

The Systematic Review: An Overview

Synthesizing research evidence to inform nursing practice.

This article is the first in a new series on systematic reviews from the Joanna Briggs Institute, an international collaborative supporting evidence-based practice in nursing, medicine, and allied health fields. The purpose of the series is to show nurses how to conduct a systematic review—one step at a time. This first installment provides a synopsis of the systematic review as a scientific exercise, one that influences health care decisions.

Research in the health sciences has provided all health care professions, including nursing, with much new knowledge to inform the prevention of illness and the care of people with ill health or trauma. As the body of research has grown, so too has the need for rigorous syntheses of the best available evidence.

Literature reviews have long been a means of summarizing and presenting overviews of knowledge, current and historical, derived from a body of literature. They often make use of the published literature; generally, published papers cited in a literature review have been subjected to the blind peer-review process (a hallmark of most scientific periodicals). The literature included in a literature review may encompass research reports that present data, as well as conceptual or theoretical literature that focuses on a concept.¹

An author may conduct a literature review for a variety of reasons, including to¹

- present general knowledge about a topic.
- show the history of the development of knowledge about a topic.
- identify where evidence may be lacking, contradictory, or inconclusive.
- establish whether there is consensus or debate on a topic.
- identify characteristics or relationships between key concepts from existing studies relevant to the topic.
- justify why a problem is worthy of further study.

All of these purposes have been well served by a “traditional” or “narrative” review of the literature. Traditional literature reviews, though useful, have major drawbacks in informing decision making in nursing practice. Predominantly subjective, they rely heavily on the author’s knowledge and

experience and provide a limited, rather than exhaustive, presentation of a topic.² Such reviews are often based on references chosen selectively from the evidence available, resulting in a review inherently at risk for bias or systematic error. Traditional literature reviews are useful for describing an issue and its underlying concepts and theories, but if conducted according to no stated methodology, they are difficult to reproduce—leaving the findings and conclusions resting heavily on the insight of the authors.^{1,2} In many cases, the author of the traditional review discusses only major ideas or results from the studies cited rather than analyzing the findings of any single study.

With the advent of evidence-based health care some 25 years ago, nurses and other clinicians have been expected to refer to and rely on research evidence to inform their decision making. The need for evidence to support clinical practice is constantly on the rise because of advances that continually expand the technologies, drugs, and other treatments available to patients.³ Nurses must often decide which strategies should be implemented, yet comparisons between available options may be difficult to find because of limited information and time, particularly among clinical staff. Furthermore, interpreting research findings as presented in scientific publications is no easy task. Without clear recommendations for practice, it can be difficult to use evidence appropriately; it requires some knowledge of statistics and in some cases extensive knowledge or experience in how to apply the evidence to the clinical setting.⁴ Also, many health care devices and drugs come with difficult-to-understand claims of effectiveness.⁵

As a result of research, the knowledge on which nursing care is based has changed at a rapid pace.

This inexorable progress means that nurses can access biomedical databases containing millions of citations pertinent to health care; these databases are growing at a phenomenal rate, with tens of thousands of publications added every year. The volume of literature is now too vast for nurses and other health care professionals to stay on top of.³ Furthermore, not all published research is of high quality and reliable; on the contrary, many published studies have used inappropriate statistical methods or have otherwise been poorly conducted.

used in the synthesis is also a hallmark of any well-conducted systematic review. Reporting standards similar to those produced for primary research designs have been created for systematic reviews. The PRISMA statement, or Preferred Reporting Items for Systematic Reviews and Meta-Analyses, provides a checklist for review authors on how to report a systematic review.¹⁰ Ultimately, the quality of a systematic review, and the recommendations drawn from it, depends on the extent to which methods are followed to minimize the risk of error

It can be difficult for nurses to know which studies from among the multitude available should be used to inform the decisions they make every day.

Such issues affecting research quality can make for research findings that are contradictory or inconclusive. Similarly, using the results of an individual study to inform clinical decision making is not advisable. When compared with other research on the topic, a study may be at risk for bias or systematic error.⁵ Therefore, it can be difficult for nurses to know which studies from among the multitude available should be used to inform the decisions they make every day. As a result, reviews of the literature have evolved to become an increasingly important means by which data are collected, assessed, and summarized.⁵⁻⁷

THE SYSTEMATIC REVIEW

Since the traditional literature review lacks a formal or reproducible means of estimating the effect of a treatment, including the size and precision of the estimate,^{2,7} a considerably more structured approach is needed. The “systematic review,” also known as the “research synthesis,” aims to provide a comprehensive, unbiased synthesis of many relevant studies in a single document.^{2,7,8} While it has many of the characteristics of a literature review, adhering to the general principle of summarizing the knowledge from a body of literature, a systematic review differs in that it attempts to uncover “all” of the evidence relevant to a question and to focus on research that reports data rather than concepts or theory.^{3,9}

As a scientific enterprise, a systematic review will influence health care decisions and should be conducted with the same rigor expected of all research. Explicit and exhaustive reporting of the methods

and bias. For example, having multiple steps in the systematic review process, including study selection, critical appraisal, and data extraction conducted in duplicate and by independent reviewers, reduces the risk of subjective interpretation and also of inaccuracies due to chance error affecting the results of the review. Such rigorous methods distinguish systematic reviews from traditional reviews of the literature.

The characteristics of a systematic review are well defined and internationally accepted. The following are the defining features of a systematic review and its conduct:

- clearly articulated objectives and questions to be addressed
- inclusion and exclusion criteria, stipulated a priori (in the protocol), that determine the eligibility of studies
- a comprehensive search to identify all relevant studies, both published and unpublished
- appraisal of the quality of included studies, assessment of the validity of their results, and reporting of any exclusions based on quality
- analysis of data extracted from the included research
- presentation and synthesis of the findings extracted
- transparent reporting of the methodology and methods used to conduct the review

Different groups worldwide conduct systematic reviews. The Cochrane Collaboration primarily addresses questions on the effectiveness of interventions or therapies and has a strong focus on synthesizing evidence from randomized controlled trials

(RCTs) (see <http://handbook.cochrane.org>).⁹ Other groups such as the Centre for Reviews and Dissemination at the University of York (<http://bit.ly/1g9WoCq>) and the Joanna Briggs Institute (<http://joannabriggs.org>) include other study designs and evidence derived from different sources in their systematic reviews. The Institute of Medicine issued a report in 2011, *Finding What Works in Health Care: Standards for Systematic Reviews*, which makes recommendations for ensuring “objective, transparent, and scientifically valid reviews” (see <http://bit.ly/1cRIAq7>).

How systematic reviews are conducted may vary; the methods used will ultimately depend on the question being asked. The approach of the Cochrane Collaboration is almost universally adopted for a clear-cut review of treatment effectiveness. However, specific methods used to synthesize qualitative evidence in a review, for example, may depend on the preference of the researchers, among other factors.⁷ The steps for conducting a systematic review will be addressed below and in greater detail throughout this series.

Adhering to the eligibility criteria stipulated in the review protocol ensures that studies selected for inclusion are selected based on their research method, as well as on the PICO elements of the study, and not solely on the study’s findings.³ Conducting the review in such a fashion limits the potential for bias and reduces the possibility of altering the focus or boundaries of the review after results are seen. In addition to the PICO elements, the inclusion criteria should specify the research design or types of studies the review aims to find and summarize, such as RCTs when answering a question on the effectiveness of an intervention or therapy.⁹ Stipulating “study design” as an extra element to be included as part of the inclusion criteria changes the standard PICO mnemonic to PICOS.

Searching for studies can be a complex task. The aim is to identify as many studies on the topic of interest as is feasible, and a comprehensive search strategy must be developed and presented to readers.^{3,10} A strategy that increases in complexity is common, starting with an initial search of major databases, such as MEDLINE (accessed through PubMed) and

As a scientific enterprise, a systematic review will influence health care decisions and should be conducted with the same rigor expected of all research.

Review question and inclusion criteria. Systematic reviews ideally aim to answer specific questions, rather than present general summaries of the literature on a topic of interest.^{5,8} A systematic review does not seek to create new knowledge but rather to synthesize and summarize existing knowledge, and therefore relevant research must already exist on the topic.^{3,5} Deliberation on the question occurs as a first step in developing the review protocol.^{5,7} Nurses accustomed to evidence-based practice and database searching will be familiar with the PICO mnemonic (Population, Intervention, Comparison intervention, and Outcome measures), which helps in forming an answerable question that encompasses these concepts to aid in the search.^{3,8,11} (The art of formulating the review question will be covered in the second article of this series.)

Ideally, the review protocol is developed and published before the systematic review is begun. It details the eligibility of studies to be included in the review (based on the PICO elements of the review question) and the methods to be used to conduct the review.

the Cumulative Index to Nursing and Allied Health (CINAHL), using keywords derived from the review question. This preliminary search helps to identify optimal search terms, including further keywords and subject headings or indexing terms, which are then used when searching all relevant databases. Finally, a manual search is conducted of the reference lists of all retrieved papers to identify any studies missed during the database searches. The search should also target unpublished studies to help minimize the risk of publication bias^{3,5}—a reality that review authors have to acknowledge. It arises because researchers are more likely to submit for publication positive rather than negative findings of their research, and scientific journals are inclined to publish studies that show a treatment’s benefits. Therefore, relying on findings only from published studies may result in an overestimation of the benefits of an intervention. To date, locating unpublished studies has been difficult, but resources for locating this “gray” literature are available and increasing in sophistication. For example, Web search engines can search across many

governmental and organizational sites simultaneously. Similarly, there are databases that index graduate theses and doctoral dissertations, abstracts of conference proceedings, and reports that aren't commercially published. Contacting experts in the field may also yield otherwise difficult-to-obtain information. Finally, studies published in languages other than English should be included, if possible, despite the added cost and complexity of doing so. (The art of searching will be addressed in the third paper in this series.)

Study selection and critical appraisal. The PICO elements can aid in defining the inclusion criteria used to select studies for the systematic review. The inclusion criteria place the review question in a practical context and act as a clear guide for the review team as they determine which studies should be included.³ This step is referred to as *study selection*.³ Once it's determined which studies should be included, their quality must be assessed during the step of *critical appraisal*. (Both of these steps will be further addressed in the fourth paper in this series.)

During study selection, reviewers look to match the studies found in the search to the review's inclusion criteria—that is, they identify those studies that were conducted in the correct population, use interventions of interest, and record the predetermined and relevant outcomes.³ The optimal research design for answering the review question is also determined. For example, for a systematic review evaluating the

of an intervention. A systematic review aims to synthesize the best evidence for clinical decision making. Assessing the validity of a study requires careful consideration of the methods used during the research and establishing whether the study can be trusted to provide a reliable and accurate account of the intervention and its outcomes.⁵⁻⁸ Studies that are of low or questionable quality are generally excluded from the remainder of the review process. Exclusion of lesser-quality studies reduces the risk of error and bias in the findings of the review.³ For the most part, critical appraisal focuses squarely on research design and the validity and hence the believability of the study's findings rather than on the quality of reporting, which depends on both writing style and presentation.¹⁰ For example, when assessing the validity of an RCT, critical appraisal generally focuses on four types of systematic error that can occur at different stages of a study: selection bias (in considering how study participants were assigned to the treatment groups), performance bias (in considering how the intervention was provided), attrition bias (in considering participant follow-up and drop-out), and detection bias (in considering how outcomes were measured).³

To aid the transparency and reproducibility of this process in the systematic review, standardized instruments (checklists, scales) are commonly used when asking the reviewers about the research they are reading.

A systematic review does not seek to create new knowledge but rather to synthesize and summarize existing knowledge.

effectiveness of an intervention, the most reliable evidence is thought to come from RCTs, which allow the inference of causal associations between the intervention and outcome, rather than from other study designs such as the cohort study, which lacks randomization and experimental "control." Any exclusion criteria should also be documented—for example, specific populations or modes of delivery of an intervention.

During critical appraisal, all studies to be included are first assessed for methodologic rigor.³ Although there are some subtle differences, this appraisal is akin to assessing the risk of bias in reviews that ask questions related to the effectiveness

Data extraction and synthesis. Once the quality of the research has been established, relevant data aligned to the predetermined outcomes of the review must be extracted for the all-important synthesis of the findings. (These steps will be addressed in the fifth paper in this series.) Data synthesized by systematic reviews are the results extracted from the individual research studies; as with critical appraisal, data extraction is often facilitated by the use of a tool or instrument that ensures that the most relevant and accurate data are collected and recorded.³ A tool may prompt the reviewer to extract relevant citation details, details of the study participants including their number and eligibility,

descriptive details of the intervention and comparator used in the study, and the all-important outcome data. Generic extraction tools for both quantitative and qualitative data are readily available.¹² The data collected from individual studies vary with each review, but they should always answer the question posed by the review. While undertaking a review, reviewers will find that data extraction can be quite difficult—often complicated by factors of the included studies such as incomplete reporting of study findings and differing ways of reporting and presenting data. When these issues arise, reviewers should attempt to contact the authors to obtain missing data, particularly for recently published research.⁵

across studies. Meta-analysis requires transforming the findings of treatment effect from individual studies into a common metric and then using statistical procedures across all of the findings to determine whether there is an overall effect of the treatment or association.^{8,9,14} The typical output from a statistical synthesis of studies is the measure or estimate of effect; the confidence interval, which indicates the precision of the estimate; and the quantification of the differences (heterogeneity) between the included studies and the statistical impact of these differences, if any, on the analysis. There are many different statistical methods by which results from individual studies can be combined during the meta-analysis. The results of the

While the methods of synthesizing quantitative data are relatively straightforward and accepted, there are numerous methods for synthesizing qualitative research.

Data synthesis is a principal feature of the systematic review.^{3,6,7,9} There are various methods available, depending on the type of data extracted that's most appropriate to the review question.⁷ An example of a systematic review addressing a question of the effectiveness of a nursing intervention is one examining nurse-led cardiac rehabilitation programs following coronary artery bypass graft surgery; the review aims to give an overall estimate of the intervention's effectiveness on patients' health-related quality of life and hospital readmission rates.¹³ Depending on the question asked, such a synthesis of the results of relevant studies also allows for exploration of similarities or inconsistencies of the treatment effect in different studies and among various settings and populations.³ Where inconsistencies are apparent they can be analyzed further. The synthesis either provides a narrative summary of the included studies or, where possible, statistically combines data extracted from the studies. This pooling of data is termed "meta-analysis."¹⁴

A *meta-analysis* may be included in a systematic review as a practical way of evaluating many studies. Meta-analysis should ideally be undertaken only when studies are similar enough; studies should sample from similar populations, have similar objectives and aims, administer the intervention of interest in a similar fashion, and (most important) measure the same outcomes.³ Meta-analysis is rarely appropriate when such similarities do not appear

meta-analysis are commonly displayed as a forest plot, which gives the reader a visual comparison of the findings.

Owing to the limited availability of relevant trials, reviews that aim to examine the effectiveness of an intervention may resort to evidence from experimental studies other than RCTs and even from observational studies; such reviews have the potential to play a greater role in evidence-based nursing, where trials, historically, have been rare.¹⁵ But when conducting a systematic review of studies using designs other than the RCT, a reviewer must take into account the biases inherent in those designs and make definitive recommendations about the effectiveness of a practice with caution.

Other types of evidence, including qualitative evidence and economic evidence addressing questions related to health care costs, can also be synthesized using methods established by organizations such as the Joanna Briggs Institute.¹² While the methods of synthesizing quantitative data are relatively straightforward and accepted, there are numerous methods for synthesizing qualitative research. Such reviews may appear as a meta-synthesis, a meta-aggregation, a meta-study, or a meta-ethnography^{7,16}; the differences between these approaches will be discussed in the fifth article in this series.

A systematic review that addresses both quantitative and qualitative studies, as well as theoretical literature, is referred to as an "integrative" or

“comprehensive” systematic review.^{6,15} The motivation for conducting a comprehensive review is often to provide further insight into why an intervention appears to have a benefit (or not). “Realist” reviews, another emerging form of evidence synthesis, often look to answer questions surrounding complex interventions, including how and for whom an intervention works.^{7,16} Formalized methods for these types of reviews are still being validated.

Reviewers must consider the quality of the studies when arriving at recommendations based on the results of those studies.

Interpretation of findings and recommendations to guide nursing practice. The conclusions of the systematic review, along with recommendations for clinical practice and implications for future research, should be based on its findings. Questions to ask when considering the recommendations of a systematic review include the following: Has a clear and accurate summary of findings been provided? Have specific directives for further research been proposed? Are the recommendations, both for practice and future research, supported by the data presented? (Such issues will be explored in the sixth and last paper in this series.)

Reviewers must consider the quality of the studies when arriving at recommendations based on the results of those studies. For example, if the best available evidence was of low quality or only observational studies were available to answer a question of effectiveness, results based on this evidence must be interpreted with caution.

Nurses are increasingly expected to make evidence-based decisions in their practice, and nursing researchers are increasingly driven to develop advanced methods of evidence synthesis. Systematic reviews aim to summarize the best available evidence using rigorous and transparent methods. We’ve provided a brief introduction to the steps taken in conducting a systematic review; the remaining papers in this series will explore each step in greater detail, addressing the synthesis of both quantitative and qualitative evidence. ▼

Edoardo Aromataris is the director of synthesis science at the Joanna Briggs Institute in the School of Translational Health Science, University of Adelaide, South Australia. Alan Pearson is the former executive director and founder of the Joanna Briggs Institute. Contact author: Edoardo Aromataris, ed.aromataris@adelaide.edu.au. The authors have disclosed no potential conflicts of interest, financial or otherwise.

The Joanna Briggs Institute aims to inform health care decision making globally through the use of research evidence. It has developed innovative methods for appraising and synthesizing evidence; facilitating the transfer of evidence to health systems, health care professionals, and consumers; and creating tools to evaluate the impact of research on outcomes. For more on the institute’s approach to weighing the evidence for practice, go to <http://joannabriggs.org/jbi-approach.html>.

REFERENCES

1. Krainovich-Miller B. Literature review. In: LoBiondo-Wood G, Haber J, eds. *Nursing research: methods and critical appraisal for evidence-based practice*. 6th ed. St. Louis: Mosby Elsevier; 2006.
2. Eger M, et al. Rationale, potentials, and promise of systematic reviews. In: Egger M, et al., eds. *Systematic reviews in health care: meta-analysis in context*. 2nd ed. London: BMJ Publishing Group; 2001. p. 3-19.
3. Averis A, Pearson A. Filling the gaps: identifying nursing research priorities through the analysis of completed systematic reviews. *JBI Reports* 2003;1(3):49-126.
4. Ubbink DT, et al. Framework of policy recommendations for implementation of evidence-based practice: a systematic scoping review. *BMJ Open* 2013;3(1).
5. Joanna Briggs Institute. An introduction to systematic reviews. *Changing practice: evidence based practice information sheets for health professionals*. 2001;5(Suppl 1): 1-6.
6. Pearson A, et al. The JBI model of evidence-based health-care. *Int J Evid Based Healthc* 2005;3(8):207-15.
7. Tricco AC, et al. The art and science of knowledge synthesis. *J Clin Epidemiol* 2011;64(1):11-20.
8. Khan KS, et al. Five steps to conducting a systematic review. *J R Soc Med* 2003;96(3):118-21.
9. Green S, et al. Introduction. In: Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions*. Chichester, West Sussex; Hoboken, NJ: Wiley-Blackwell; 2008.
10. Moher D, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;151(4):264-9, W64.
11. Stone PW. Popping the (PICO) question in research and evidence-based practice. *Appl Nurs Res* 2002;15(3): 197-8.
12. Joanna Briggs Institute. *Joanna Briggs Institute reviewers’ manual: 2011 edition*. Adelaide, South Australia: University of Adelaide; 2011. <http://joannabriggs.org/assets/docs/sumari/ReviewersManual-2011.pdf>.
13. Mares MA, McNally S. The effectiveness of nurse-led cardiac rehabilitation programs following coronary artery bypass graft surgery: a systematic review protocol. *JBI Database of Systematic Reviews and Implementation Reports*. 2013;11(11):21-32.
14. Crowther M, et al. Systematic review and meta-analysis methodology. *Blood* 2010;116(17):3140-6.
15. Whittlemore R, Knafl K. The integrative review: updated methodology. *J Adv Nurs* 2005;52(5):546-53.
16. Pawson R, et al. Realist review—a new method of systematic review designed for complex policy interventions. *J Health Serv Res Policy* 2005;10 Suppl 1:21-34.