Systematic review protocols: an introduction

Lisa S Whiting looks at the steps necessary to create a review protocol for appraising literature

Abstract
Systematic reviews are a thorough and efficient method of appraising literature and providing an evidence base for practice. However, to ensure that they are undertaken in a structured and comprehensive manner, it is crucial that a systematic review protocol is first formulated. While resources such as the Cochrane Collaboration provide extensive and valuable information in relation to the issues that should be considered, there is little immediately available in the nursing literature.

key words
- systematic reviews
- qualitative research
- protocols
- literature search

Despite the recent widespread inclusion of qualitative research studies, the majority of systematic reviews rely heavily on the aggregation of quantitative studies – in particular, randomised controlled trials (RCTs). Since nurses may be involved in the writing, reading and appraising of systematic reviews and associated protocols, it is imperative that they understand the necessary components. This paper will identify the key sections of a systematic review protocol that uses RCTs.

Mulrow et al (1997) state that ‘systematic reviews use explicit and rigorous methods to identify, critically appraise, and synthesize studies’ and ‘seek to assemble and examine all of the available high-quality evidence that bears on the clinical question at hand’. Systematic reviews use a structured and reproducible approach to appraise research for consistency, generalisability and precision (Evans 2001, Kitchenham 2004, Higgins and Green 2006). Given their importance, clear guidance on their development has been provided
Despite their advantages (Box 1), they are not without problems: they need to be conducted thoroughly (Davies and Crombie 2001a); the results may not echo the results from a large single trial (Egger and Davey Smith 1995, LeLorier et al 1997), and there can be unsuitable ‘aggregation’ of research studies that influence the overall findings (Davies and Crombie 2001a).

**Box 1. Advantages of systematic reviews**
- Explicit methods limit bias in identifying and rejecting studies.
- Conclusions are more reliable and accurate because of methods used.
- Large amounts of information can be assimilated.
- Delay between research discoveries and implementation of effective diagnostic and therapeutic strategies may be reduced.
- Results can be formally compared to establish generalisability of findings and consistency (lack of heterogeneity) of results.
- Reasons for heterogeneity (inconsistency in results across studies) can be identified and new hypotheses generated about particular subgroups.
- Quantitative systematic reviews (meta-analyses) increase the precision of the overall result.
  (Chalmers and Altman 1995)

**Developing a review protocol**
To plan and structure a systematic review, a detailed protocol needs to be prepared (Evans 2001, Kitchenham 2004). This comprises information concerning the review question, retrieval, appraisal and synthesis of studies (Evans 2001). Clear details about the review process enhance transparency and objectivity, and help to minimise bias (Khan et al 2001). The protocol development stages are identified in Box 2.

**Box 2. Components of a review protocol from Khan et al (2001)**
- Background.
- Review questions.
- Search strategy including terms and resources to be searched.
- Study selection criteria and procedures.
- Study quality assessment checklists and procedures.
- Data extraction strategy.
- Synthesis of the extracted evidence.
- Project timetable.
Background
This should provide an insight into the selected topic area and the need for a systematic review of the literature.

Review questions
Systematic reviews are costly and labour intensive, so duplicate ones should be avoided (Cook et al 1997). It may be unethical to write a protocol for a review that already exists, so databases (Box 3) should be searched to avoid this.

A clear, well-focused question is key to a systematic review and should be clearly identified in the protocol (Counsell 1998). Higgins and Green (2006) suggest that this provides structure, leads the whole review process and provides the reader with insight into the chosen topic. It may be written as a specific question, or as an overall aim.

Review questions have key components (Oxman et al 1994, Counsell 1998):

- **Population**: the type of person involved.
- **Interventions**: the type of exposure.
- **Comparison**: the type of control with which the exposure is being compared.
- **Outcomes**: the factors that will be used to compare the interventions.

Higgins and Green (2006) urge clear identification of the population under consideration, clarifying any restrictions that may be imposed. Authors should also consider the type of interventions that are of interest, and the environment in which they take place.

To identify appropriate RCTs for the systematic review, it is crucial to state the type of control (or comparison method) that will be examined (Counsell 1998).

Box 3. Example of review databases

- **The Cochrane Collaboration**: [www.cochrane.org](http://www.cochrane.org)
- **Centre for Evidence-Based Medicine at Oxford**: [www.cebm.net](http://www.cebm.net)
- **The Centre for Reviews and Dissemination (CRD)**: [www.york.ac.uk/inst/crd](http://www.york.ac.uk/inst/crd)
- **The Campbell Collaboration**: [www.campbellcollaboration.org](http://www.campbellcollaboration.org)
- **Systematic Reviews Training Unit**: [http://tinyurl.com/ycvny39](http://tinyurl.com/ycvny39)
- **Bandolier**: [www.medicine.ox.ac.uk/bandolier](http://www.medicine.ox.ac.uk/bandolier)
- **The EPPI-Centre**: [http://eppi.ioe.ac.uk](http://eppi.ioe.ac.uk)
Examples include:

- One intervention with no intervention at all: for example, is one form of treatment preferable to none at all?
- Interventions alternatable with one another: for example, is one form of treatment preferable to another?
- Different interventions experienced by the same population/sample group: for example, does the population prefer one treatment to another?

Outcomes are identified to assist with comparing the effectiveness of interventions (Kitchenham 2004). They should be relevant to the population (Eddy 1990) and to practitioners (Kitchenham 2004).

**Search strategy**

To address the identified aim and objectives, a comprehensive literature search will retrieve relevant studies (Parahoo 2006). Search strategies should be included in the protocol. Forward and Hobby (2002) suggest that a systematic review is the ‘gold’ standard of literature-based research, allowing objective assessment of research evidence.

A number of authors have advocated the use of computerised databases (Fink 1998, Hek et al 2000, Playle 2000), but it is acknowledged that no single database records publications from all health journals, (Tait and Slater 1999, Suarez-Almazor et al 2000). Literature from all languages should be sought to maximise data retrieval and minimises bias (Egger et al 1997a), but time and financial restrictions may not allow this (Meade and Richardson 1997).

Key search terms are beneficial, although due to the complex nature of database indexing, literature could be missed (Sindhu and Dickson 1997). Hence, a range of search strategies is imperative. Manual searching, although time-consuming and labour intensive, may retrieve literature not already identified and can highlight alternative synonyms for database searching.

Bibliographies and reference lists of all studies accessed should be scanned for further research. Greenhalgh and Peacock (2005) found that only 30 per cent of their sources were retrieved via the strategies identified in their protocol; the remainder of their material was acquired using a ‘snowballing’ technique of seeking references used by the research (51 per cent) and by personal contact (24 per cent).
Bowling (2002) suggests that grey literature – publications issued by government, academia and industry, uncontrolled by commercial publishing interests – should also be examined (Weintraub 2000). These are particularly difficult to retrieve (Hek 1996), but there are several relevant websites and databases that can help (Box 4).

**Box 4. Databases and websites that can help retrieval grey literature**

- System for Information Retrieval on Grey Literature (SIGLE).
- Key websites, such as the Department of Health.
- Bibliographic searching – unpublished research may be referenced and the work retrieved via library services or author contact.
- Search of theses and dissertations in accessible libraries and via CINAHL.
- Search engines, such as Google and Google Scholar.

Researchers need to consider ‘publication bias’: not all research studies are readily available due to difficulty in retrieval or language issues (Dickersin 1990). In addition, as studies are more likely to be published if findings demonstrate positive significant effects of treatment (Streiner and Norman 2003), those that do not – and therefore remain unpublished – can hold valuable information about the limited effect of an intervention.

**Study selection criteria and procedures**

The protocol should identify the inclusion criteria for study selection to ensure that appropriate and reliable data are provided and the research fulfils the areas stipulated (Khan et al 2003). To minimise the risk of relevant papers being discarded, more than one person should assesses the suitability of the study for inclusion (Edwards et al 2002).

If after scrutinising the study title and abstract, it is unclear whether the paper meets the inclusion criteria or not, the full text should be read. If studies are excluded, a rationale should be provided (Higgins and Green 2006).

Once the relevant papers have been identified, they need to be managed (Higgins and Green 2006) using, for example, a computer assisted package such as EndNote.
Study quality assessment, checklists and procedures
Studies selected should be assessed for methodological robustness (Sanna and Aro 2005), by two reviewers to prevent ‘false positives’ (Meade and Richardson 1997). The assessment process should be stated in the protocol (Higgins and Green 2006). A number of tools exist to help reviewers (Moher et al 1995).

Data extraction strategy
‘Data extraction is the process by which the information needed for data synthesis is obtained’ (Khan et al 2003). Thus, the intended data extraction form should be stated in the protocol (Higgins and Green 2008) (Box 5).

Box 5. Key components of data extraction forms

**General information**
- Date of data extraction.
- Title, authors, journal, publication details, or any other identifying features of the study.
- Identification of the reviewer.
- Notes.

**Specific information**

*Study characteristics*
- Re-verification of study eligibility.
- Population characteristics and care setting.
- Methodological quality of the study.
- Interventions.
- Outcomes.
- Notes.

*Outcome measures and results*
- Length of follow up.
- Drop-outs.
- Missing data.
- Discrete data (events, total numbers, $p$-value).
- Continuous data (mean, SE, SD, numbers, $p$-value).
- Survival data (observed and expected number of events, survival plots, $p$-value).
- Effect measures.
- Notes.

(Adapted from Khan and Kleijnen 2001)
Synthesis of the extracted evidence

The concluding stage of the systematic review is the summary of the data (Lau et al 1997). The proposed approach should be explained in the review protocol. A much-used approach following a review of RCTs is meta-analysis, which combines the results of studies (Sanna and Aro 2005) and ‘calculates a pooled estimate of data and its confidence interval’ (Deeks et al 2001). Of course, the appropriateness of using meta-analysis should be established; it is not, for example, suitable for studies without homogeneity (a degree of similarity).

Meta-analysis has a range of benefits. It can: increase power; use more data to enhance the estimated effect of treatment; address questions that may not have been asked by an individual study; and assess the differences in findings between studies (Higgins and Green 2006). Meta-analysis can be undertaken using appropriate software such as Comprehensive Meta-Analysis Version 2 (www.meta-analysis.com).

Deeks et al (2001) suggest that having identified the feasibility of meta-analysis, the researcher should consider and include in the protocol three questions:

■ What comparisons should be made?
■ Which outcome measures should be used?
■ Which effect measure should be used to describe the effectiveness?

How to measure effect

Measuring effect enables the reviewer to present the results of each study in a standard manner, which helps when comparing the studies. Effect measure is defined as: ‘The observed relationship between an intervention and an outcome… summarised as a p-value, odds ratio, relative risk, risk difference, number needed to treat, standardised mean difference, or weighted mean difference’ (Deeks et al 2001).

The protocol should explain the measurement method used; ‘Odds ratio’ is widely used in meta-analysis of systematic reviews of RCTs (Davies and Crombie 2001b). However, it is only appropriate if the studies have generated binary data; studies that have produced continuous data require a ‘mean difference’ (that is, the difference between the means of the two groups) measurement.
Approaches to meta-analysis
Having summarised the findings of each study, the results are then ‘pooled’ using one of two models; the random effects model or the fixed effect model (Deeks et al 2001). There is debate regarding which is most appropriate (Thompson and Pocock 1991); while some suggest the random effects model may be preferable (Hunter and Schmidt 2000), there is no ‘correct’ approach and there is usually only a marked difference in the combined effect of the studies between the two models if there is a large degree of heterogeneity (Egger et al 1997b).

Presentation of results
The protocol should outline how the findings of the systematic review will be presented. A narrative approach can be used to address key areas, diagrams for the meta-analysis results can demonstrate confidence intervals and outcome measures. One of the most commonly used is the ‘forest plot’ (Lewis and Clarke 2001), which enables the reader to assess the effect and confidence interval of each study at a glance (Lewis and Clarke 2001). Software will produce forest plots almost instantaneously.

Project timetable
Khan et al (2001) suggest that a timescale of between nine and twenty-four months is reasonable to conduct a systematic review. Arguably, this may depend on funding and release from other work commitments, but should be considered in the protocol.

Conclusion
While the focus of this article has been RCTs, it is important to stress that a range of other studies could be drawn on in a systematic review, particularly time-series analysis and controlled before-and-after studies. In addition, more recent consideration has been given to the inclusion of qualitative research, an area of systematic review that is developing rapidly. However, it is clear that whatever studies are used, a protocol is an essential pre-requisite to any systematic review; without it, there is a danger that valuable resources and time will be inappropriately used, potentially jeopardising health care.
In addition, nurses must be encouraged not to ‘shy’ away from evaluating quantitative research, it can be challenging, but with the assistance of appropriate software and advice, it is achievable.

Lisa Whiting MSc, BA(Hons), RGN, RSCN, RNT, LTCL is professional lead, children’s nursing, University of Hertfordshire, Hatfield, UK

This article has been subject to double-blind review

references


Nurse Researcher. 7, 3, 40-57.


Copyright of Nurse Researcher is the property of RCN Publishing Company and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder’s express written permission. However, users may print, download, or email articles for individual use.