The art and science of knowledge synthesis

Andrea C. Triccoa, Jennifer Tetzlaffa,b, David Mohera,b,*

aOttawa Hospital Research Institute, The Ottawa Hospital, Ottawa, Ontario, Canada
bDepartment of Epidemiology & Community Medicine, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada

Accepted 25 November 2009

Abstract

Objectives: To review methods for completing knowledge synthesis.

Study Design and Setting: We discuss how to complete a broad range of knowledge syntheses. Our article is intended as an introductory guide.

Results: Many groups worldwide conduct knowledge syntheses, and some methods are applicable to most reviews. However, variations of these methods are apparent for different types of reviews, such as realist reviews and mixed-model reviews. Review validity is dependent on the validity of the included primary studies and the review process itself. Steps should be taken to avoid bias in the conduct of knowledge synthesis. Transparency in reporting will help readers assess review validity and applicability, increasing its utility.

Conclusion: Given the magnitude of the literature, the increasing demands on knowledge syntheses teams, and the diversity of approaches, continuing efforts will be important to increase the efficiency, validity, and applicability of systematic reviews. Future research should focus on increasing the uptake of knowledge synthesis, how best to update reviews, the comparability between different types of reviews (e.g., rapid vs. comprehensive reviews), and how to prioritize knowledge synthesis topics. © 2010 Elsevier Inc. All rights reserved.

Keywords: Systematic reviews; Knowledge syntheses; Methods; Risk of bias; Reporting; Dissemination

1. Introduction

“Knowledge synthesis” is not a new concept, with examples dating back to the early 1900s [1]. In the 1960s, research syntheses were common within social science, education, and psychology [1]. With the recognition of the importance of evidence-based decision making in health care came an understanding of the need for making decisions on the best available evidence and a demand for this evidence [2]. Since this time, knowledge syntheses have become increasingly important in health care. Knowledge translation focusing on the results of individual studies may be misleading due to bias in their conduct or random variations in findings [3]. This suggests that knowledge syntheses that interpret the results of individual studies within the context of global evidence should be considered as the basic unit of knowledge translation [4]. Syntheses provide the evidence base for other knowledge translation tools, such as policy briefs, patient decision aids, and clinical practice guidelines [4]. Additionally, granting agencies, such as the Canadian Institutes of Health Research, require knowledge syntheses to justify the need to fund and conduct randomized controlled trials (RCTs) [5]. Knowledge synthesis is central to knowledge translation, bridging the gap between research and decision making [6].

The terminology used to describe specific approaches to synthesis has evolved over time. There is no consistent use of these terms, but we will adopt the definitions used by The Cochrane Collaboration. Systematic reviews consist of a clearly formulated question and use systematic and explicit methods to identify, select, critically appraise, and extract and analyze data from relevant research. A meta-analysis consists of statistical techniques to quantitatively integrate the results of included studies. A knowledge synthesis may not necessarily include a meta-analysis.
What is new?

- Knowledge synthesis is used to interpret the results of individual studies within the context of global evidence and bridges the gap between research and decision making.
- Many groups worldwide conduct knowledge syntheses and some general methods are applicable to most reviews. However, variations of these methods are apparent for different types of reviews, such as realist reviews and mixed-model reviews.
- Review validity is dependent on the validity of the included primary studies and the review process itself. Steps should be taken to avoid bias in knowledge synthesis conduct. Transparency in reporting will help readers assess review validity and applicability, increasing its utility.
- Given the magnitude of the existing literature, the increasing demands on systematic review teams, and the diversity of approaches, continuing methodological effort will be important to increase the efficiency, validity, and applicability of systematic reviews.
- Future research should focus on increasing the uptake of knowledge synthesis, how best to update reviews, the comparability between different types of reviews (eg, rapid vs. comprehensive reviews), and how to prioritize knowledge synthesis topics.

Many groups worldwide conduct systematic reviews. The Cochrane Collaboration often answers questions regarding the efficacy and effectiveness of an intervention [7]. As such, they strongly rely on synthesizing evidence from RCTs. Other organizations will include other study designs in their systematic reviews. Examples include the Campbell Collaboration [8], which addresses questions related to crime and justice, education, and social welfare; the Evidence-based Practice Center program [9], which completes knowledge syntheses in a wide variety of health care and policy areas; and the Joanna Briggs Institute, which conducts knowledge syntheses on health issues of interest to health care professionals, such as nurses and midwives [10]. These groups may use different templates for conducting systematic reviews. The explicit methods involved will depend, in part, on the question being considered.

Systematic reviews addressing questions about the effectiveness of interventions will usually include (or attempt to include) quantitative data (eg, odds ratio for a particular drug vs. placebo). However, these systematic reviews could also include qualitative evidence to provide contextual insight [11]. Examples of the types of reviews that incorporate rich information and address broader questions include realist reviews and metanarrative reviews (Table 1). Realist reviews aim to identify and test theories [24], while metanarrative reviews also aim to identify and test theories, but the focus is on the overarching story line that unfolds over time within a community of scholars as they address a particular research area [22].

Including qualitative evidence in systematic reviews, although important, is often difficult to do [26]. Some of the challenges that systematic reviewers face include searching for and identifying qualitative studies [26], appraising the quality of qualitative research [27], and synthesizing qualitative and quantitative research [28]. The Cochrane Qualitative Research Methods Group has been working on the issues involved with including qualitative research in systematic reviews since 1998 (http://www.joannabriggs.edu.au/cqrmg/about.html), and other researchers are working on ways to combine quantitative and qualitative evidence into a systematic review via a mixed-model approach [29].

Variations in how systematic reviews are conducted demonstrate the creative component (or the “art”) of knowledge synthesis. Although these variations in approaches to conducting systematic reviews are innovative, there are limited examples of their use and even less methodological research validating them. As such, the following sections will focus on general methods for conducting systematic reviews that are applicable to most reviews (Table 2).

2. The review team

Before beginning a review, a systematic review team should be identified. The optimal team required is determined by the type of question being addressed and generally consists of clinical or content experts with extensive knowledge of the review topic, methodologists with expertise in systematic reviews, a librarian to help search the literature comprehensively [30], and epidemiologists or other researchers with experience in conducting primary research on the topic. The funder or commissioning agency may help inform the context of the question, and a statistician may be consulted if statistical pooling (ie, meta-analysis) is being considered.

Some review teams also involve end users of the review, such as policy makers, health care professionals, or patients, throughout the review process to ensure that the review is relevant and user friendly. For example, the Canadian Health Services Research Foundation fosters relationships between decision makers and researchers by funding knowledge syntheses that are conducted using a collaborative approach (http://www.chsrf.ca/dss/framework_e.php). Decision makers lead in identifying the key issues and translate the results to policy makers with help from researchers. Researchers conduct the knowledge synthesis in a scientifically rigorous manner and formulate the implications of their results with help from the decision makers.
Table 1: Variations in questions and methods for different types of systematic reviews

<table>
<thead>
<tr>
<th>Type of review</th>
<th>Sources of evidence used</th>
<th>Type of question</th>
<th>Example of question</th>
<th>Specific methods used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention review [12]</td>
<td>Qualitative or quantitative</td>
<td>Does the intervention of interest work for a particular outcome?</td>
<td>What is the effectiveness of vitamin E for the treatment of cardiovascular disease? [13]</td>
<td>• Studies: often limited to data from experimental studies</td>
</tr>
<tr>
<td>Network meta-analysis [14]</td>
<td>Quantitative</td>
<td>Does the intervention of interest work for a particular outcome?</td>
<td>What are the odds of developing diabetes during long-term treatment with an initial class of antihypertensive drug? [15]</td>
<td>• Risk of bias assessment: often focused on experimental studies</td>
</tr>
<tr>
<td>Diagnostic test review [16]</td>
<td>Qualitative or quantitative</td>
<td>How well does the diagnostic test work for a particular group of patients?</td>
<td>What is the diagnostic accuracy of sentinel node biopsy, positron-emission tomography, magnetic resonance imaging, and computed tomography in determining lymph node status in patients with cervical cancer? [17]</td>
<td>• Analysis: may be qualitative or quantitative (meta-analysis)</td>
</tr>
<tr>
<td>Human genome epidemiology reviews [18]</td>
<td>Quantitative</td>
<td>Which genes are associated with particular outcomes?</td>
<td>What is the susceptibility of 160A allele carriers to seven types of cancers? [19]</td>
<td>• Studies: often limited to data from observational studies</td>
</tr>
<tr>
<td>Prognostic review [20]</td>
<td>Quantitative</td>
<td>How can you predict a disease outcome more accurately or efficiently?</td>
<td>Does B-type natriuretic peptide (BNP) predict mortality or other cardiac endpoints in persons diagnosed with coronary artery disease? [21]</td>
<td>• Risk of bias: focuses on issues pertinent to genome association studies</td>
</tr>
<tr>
<td>Meta-narrative review [22]</td>
<td>Qualitative</td>
<td>How best can one explain complex bodies of evidence?</td>
<td>How best to explain the diffusion of innovation in evidence-based medicine? [22]</td>
<td>• Analysis:</td>
</tr>
<tr>
<td>Realist review [23]</td>
<td>Qualitative</td>
<td>How do complex programs work (or why do they fail) in certain contexts and settings?</td>
<td>Which aspects of school feeding programs in disadvantaged children determine success and failure in various situations? [24]</td>
<td>• Questions: framed in broad open-ended format</td>
</tr>
<tr>
<td>Meta-ethnography review [25]</td>
<td>Qualitative</td>
<td>How can qualitative evidence explain why certain interventions work and others do not?</td>
<td>What are the types of factors that could influence adherence to tuberculosis treatment from the patient’s experience? [25]</td>
<td>• Literature search: involves browsing relevant perspectives and approaches, finding seminal conceptual articles by tracking references of references</td>
</tr>
</tbody>
</table>

A.C. Tricco et al. / Journal of Clinical Epidemiology - (2010) - ARTICLE IN PRESS
3. How do we formulate the question, protocol, and eligibility criteria?

Developing a clear and concise question is the first and one of the most important steps in conducting a systematic review, as this will guide the review process. Formulating relevant questions can be complex, and the Population, Intervention, Comparators, Outcome, and (Study design or Time period) (PICO(S or T)) acronym has been proposed as a structured way to facilitate this process [31]. With this approach, the question addresses the population, participants, or problem (P); the interventions, independent variables, or index test (for diagnostic reviews) (I); the comparators (C; eg, placebo or standard of care or context, in the case of qualitative studies); and the dependent variables, endpoints, or outcomes of interest (O). Sometimes, additional components, such as study design (S) or time period (T), are used to limit the systematic review to certain study types (eg, RCTs, cohort studies) or periods. This paradigm will not fit all questions but may be a useful guideline. For example, Intervention can be substituted with Exposure in systematic reviews of epidemiological studies.

Once the review team has been assembled and the objectives defined, a protocol, prespecifying the systematic review methods, should be developed to guide the process. Protocol use may decrease the likelihood of biased post hoc changes to methods and selective reporting. Important elements of the review protocol are provided elsewhere but include details on the methods used for the search, retrieval and appraisal of the literature, and data abstraction [12]. The systematic review process may be iterative and, as such, the protocol may change over time; this is especially the case for qualitative reviews, which are iterative by nature [32]. Changes to the review protocol are acceptable; however, they should be transparently described in the final report.

The eligibility criteria of studies to be included in the synthesis should extend from the components of the review question and may be based on study characteristics or report characteristics. Study characteristics include the PICO(S or T) criteria as well as others, such as length of follow-up. Report characteristics include publication language and status (eg, published or unpublished material). Evidence suggests that selective inclusion of studies based on report characteristics (eg, publication status or publication language) may bias the effect estimate in meta-analyses [33–38]. Conversely, caution may need to be exercised in all-inclusive approaches to study selection due to potential differences in the risk of bias from some sources, such as selective reporting in abstracts vs. full text reports [39–43]. Regardless of the choices made, it is recommended that the eligibility criteria are thoroughly considered, properly defined, and transparently reported to avoid ambiguity in the review process and to inform the validity of the review.

4. How do we find relevant studies?

The review question or PICO(S or T) components are used to guide the location of relevant studies, usually entailing bibliographic database searches and other methods. Commonly searched electronic databases for locating health-related research are MEDLINE [44], EMBASE [45], and The Cochrane Central Register of Controlled Trials [46]. Reviewers may also search subject-specific databases, such as the Cumulative Index of Nursing and Allied Health (CINAHL), or geographical databases, such as the Latin American Caribbean Health Sciences Literature (LILACS). The scope of the databases and the review should be considered to select the most relevant databases for the review, and searching more than one database is highly recommended to overcome potential indexing biases [47]. Consulting a librarian ensures that the search strategies are comprehensive.

Qualitative evidence is cataloged in a variety of different databases, and a register of qualitative research does not
exist [26]. This makes searching for qualitative evidence difficult to accomplish. Working with a librarian who has experience with locating qualitative research may facilitate the identification of qualitative evidence.

Systematic reviewers often use other sources to supplement their electronic searches, such as hand searching journals, searching the reference lists of included studies, or searching trial registries. Only including published literature in a systematic review may lead to publication bias because published studies have a greater likelihood of having statistically significant results than unpublished studies [48]. As such, searching the gray literature\(^1\) should be considered, including Web sites from funding agencies, health policy groups, and ministries of health, among others. More detailed searching information can be found elsewhere [12].

5. How do we select studies?

Systematic reviewers generally separate the study selection process into two stages: (1) a broad screen of the titles and abstracts of the citations retrieved from the literature search and (2) a strict screen of the full text articles passing the broad screen to select the final included studies. This selection is usually facilitated through the use of eligibility criteria. Two or more reviewers should screen the material independently to ensure that relevant material is not excluded from the systematic review. Results are then compared and conflicts may be resolved by discussion or with the involvement of a third reviewer. Agreement across reviewers should also be recorded and reported. Interrater agreement can be statistically assessed using the kappa statistic [49].

The process of identifying and selecting studies requires detailed record keeping because it must be reported in sufficient detail for the end user to determine validity. The ratio of reports to studies is not always 1:1, as some reports will describe multiple studies, and some studies are described in more than one report. Duplicate publications are not always obvious [50–53], and authors should attempt to identify such duplicate data. Using data from the same participants more than once may exaggerate the estimated treatment effect [51]. As such, reviewers should use the major study publication for data abstraction, and the other publication(s) can be used as supplementary material.

6. How do we assess risk of bias of included studies?

The validity of the results of a systematic review will greatly depend on the risk of bias in the individual studies. Risk of bias assessment can be completed using scales, checklists, or components, and many tools are available for different study designs [54,55]. Numerous tools have been developed to appraise the quality of qualitative evidence [32], and an example of three different quality appraisal strategies for a systematic review including qualitative evidence has been published [27]. Regardless of the instrument used, the assessments for each criterion from the quality appraisal should be reported for each study. Simply reporting a score from an assessment scale might be flawed [36] and is not as helpful to the end user because it lacks sufficient detail to describe where sources of bias may arise. Excluding studies from a review based on their risk of bias is not advisable. Rather, the impact of this potential bias can be addressed through sensitivity analyses, which explores whether results of the review are robust to differences in the trials, such as methodology (eg, examining studies with and without allocation concealment analyzed separately) and populations examined.

7. How do we extract data from the individual studies?

At the time of protocol development, the information sought from the included studies should be considered. The outcome(s) of primary importance (eg, clinical, patient or policy relevant) should be differentiated from the “secondary” outcomes. Recent surveys have found that authors of RCTs modified primary outcomes between the protocol and the final report in approximately 50% of trials [56] and that the outcomes selectively included in the final reports were more likely to be statistically significant than those omitted [57,58]. Therefore, if a review is limited only to variables that are reported in the included studies rather than identifying those considered important at the outset, the review risks being subject to bias.

It is advisable to develop a data extraction form a priori including the variables to be collected and clear definitions for them. The form can be pilot tested by the review team members to increase the reliability of the data extraction process. Having two or more people extract study data independently may also decrease the potential for error. Reviewers should consider contacting authors to verify assumptions made for missing or unclear information.

8. How do we analyze the data?

The analysis method will depend on the question(s) being asked and the type of data collected; however, all systematic reviews should at least include a narrative synthesis describing the results and risk of bias in the included studies. For a typical intervention review including quantitative data, standard effect measures will need to be chosen, if possible, to compare studies (eg, odds ratio, standardized mean difference, hazard ratio). The next step usually involves determining whether statistical synthesis (ie, meta-analysis) is

\(^1\) Information produced on all levels of government, academics, business and industry in electronic and print formats not controlled by commercial publishing, that is, where publishing is not the primary activity of the producing body [85].
possible and appropriate. For example, quantitative synthesis may not be suitable if outcome assessment across the included studies is inconsistent or evidence of substantial clinical, methodological, or statistical heterogeneity of the included studies exists [12]. Guidance on effect measures, approaches to detect heterogeneity, and meta-analysis techniques are available elsewhere [12,59,60].

Qualitative approaches of analysis differ from quantitative methods. For example, qualitative evidence may be inputted into matrices or tables to allow comparison across studies [61]. Some knowledge syntheses will include both qualitative and quantitative data for which a variety of methods are available [61]. Examples include a quantitative case survey, where qualitative data are converted into quantitative form and analyzed statistically (eg, through meta-analysis) [61], and Bayesian meta-analysis, which allows the incorporation of qualitative research into a quantitative synthesis to provide policy makers with decision support [62,63].

9. How can we present the results of the review?

Results of knowledge syntheses may be presented in numerous ways. The screening process may be described in the text and/or presented as a flow diagram (Fig. 1) [64]. Many journals are requiring this information to be presented as a flow diagram to facilitate transparency of the process. Characteristics of included studies, such as descriptions of study designs, participant populations, and interventions, are generally presented by the study in tabular form and/or synthesized textually. The results of risk of bias assessments may also be presented in a table or text, and sufficient detail should be presented to allow the end user to be able to determine the potential threats to validity.

Quantitative data should be presented as summary data (eg, 2 × 2 tables of counts, mean, and standard deviations) and effect estimates (eg, odds ratio, difference in means) with confidence intervals for each study, where possible. These data may be presented for each outcome in a table or in a forest plot, with the combined effect estimate of the meta-analysis, if relevant (Fig. 2). Qualitative data may also be presented visually, for example, through a conceptual framework diagram. Results of all other analyses, such as assessment of publication bias, should also be reported.

10. How can we interpret the results?

Reviewers should discuss the quality, strength, and applicability of the evidence for each main outcome when summarizing the results. Formal assessment approaches do exist [65], although no standard method is universally recommended. The relevance of the results should also be considered for key stakeholders (eg, policy makers, patients, health care providers) because this will help increase the applicability of the results for these groups. For example, policy makers involved with a systematic review from the beginning might ensure that the interpretation and
implications of the review’s results are politically relevant. As mentioned above, involving these groups at the outset of the review (e.g., in defining the research question, choosing eligibility criteria and outcomes) will increase the applicability of the knowledge synthesis results.

Qualitative evidence can help with the interpretation of the systematic review results by explaining how the intervention worked and whether it will work in a different setting (http://www.joannabriggs.edu.au/cqrmg/about.html). For example, a Cochrane review that assessed how to improve communication with children about their cancer [66] could include qualitative evidence to provide information about why particular interventions facilitate or hinder successful communication and how these strategies could be improved from the child’s perspective [11,67]. Qualitative evidence can also be used to identify the facilitators or barriers involved with uptake of the intervention, as well as the subjective experience of individuals who received, developed, or delivered the intervention (http://www.joannabriggs.edu.au/cqrmg/about.html).

Reviewers should consider both study- and review-level limitations. If the conduct or reporting of included studies is poor, the review conclusions may be biased and this should be stated explicitly. Furthermore, knowledge syntheses themselves can be susceptible to bias. A recent systematic review summarized the evidence concerning possible bias and confounding in systematic reviews and found that they should include unpublished material, hand search for additional material, search multiple databases, assess for publication bias, and periodically update the systematic review [47].

Despite efforts to decrease bias in knowledge syntheses, some review limitations will exist and should be noted. One way to assess a review’s strengths and limitations is to use the “Assessment of Multiple SysTematic Reviews” (AMSTAR) tool, which has high interrater reliability and construct validity [68]. It addresses the use of a protocol for the review, study selection and data abstraction process, comprehensiveness of the literature search, inclusion of unpublished literature, documentation of included and excluded studies, reporting of included study characteristics, assessment and interpretation of study quality, combining methods of the results, assessment of publication bias, and reporting of conflicts of interest.

Finally, reviewers should carefully draw conclusions based on the available evidence. Conclusions may include specific recommendations for decision making or for research [69]. If no conclusions can be drawn due to insufficient reliable evidence, this should be stated as it may indicate the need for further research.

11. How do we disseminate the results of our review?

The final step in the review process is making the results accessible. The most common form of dissemination is publication in peer-reviewed journals, with recent estimates...
suggesting over 2,500 English language systematic reviews being indexed annually in MEDLINE [70]. Open access journals offer broader readership base to low- and medium-income countries. Other forms may include targeted dissemination via media for the public [71], brief reports for health care providers, policy makers and consumers [72], and decision aids for patients [73]. Uptake of the results of systematic reviews may be impeded by many factors, but one that is in the author’s control is the quality of the review report. Transparent descriptions of the review methods and results allow readers to assess the methods, risk of bias of the included studies, and the review and inform them of the applicability of the review.

Recent evidence suggests that systematic review reporting is suboptimal [70]. The QUOROM Statement (QUality Of Reporting Of Meta-analyses) was developed to improve the reporting of systematic reviews and meta-analysis [74]. This reporting guideline has been updated and expanded recently (PRISMA—Preferred Reporting Items for Systematic reviews and Meta-Analyses) to address advances in systematic review methodology and changes in terminology [64,75]. Similar initiatives are also available for reviews of observational studies [76].

12. How do we increase the uptake of our review results?

There is limited evidence to support how systematic reviews should be presented to enhance uptake in decision making. Despite advances in the conduct and reporting of systematic reviews (and recognition of their importance in knowledge translation), current evidence suggests that they may be used infrequently by health care providers, patients, and others to make decisions. For example, a systematic review of the information-seeking behavior of physicians found that textbooks (many of which do not rely on evidence from systematic reviews) are still the most frequent source of information followed by advice from colleagues [77].

Given that systematic reviews of randomized trials are less susceptible to bias than the opinions of experts and observational data, why are they used so infrequently? There are many answers to this question, which can be broadly categorized into the relevance of the questions the reviews are addressing, the lack of contextualization, and the format of presentation. Although much attention has been paid to enhancing the quality of systematic reviews, relatively little attention has been paid to the format for presenting the review. Because the reporting of systematic reviews tends to focus on methodological rigor more than clinical context, they often do not provide crucial information for clinicians. In one study, the researchers found that of systematic reviews published in ACP JC and EBM Journal (journals of secondary publication), less than 15% of these had sufficient information about the intervention to allow clinicians or policy makers to implement it [78].

Current efforts are trying to improve the utility of systematic review by making them more user friendly. Resources aiming to provide easy access to current evidence from systematic reviews to health care providers include Clinical Evidence in the UK (http://clinicalevidence.bmj.com/ceweb/index.jsp) and Up-to-Date in the USA (http://www.uptodate.com/home/index.html). Resources for policy makers include Rx for Change (http://www.cadth.ca/index.php/en/compus/optimal-ther-resources/interventions) and the Program in Policy Decision-Making/Canadian Cochrane Network and Centre (PPD/CCNC) database (http://www.researchtopolicy.ca/Search/Reviews.aspx), which are based in Canada. In addition, enabling The Cochrane Library to be publicly available may increase the uptake of Cochrane reviews by health care providers, patients, and policy makers.

13. Future research

Future research is needed to assess various aspects of systematic review methodology and to help guide review prioritization and updating. For example, is there sufficient evidence for duplicate screening in reviews given the resource implications or are sampling methods to check validity sufficient? Although new research is emerging on the importance of updating systematic reviews [79,80] and suggested frequency and methods of updating [81,82], more research in this topic is required. Additionally, the comparability of different systematic review approaches to answer the same question is unknown. For example, as rapid reviews emerge as a method to address timely but pressing topics, is there any evidence that the different methods provide consistent results compared with more comprehensive reviews? Also, the utility of efforts to increase the uptake of systematic reviews, such as Clinical Evidence, Up-to-date, Rx for Change, and the PPD/CCNC database, needs to be formally evaluated. Finally, methods to prioritize systematic review topics, such as efforts by the Agency for Healthcare Research and Quality Effective Healthcare program [83] and the “Listening for Direction” initiative by the Canadian Health Services Research Foundation [84], are becoming of interest and may warrant future research.

References


