEMATIC REVIEWS

No standardized term and definition have yet been established for systematic reviews. Authors inconsistently and loosely use several terms, such as systematic review, research synthesis, structured review, meta-analysis, and overview. Occasionally, a traditional narrative review is wrongly labeled a systematic review when it does not meet the strict standards for conducting systematic reviews.

Four types of systematic reviews exist based on purpose and on analytical techniques:

- Scoping systematic reviews are exploratory.
- They are also known as scoping reviews, scoping studies, scoping literature reviews, and scoping exercises. They provide a frame of reference for a topic by outlining the range, extent, and breadth of relevant evidence, literature, and other information sources that are available. Typically, scoping studies do not examine the depth
Table 1. Definitions of Key Terms

| Key Term                        | Definition                                                                                                                                 |
|--------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| **Applicability**              | Extent to which the evidence is likely to be applicable to populations, interventions, and settings of interest to the user\(^5\)          |
| **Assessment of risk of bias** | Assessing the risk that the results of individual studies included in a systematic review reflect bias in design or execution in addition to the true effect of the intervention or exposure. Sometimes equated to internal validity. |
| **Backward tracking**          | Process of manually reviewing the reference lists of the originally identified articles to find additional articles and studies\(^6\)       |
| **Clinical diversity**         | Variability among populations' characteristics, interventions, and outcomes of studies\(^7\)                                            |
| **Cochran Q test**             | Classical measure of heterogeneity, which is calculated as the weighted sum of squared differences between effects of individual studies and the pooled effect across studies, with the weights being those used in the pooling of the studies (also known as \(\chi^2\), chi\(^2\), chi-square)\(^8\) |
| **Content analysis**           | Categorization of textual data, in a review’s evidence and sample of articles, into themes and potentially with tabulation of the categories to identify dominant findings and to make generalizations\(^9\) |
| **Cumulative meta-analysis**   | Specialized type of meta-analysis that adds new data to the pool as new studies are published. Cumulative meta-analyses show trends and evolution. Can be performed retrospectively with sequential addition of new data in chronological order of studies’ publication |
| **Directness**                 | Evidence links the interventions directly to health outcomes rather than indirectly through intermediate or surrogate outcomes. \(^5\)         |
| **Effect size (estimate of effect, treatment effect)** | Overall outcome of a meta-analysis achieved by standardizing the individual effect sizes across studies, thus eliminating units and thereby allowing the combination of the studies’ effect sizes. Examples include weighted mean differences, standardized mean differences, odds ratios (ORs), relative risks, risk differences (RDs), Glass $\Delta$, Cohen $d$, Hedge $g$, and the Pearson $r$ correlation coefficient. |
| **Eligibility criteria**       | Explicit standards, established before a systematic review, for inclusion or for exclusion of original studies in the review                  |
| **Forward tracking**           | Process of using citations of the originally identified articles to find additional subsequent articles that cite the originally identified article\(^6\) |
| **Global filter**              | Screening mechanism that uses data about documents to limit results of the search, such as time period or the representation of 1 research study to 1 article per year\(^10\) |
| **Grey (gray) literature**     | Body of publications produced by government, academia, or business and industry, in print or electronic forms or both, which is not published in easily accessible journals and may not appear in databases or through Web searches |
| **Hand-searching**             | Process of scanning the contents of journals, conference proceedings, and abstracts, page by page, to identify relevant publications.       |
| **Heterogeneity**              | Variations among original studies in terms of study designs, populations, interventions, settings, date, and other variables              |
| **I\(^2\) index**              | Measure that quantifies the extent of heterogeneity by describing the percentage of the variability in effect estimates due to heterogeneity rather than chance. A value greater than 50% may be considered to represent substantial heterogeneity.\(^11\) |
| **Individual patient data meta-analysis** | Specialized type of meta-analysis that pools individual patient data (rather than summary data) from multiple original studies and thereby allows analyses of data from subgroups that may not have been possible in the original studies because the subgroups were too small |
| **Meta-analysis**              | Quantitative specialized type of systematic review that introduces statistical techniques to combine the summary-level information from of primary studies (sometimes also known as quantitative systematic review or overview) |
| **Methodological diversity**   | Variability in study design, conduct, and quality, such as blinding and concealment of allocation\(^12\)                                |
| **PICOS**                      | Model to frame research questions, participants (patients, population), intervention (exposure), comparison, outcome, study design (alternately PICOC, substituting context for study design) |
| **Power**                      | Likelihood of detecting, within a sample, that an effect exists within the population; probability that a null hypothesis will be rejected when it is truly false (and the (continues))
of evidence. By mapping a topic, scoping reviews assist researchers in refining their research questions and in selecting appropriate research designs and methods. Scoping reviews support the determination of whether additional research, such as an experimental study or a full systematic review, is warranted.

- Rapid reviews are simplified versions of a systematic review. Rapid reviews are also known as ultra-rapid review, rapid literature reviews, rapid health technology

### Table 1. Definitions of Key Terms, Continued

| Key Term                    | Definition                                                                                                                                 |
|-----------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| Conclusion made that the phenomenon exists when it truly does; increasing power decreases the likelihood of type 2 errors |
| Degree of certainty surrounding the estimate of an outcome’s effect; includes sample size, number of studies, and heterogeneity within or across studies |
| Tendency for selective submission and acceptance of studies with statistically significant (positive) results to be published and with nonsignificant (negative, supporting the null) or inconclusive results not to be published |
| Simplified version of a systematic review that generally conceptualizes questions and succinctly summarizes broad scopes of scientific evidence |
| Procedure to document the search and its comprehensiveness by using several articles to identify and record related Medical Subject Headings terms, key terms, synonyms, analogies and similar terms, conceptual terms, vocabulary, and jargon used by various authors |
| See assessment of risk of bias |
| Exploratory study that outlines the range or scope of evidence or literature that is available on a topic |
| Screening mechanism that removes irrelevant articles from the search by excluding articles that share terms or abbreviations having different meanings in different topics, such as COB for “chronic obstructive bronchitis” and COB for “coordination of benefits” |
| Variability in observed treatment effects across studies resulting from clinical and methodological diversity |
| Overall, cumulative quality of the information from studies included in a systematic review taking into account their individual quality (risk of bias) but emphasizing the quality of their cumulative evidence in terms of precision, directness, and applicability |
| Analysis of subsets (subgroups) of studies in a meta-analysis to estimate the effect of the intervention on the subgroups of interest and to make formal statistical comparisons across the subgroups |
| Effect measure for each outcome, such as likelihood ratios, risk ratios, ORs, RDs, sensitivity and specificity, difference in means, and hazard ratios |
| Comprehensive review of the evidence on a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant published and unpublished research studies; to extract and analyze data from the studies that are included in the review; and to present integrated and synthesized information (sometimes also known as qualitative systematic review or narrative systematic review) |
| Identification of recurrent and important themes, trends, patterns, and other issues across evidence and a review’s sample of articles |
| Procedure to document the search and its comprehensiveness by recording and organizing, for each database, all search (key) terms, their synonyms, and their similar words in a table |
| Critical appraisal, conducted by summarizing, classifying, and comparing the published knowledge available about a topic from sources such as research and theoretical articles, annual reviews, books, and dissertations |
| Method used to combine measures on continuous scales where the mean, standard deviation, and sample size in each group are known; the weight given to the difference in means from each study is determined by the precision of its estimate of effect; this weight equals the influence that the individual study has on the overall estimate of effect size of the meta-analysis (also called mean difference) |

Abbreviations: ORs, odds ratios; RDs, risk differences.
assessments, rapid response, rapid evidence assessments, rapid assessments, realist reviews, evidence summaries, and accelerated systematic reviews. Unlike the tightly focused systematic reviews, rapid reviews generally conceptualize questions and succinctly summarize broad scopes of scientific evidence.27 Rapid reviews are conducted when practitioners, decision makers, or policymakers need information quickly, typically within 1 to 6 months.28 In rapid reviews, researchers and analysts shorten the length of time needed to conduct the review by simplifying the methods of systematic reviews. However, despite the simplified methods, the overall conclusions in rapid reviews are similar to the conclusions of systematic reviews.28

• Systematic reviews are syntheses of evidence on a topic. Systematic reviews are characterized by their comprehensiveness and transparency. Additional characteristics of systematic reviews include the following:
  • Laser focus on a well-defined question for which appropriate study designs can be identified.
  • Provision of answers to questions from a relatively narrow range of homogeneous studies whose quality has been assessed.
  • Systematic reviews can take from 6 months to 24 months to complete.27,28
  • Systematic reviews are sometimes called qualitative systematic reviews or narrative systematic reviews because researchers use qualitative techniques to analyze their findings. It should be emphasized, however, that qualitative (narrative) systematic reviews are not the same as traditional, narrative literature reviews, although their name is very similar.

• Meta-analyses are a specialized type of systematic review that introduces statistical techniques to combine summary-level information from at least 2 studies. In 1976, Gene Glass coined the term *meta-analysis.*29 Meta-analysis is “the statistical analysis of a large collection of results from individual studies for the purpose of integrating findings.”30(p3) Sometimes, a meta-analysis is called an overview or a quantitative systematic review.

To conduct a meta-analysis, researchers must have a highly focused research question. Meta-analyses estimate the overall, combined effect of several studies’ outcomes (effects) for an intervention or exposure. Specifically, this estimate of effect (also called treatment effect) is the observed relationship between an intervention (or exposure) and an outcome expressed as, for example, an odds ratio, a relative risk, a risk difference, a weighted mean difference, a standardized mean difference, Pearson $r$ correlation coefficient, Glass $\Delta$, Cohen $d$, or Hedge $g$.17,21 Thus, meta-analysis is an integrative analysis of all available evidence.

Studies that are included in a meta-analysis have common underlying characteristics.29

The studies do not need to have identical

### Table 2. Comparison of Characteristics: Traditional Narrative Literature Review Versus Systematic Review

| Traditional Narrative Literature Review | Systematic Review |
|----------------------------------------|-------------------|
| Opaqueness                             | Transparency      |
| Lack of methods section with review conducted using informal methods | Methods section with preestablished methods and criteria |
| Subjectivity and idiosyncrasy          | Application of scientific procedures |
| Irreproducibility and disposition toward biased conclusions | Credibility, reproducibility, defendability, and accountability of conclusions |
| Ability to trace history of topics     | Extent of consensus in the evidence |
| Rare updates                           | Periodic updates in the case of electronic reviews, such as the Cochrane Reviews |
| Support for practitioners’ continuing education | Support for evidence-based decision making and policy making |
methodologies; however, the individual studies must be comparable. For example, they share a single intervention or a single outcome (dependent variable). Thus, they are sufficiently similar that the pooled data arise from reasonably homogenous study groups.31

Advantages are derived from the quantitative procedures of meta-analyses. First, in meta-analyses, findings from many studies are combined, some of which are contradictory. These studies examine the same question; however, individually, they are either too small or give conflicting results. Second, researchers weight the studies before combining them. Studies that provide more information, such as having greater precision, are assigned greater weight than less precise studies are.32 Thus, well-conducted studies have greater influence on the meta-analysis' results than poorly conducted studies do. Third, power and precision are increased.33 Fourth, a special type of meta-analysis, individual patient data meta-analysis, pools individual patient data (rather than summary data) from multiple original studies.35 Individual patient data meta-analysis allows analyses of data from subgroups that may not have been possible in the original studies because the subgroups were too small. Finally, cumulative meta-analyses add new data to the pool as new studies are published.34 Cumulative meta-analyses show trends.

It should be noted that the Cochrane Collaboration, a leading organization in evidence-based health care and systematic reviews, does not categorize a meta-analysis as a type of review.14 The Cochrane Collaboration states that a meta-analysis is only the statistical techniques that combine multiple studies’ results into a single estimate of effect. The Collaboration elaborates stating that the final estimate of effect may not always be the result of a systematic review of the literature.

PROCESS

A systematic review is a methodical approach to the body of published and unpublished evidence that reduces the possibility of bias. The method must be so clear, detailed, and exact that another person can replicate it. The process is characterized by the following:

- Exhaustiveness of the search for evidence.
- Transparency of the author’s process of selecting evidence. The method of the search for evidence (literature and other data) is made explicit, including sources of evidence, search terms, and criteria to identify the evidence.
- Evaluation of the quality of the evidence against consistent, specified methodological standards.
- Transparent explanation of how evidence, data, and information from studies, other published works, and other sources were included in or excluded from the review.

Narrative systematic reviews and meta-analyses follow the same general process, with extra procedures being added for the meta-analysis. The steps of conducting narrative systematic reviews are presented in this first article of the 2-part series. The extra procedures required for meta-analyses are presented in the second article of the 2-part series. It should be noted that the following steps incorporate the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses.35

**Narrative systematic review**

**Step 1**

Provide an introduction that includes both the context of the systematic review and an explicit statement of its question, aim, or objective. The context justifies the review in terms of current knowledge on the topic and existing gaps in that knowledge. The introduction concludes with the question, aim, or objective that has been established a priori (before the review begins). For scoping or rapid reviews, the question, aim, or objective may be broad. For systematic reviews and meta-analyses, the research question, aim, or objective must be narrowly focused and stated with reference to participants (patients, populations), interventions (exposures), comparisons, outcomes, and study design (PICOS). Finally, and most importantly,
the research question, aim, or objective must be stated in a way that is answerable.

**Step 2**

Delineate the method of the review. The method explicitly addresses eligibility, selection of information resources, the specifics of the search’s strategy, data extraction, details of the analytic techniques, and the assessment of potential biases.

- Predetermine eligibility criteria of both inclusion criteria and exclusion criteria for the selection of articles. Categories of eligibility criteria include types of research design, studies’ characteristics, and type of publication.
- Query and search information resources.
  - Reference databases include MEDLINE, EMBASE, PsychINFO, and many others. Searching multiple databases is important because the databases contain different sets of journals. For example, in 2012, 1800 journals referenced in MEDLINE were not referenced in EMBASE and vice versa.36
  - Forward tracking and backward tracking use citations from the originally identified articles to find additional articles and relevant information resources.16
  - Hand-searching the contents of journals, conference proceedings, and abstracts, page by page, to identify very recent publications that are not yet indexed in electronic reference databases and publications in nonindexed journals.13,37
  - The grey (gray) literature is the body of publications produced by government, academia, or business and industry, in print or electronic form or both, that is not published in easily accessible journals and may not appear in databases or through Web searches.38 Tools to find the grey literature include the National Technical Reports Library of the US Department of Commerce,39 the Health Services Research Projects in Progress of the US National Library of Medicine,40 the Canadian search tool Grey Matters,41 and others.42-44
  - Contacts are made with authors and researchers in the topic, requesting whether they have unsubmitted manuscripts or suggestions for additional articles or studies.
  - Google Scholar and other Internet search engines, such as Google and Bing, are queried.
- Present the search strategy with the search terms; the limits or filters; the search’s start and end dates for each database; the details of 1 database’s full electronic search strategy, such that it could be replicated; and the method of selecting the articles for inclusion in and exclusion from the review.
  - To ensure the comprehensiveness of the search, the list of search terms (also known as key terms in some reference databases) must be as complete as possible. Two procedures that contribute to the documentation of the search and its comprehensiveness are thethesauri-based expansion and the report-based expansion.20
  - Also reported are any limitations in addition to the exclusion criteria, such as language and date restrictions. Two filtering strategies that can assist researchers are the global filter and the semantic filter.20
  - The start and end dates for the search of each database are provided. These dates allow readers to assess the currency of the review and to take into account the time lag in publication.45
  - The full electronic search strategy of 1 major reference database is reported (Figure 1).
  - The process for selecting studies for inclusion is described. The process includes how the articles were screened, such as title and abstract; who conducted the screening, such as authors or research assistants; independence of the review; interrater reliability; and how disagreements were resolved, such as consensus or the principal investigator’s decision.
- Describe the process of extracting the data. Data extraction includes selecting and
piloting extraction forms, compiling lists of all the variables and outcomes with operational definitions, extracting the data, and finally, describing any procedures to ensure the quality of the data.

- Forms can be researcher designed or based on a published template. Any piloting of forms is noted.
- Researchers list and operationally define all the variables and outcomes, such as PICOS, that they originally sought. They also list and justify any variables or outcomes that they later added to the list.
- Researchers develop a protocol to code and extract the data that they will present, critique, and summarize in the review. The coded data usually fall into 4 basic categories: (a) methodological and substantive features (process and content), (b) study quality, (c) intervention descriptors, and (d) outcome measures. Researchers extract the data.
- Any procedures taken to prevent bias or errors in data extraction are explained. For example, procedures could be that 1, 2, or more researchers extracted data independently and that interextractor reliability was checked. Another procedure noted is the way that the researchers resolved their disagreements, such as consensus obtained through discussion or consulting a third researcher.
- Any contacts or attempted contacts with the articles' authors to obtain or clarify data are also described.
- Researchers describe any procedures that they took to identify multiplicative articles on the same study because, often, studies are published multiple times. This precaution must be taken to avoid double (triple, etc) counting.
- Provide details of the planned analytical techniques. Researchers state the summary statistic (measure) that they intend to use for each outcome, such as likelihood ratios, risk ratios, odds ratios, risk differences, sensitivity and specificity, difference in means, and hazard ratios. Researchers also explain any analytical techniques to examine data from subgroups, to assess heterogeneity, and to determine whether a meta-analysis is feasible. To perform the statistical techniques of meta-analysis, the outcomes of a set of studies need to be reasonably homogeneous (similar). Therefore, meta-analysts conduct the Cochran Q test and the $I^2$ index to check for statistical heterogeneity (homogeneity). Potential causes of statistical heterogeneity are clinical and methodological diversity, bias, and mere chance. Should the researchers conclude, after interpreting the Cochran Q...
test and the $I^2$ index, that the review’s set of studies is not reasonably homogeneous and, therefore, that a meta-analysis is inadvisable, the researchers have analytic methods for the resulting qualitative systematic review. These methods deal with qualitative evidence and with the combination of quantitative and qualitative evidence.\textsuperscript{19} This qualitative synthesis includes thematic analysis and content analysis.\textsuperscript{19}

- Describe the strategies to assess the quality of the individual studies included in the review and to appraise the overall strength of the body of evidence (“strength of evidence”) across the studies. It should be emphasized that the terminology for quality varies across appraisal systems and throughout the literature on systematic reviews and meta-analyses.\textsuperscript{49} Experts at the Agency for Healthcare Research and Quality prefer the term assessment of risk of bias or risk of bias when referring to the appraisal of the quality of the individual studies in the systematic review.\textsuperscript{49} Specifically, “assessing the risk of bias of a study can be thought of as assessing the risk that the study results reflect bias in study design or execution in addition to the true effect of the intervention or exposure under study.”\textsuperscript{49(p3)} This term is preferred because flaws in the design of a study or its implementation, such as a sample unrepresentative of the target population, result in bias. Importantly, bias undermines the overall strength of the evidence.

- An overall rating of risk of bias (quality) is made, such as low, medium, high, or unclear risk of bias.\textsuperscript{49}

- Instruments to assess the risk of bias in individual studies include the Cochrane Collaboration’s tool\textsuperscript{50} and checklists, such as the Quality Assessment of Diagnostic Accuracy Studies\textsuperscript{51} and the Standards for Reporting of Diagnostic Accuracy.\textsuperscript{52}

- Studies are weighted by their assessed quality (risk of bias). Higher weights are given to studies with greater sample sizes and narrower confidence intervals (certainty is linked to perceived quality).\textsuperscript{53}

Weighting can also be based on subjective ratings, such as low, medium, high, or unclear risk of bias. The appraisal of the strength of the body of evidence (“strength of evidence”) takes into account the risk of bias in the individual studies but focuses on the overall, cumulative quality of the information across all the studies (“interstudy bias”). To assess the strength of the evidence, researchers appraise the studies’ evidence overall in terms of its precision, directness, and applicability.\textsuperscript{49} The researchers specify how they will assess phenomena that may affect the strength of the cumulative evidence, such as publication bias and selective reporting of data and outcomes.

- The strength of the evidence is graded for each of the major outcomes of interest, such as PICO\textsubscript{S}. In addition, the strength of evidence is graded for comparisons of special interest to clinicians, decision makers, and policymakers.

- Three commonly used systems to grade the strength of the evidence are the US Preventive Services Task Force system\textsuperscript{54}; the Strength of Evidence system used by the Evidence-Based Practice Centers, supported by the Agency for Healthcare Research and Quality; and the Grading of Recommendations Assessment, Development, and Evaluation system.\textsuperscript{55}

### Step 3

Present the results of the review. The results section documents the results of the search strategy, the characteristics of each of the studies included in the review, the analyses and syntheses, and the assessments of the risk of bias and the strength of the evidence.

- Present the results of the search at each stage of the multistage filtering process. A flow diagram (Figures 2 and 3) shows the number of studies that the search initially identified (by information source), the number of studies screened, the number of studies assessed for eligibility, and the number of studies included in the review. At each stage, the reasons for the exclusion...
Summarize the characteristics of each of the retained studies included in the review in the review's text and in a summary table.

Figure 2. Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) template of a flow diagram. Source: Moher et al.35 For more information, visit www.prisma-statement.org. The PRISMA Statement and the PRISMA Explanation and Elaboration document are distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Figure 3. Example of completed flow diagram of search results. Source: Butler et al.56 (This document is in the public domain and may be used and reprinted without permission except those copyrighted materials noted for which further reproduction is prohibited without the specific permission of copyright holders.)
The summary table includes pertinent information such as the citation, subjects, study size, setting, duration, intervention or exposure, major outcomes, and other key information. The summary table standardizes the results to clearly show missing or unclear information.45

- Present the results of the analyses and syntheses. Summary data, such as means, standard deviations, and odd ratios, from each article are provided in narrative and in graphics. Subgroup analysis on particular populations may be performed and reported. Qualitative synthesis is performed and patterns are reported. The reasons precluding meta-analysis are explained. Finally, the synthesized major findings of the review are presented.
- Report the outcome of the assessment of the risk of bias for each retained study. Potential biases may be design or methodological flaws related to the selection of subjects, the blinding of subjects and researchers, attrition and other sources of incomplete data, and researchers’ selective reporting of outcomes. Finally, an assessment of the aggregate result of the individual studies' biases on the strength of the evidence for the review’s major findings is stated.

**Step 4**

Provide a discussion that summarizes the evidence, addresses limitations, and presents a conclusion.

- Present a summary of the evidence, including the evidence’s strength, for each of the main outcomes. The outcomes should be framed in terms of their relevance for various target audiences, such as policy makers, decision makers, providers, payers, and other stakeholders. For example, many target audiences desire information on the outcomes' impact on individuals' and populations' disease and health; on quality of life; on costs across the health care delivery system; and on beneficial, harmful, and unintended consequences.57 Moreover, the paucity or absence of data for some outcomes may be relevant to policymakers and future researchers and, therefore, should be mentioned.45 Finally, the generalizability of the major outcomes across populations or subgroups, settings, and geographic areas is also important.
- Discuss the review’s limitations in terms of the limitations both of the studies and of the process of the review itself. Examples of the studies’ limitations include design flaws (risk of bias), incomplete reporting of data, substantial missing data, biased reporting of outcomes, and implausible conclusions. Limitations of the process itself include the inability to obtain published articles or unpublished manuscripts, to assess the credibility of the grey literature, or to translate non-English articles.
- Conclude the systematic review with a general interpretation, being careful to restrict the interpretations to those fully supported by the evidence. Neutrally and objectively, the general interpretation puts the results of the systematic review into the context of the current knowledge on the topic, offers implications for practice, and provides recommendations for future research.45 Authors should state when they are unable to interpret results because the evidence is inadequate, insufficient, or absent.

These 4 steps are the basis of the scientific rigor of systematic reviews and why their results can be used as evidence. The second article in this 2-part series will present the additional steps for meta-analyses.

**CONCLUSION**

Systematic reviews underpin evidence-based decision making. Systematic reviews critically appraise all relevant research. This first article of a 2-part series explained how systematic reviews differ from traditional, narrative literature reviews. It defined key terms and differentiated the 4 types of systematic reviews. Most importantly, health care managers and supervisors were introduced to the process of conducting systematic reviews and meta-analyses. Thus, this article provided information essential to health care manager's and supervisor's understanding of systematic reviews.
REFERENCES

1. Pfeffer J, Sutton RI. Evidence-based management. Harv Bus Rev. 2006;84(1):62-74, 153.
2. Dyba T, Kitchenham BA, Jørgensen M. Evidence-based software engineering for practitioners. IEEE Software. 2005;22(1):58-65.
3. Ammenwerth E, de Keizer N. A viewpoint on evidence-based health informatics, based on a pilot survey on evaluation studies in health care informatics. J Am Med Inform Assoc. 2007;15(3):368-371.
4. Green J. The role of theory in evidence-based health promotion practice. Health Educ Res. 2000;15(2): 125-129.
5. Evidence-Based Medicine Working Group. Evidence-based medicine: a new approach to teaching the practice of medicine. JAMA. 1992;268(17):2420-2425.
6. Melnyk BM, Fineout-Overholt E. Key steps in implementing evidence-based practice: asking compelling, searchable questions and searching for the best evidence. Pediatr Nurs. 2002;28(3):262-263, 266.
7. Jennings ET, Hall JL. Evidence-based practice and the use of information in state agency decision making. J Public Adm Res Theory. 2012;22(2):245-266.
8. Anderson LM, Brownson RC, Fullilove MT, et al. Evidence-based public health policy and practice: promises and limits. Am J Prev Med. 2005;28(5S): 226-230.
9. Briggs HE, McBeath B. Evidence-based management: origins, challenges, and implications for social service administration. Adm Soc Work. 2009;3(3):242-261.
10. Kovner AR, Rundall TG. Evidence-based management reconsidered. Front Health Serv Manage. 2006;22(3): 3-22.
11. Shortell SM. Promoting evidence-based management. Front Health Serv Manage. 2006;22(3):25-29.
12. Fox DM. Systematic reviews and health policy: the influence of a project on perinatal care since 1988. Milbank Q. 2011;89(3):425-449.
13. University of York, Center for Reviews and Dissemination. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care. York, United Kingdom: York Publishing Service; 2009. http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf. Accessed June 12, 2013.
14. Higgins JPT, Green S. What is a systematic review (1.2.2.)? In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [Updated March 2011]. The Cochrane Collaboration; 2011. http://www.cochrane-handbook.org. Accessed June 12, 2013.
15. Agency for Healthcare Research and Quality. Methods Guide for Effective and Comparative Effectiveness Reviews. Rockville, MD: Agency for Healthcare Research and Quality; 2012. AHRQ publication no. 10(12)-EHC003-ED. http://www.effectivehealthcare.ahrq.gov or http://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&amp;1=1&amp;productID=518. Accessed June 3, 2013.
16. Relevo R, Balshem H. Finding evidence for comparing medical interventions: AHRQ and the Effective Health Care Program. J Clin Epidemiol. 2011;64(11):1168-1177.
17. Fu R, Gartlehner G, Grant M, et al. Conducting quantitative synthesis when comparing medical interventions: AHRQ and the effective health care program. J Clin Epidemiol. 2011;64(11):1187-1197.
18. Neyeloff JL, Fuchs SC, Moreira LB. Meta-analyses and forest plots using a Microsoft excel spreadsheet: step-by-step guide focusing on descriptive data analysis. BMC Res Notes. 2012;5:52.
19. Mays N, Pope C, Popay J. Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field. J Health Serv Res Policy. 2005;10(Suppl 1):6-20.
20. Landa AH, Szabo I, Le Brun L, Owen I, Fletcher G, Hill M. An evidence-based approach to scoping reviews. Electron J Inf Syst Eval. 2011;14(1):46-52.
21. Cochrane Collaboration. Glossary. http://www.cochrane.org/glossary. Accessed June 12, 2013.
22. Norris SL, Atkins D, Bruening W, et al. Observational studies in systematic reviews of comparative effectiveness: AHRQ and the Effective Health Care Program. J Clin Epidemiol. 2011;64(11):1178-1186.
23. Collins JA, Fauser BCJM. Balancing the strengths of systematic and narrative reviews. Hum Reprod Upd. 2005;11(2):103-104.
24. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 2005;8(1):19-32.
25. Brien SE, Lorenzetti DL, Lewis S, Kennedy J, Ghali WA. Overview of formal scoping review on health system report cards. Implement Sci. 2010;5:2. doi:10.1186/1748-5908-5:2.
26. Davis K, Drey N, Gould D. What are scoping studies? a review of the nursing literature. Int J Nurs Stud. 2009;46(10):1386-1400.
27. Khangura S, Konnyu K, Cushman R, Grimshaw J, Moher D. Evidence summaries: the evolution of a rapid review approach. Syst Rev. 2012;1:10.
28. Glass GV, Colosi D, Thomas H. Expediting systematic reviews: methods and implications of rapid reviews. Implement Sci. 2010;5:56. doi:10.1186/1748-5908-5:56.
29. Cohen PA. Meta-analysis: application to clinical dentistry and dental education. J Dent Educ. 1992;56(3):172-175.
30. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. A basic introduction to fixed-effect and random-effect models for meta-analysis. Res Synth Methods. 2010; 1(2):97-111.
31. Imperiale TF. Meta-analysis: when and how. Hepatology. 1999;29(6 suppl 1):268-315.
Evidence-Based Decision Making for Health Care Managers

109

34. Ioannidis JPA, Schmid CH, Lau J. Meta-analysis in hematology and oncology. Hematol Oncol Clin North Am. 2000;14(4):973-991.

35. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(6):e1000097.

36. Impellizzeri FM, Bizzini M. Systematic review and meta-analysis: a primer. Int J Sports Phys Ther. 2012; 7(5):493-503.

37. Hopewell S, Clarke MJ, Lefebvre C, Scherer RW. Handsearching versus electronic searching to identify reports of randomized trials. Cochrane Database Syst Rev 2007. 2008;(2):MR000001. doi:10.1002/14651858.MR000001.pub2.

38. Finnegan GA. New frontiers in grey literature: fourth international conference on grey literature. Coll Res Libr News. 1999;60(11):909.

39. US Department of Commerce. National Technical Information Service. National Technical Reports Library. http://www.ntis.gov/products/ntrl.aspx. Accessed June 7, 2013.

40. US National Library of Medicine. Health Services Research Projects in Progress (HSRProj). Last updated March 26, 2013. http://www.cvicf.nlm.nih.gov/hsr_project/home_proj.cfm. Accessed June 8, 2013.

41. Canadian Agency for Drugs and Technologies in Health. GreyMatters: a practical search tool for evidence-based medicine. 2013. http://www.cadth.ca/en/resources/finding-evidence-is/grey-matters. Accessed June 8, 2013.

42. American Psychological Association. PsychEXTRA®. http://www.apa.org/pubs/databases/psychextra/index .aspx. Accessed June 7, 2013.

43. Scirus. 2013. http://www.scirus.com/. Accessed June 8, 2013.

44. New York Academy of Medicine. Grey literature report. http://www.greylit.org/. Accessed June 5, 2013.

45. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009; 6(7):e1000100.

46. Gutman SI, Piper M, Grant MD, Basch E, Oliansky DM, Aronson N. Progression-Free Survival: What Does It Mean for Psychological Well-being or Quality of Life? Methods Research Report. Rockville, MD: Agency for Healthcare Research and Quality. 2013. Prepared by the Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-Based Practice Center under contract no. 290-2007-100584. AHRQ publication no. 13-EHC047-1. http://www.effectivehealthcare.ahrq.gov/ehc/products/522/998/MethodsGuideforCERs_Viswanathan_IndividualStudies.pdf or http://www.effectivehealthcare.ahrq.gov/. Accessed June 12, 2013.

47. Higgins JPT, Altman DG, Sterne JAC. The Cochrane Collaboration’s tool for assessing risk of bias (8.5). In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [Updated March 2011]. The Cochrane Collaboration; 2011. http://www.cochrane-handbook.org. Accessed June 12, 2013.

48. Whiting P, Rutjes AWS, Reitsma JB, Bossuyt PMM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003;3(25).

49. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Ann Intern Med. 2003;138(5):410-414.

50. Ried K. Interpreting and understanding meta-analysis graphs—a practical guide. Aust Fam Physician. 2006; 35(8):635-638.

51. Barton MB, Miller T, Wolff T, et al. How to read the new recommendation statement: methods update from the U.S. Preventive Services Task Force. Ann Intern Med. 2007;147(2):123-127.

52. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knothtnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol. 2011;64(4):380-382.

53. Butler M, Kane RL, McAlpine D, et al. Integration of Mental Health/Substance Abuse and Primary Care. Rockville, MD: Agency for Healthcare Research and Quality; 2008. Evidence Reports/Technology Assessments, no. 173. 3, Results. http://www.ncbi.nlm.nih.gov/books/NBK38628/. Accessed June 12, 2013.

54. Green LW, Glasgow RE, Atkins D, Stange K. Making evidence from research more relevant, useful, and actionable in policy, program planning, and practice slips ‘twixt cup and lip.’ Am J Prev Med. 2009; 37(6 suppl 1):S187-S191.