International Journal of Nursing Education Scholarship

Volume 1, *Issue* 1 2004 *Article* 15

Systematic Reviews of Health Care Interventions: An Essential Component of Health Sciences Graduate Programs

Shelley Peacock* Dorothy Forbes†

*Saskatchewan Institute of Applied Science & Technology, shelliepeacock@netscape.net

†University of Saskatchewan, dforbes6@uwo.ca

Systematic Reviews of Health Care Interventions: An Essential Component of Health Sciences Graduate Programs*

Shelley Peacock and Dorothy Forbes

Abstract

Systematic reviews are an objective, rigorous assessment of both published and unpublished research that enable the reviewer to make recommendations to clinicians, policy-makers, consumers, and researchers. The steps in a systematic review include: (a) developing a research question, (b) developing relevance and validity tools, (c) conducting a thorough literature search of published and unpublished studies, (d) using relevance and validity tools to assess the studies, (e) completing data extraction for each study, (f) synthesizing the findings and, (g) writing the report. The purpose of this paper is to demonstrate the value of providing health science graduate students with the opportunity to learn about the conduct of a systematic review. An example of a thesis utilizing the method of a systematic review is presented.

KEYWORDS: systematic reviews, effectiveness, graduate program

---

*The authors wish to acknowledge the assistance of librarians Janet Bangma and Erin Watson, HSL, University of Saskatchewan, Saskatoon, Canada. Partial funding is acknowledged from the College of Graduate Studies & Reasearch, University of Saskatchewan, Saskatoon, Canada.

In light of the enormous amount of literature available and the barriers to accessing journals (e.g., lack of time, resources or appraisal skills) systematic reviews are invaluable to aid policy-makers, clinicians and consumers in making health care decisions (Ciliska, Hayward, Dobbins, Brunton, & Underwood, 1999). Often nurses and other health care professionals do not have adequate time or skills to search and assess available research in a particular area. Systematic reviews can help manage the large amount of information available by synthesizing data from primary studies and completing a summary of the effectiveness of interventions (Forbes & Clark, 2003).

Systematic reviews are a unique way to conduct research. Carrying out a systematic review is comparable to conducting a research study (Forbes, 2003). The unit of analysis is the primary difference (Moher, Jadad, & Klassen, 1998); research reports or articles are the unit of analysis instead of participants or subjects (Forbes). The method of appraisal and synthesis in a systematic review are explicitly described for readers, unlike other reviews (e.g., narrative reviews) (Evans, 2001; Jones, 1994; Klassen, Jadad, & Moher, 1998). A valuable source of reliable systematic reviews is The Cochrane Library. It includes nearly two thousand systematic reviews, in addition to the Database of Abstracts of Reviews of Effectiveness (DARE).

Several graduate programs offer the opportunity to learn how to conduct systematic reviews. The following is a brief discussion of the method of systematic reviews with an example that was conducted to complete the requirements for a Master's in Nursing degree. This review will be referred to as the *caregiver review*. The purpose of this paper is to demonstrate the value of providing graduate students in nursing the opportunity to learn about the conduct of a systematic review.

METHOD AND DISCUSSION

The steps of a systematic review outlined in this paper are intended to inform graduate students and their supervisors who may be interested in conducting a systematic review as part of the requirement for their graduate program. The framework that guided the conduct of the caregiver review was based on the work of Forbes and Strang (1997) and Forbes (1998). Conclusions about evidence are attained by assessing studies using defined steps. Forbes (2003) identifies the steps of a systematic review as follows: (a) developing a research question, (b) developing relevance and validity tools, (c) conducting a thorough literature search to include both published and unpublished studies, (d)

using relevance and validity tools to assess the studies, (e) extracting the data, (f) synthesizing the findings, and (g) writing the report. Each is discussed and the caregiver review is used to illustrate the step.

Developing the Research Question

Formulating an appropriate question will drive how the review is carried out. The question should be meaningful and relevant to an area that will impact patient/client care and/or outcomes. Ideally, focus groups comprised of consumers and/or health care providers should be conducted to identify what questions are important to them (Forbes & Phillipchuk, 2001). A good research question should include a description of the population, intervention, and outcome (Forbes, 2003).

The question for examination in the caregiver review was a result of personal clinical experience. Prior study and literature searches in the areas of family caregivers and dementia resulted in refinement of the research question. The research question for the caregiver review was 'What interventions are effective in supporting unpaid caregivers' well-being when caring for elderly persons with dementia in the community?' The population was identified as caregivers of elderly individuals with dementia in the community; the interventions included any approaches to support carers; and the outcomes were attributes of well-being. The question is broad in order to gather a sufficient number of research studies. The population is limited to those who care for someone with dementia in the community and does not include caregivers of persons with other conditions or diseases.

Developing Relevance and Validity Tools

Relevance tool. The relevance tool is essential to screen for studies that will ultimately be included in the review. All potentially eligible studies should be assessed (Clarke & Oxman, 2000). The criteria included in the relevance tool should evolve from the research question and include the population, the intervention(s), and outcome(s) as well as the study design(s) that will best answer the question (Forbes, 2003). Clarke and Oxman suggest that two reviewers independently assess studies for relevance as some reviewers may have pre-formed opinions in the area under consideration.

The relevance tool for the caregiver review was modified from the work of Forbes and Strang (1997). The relevance criteria for the caregiver review were used to determine if the study: (a) was conducted or published in 1992 or later; (b)

DOI: 10.2202/1548-923X.1042

evaluated an intervention directed at caregivers of an elderly individual with dementia living in the community; (c) measured one of the following caregiver outcomes: well-being (includes: physical, mental, social, and financial consequences), depression, strain, and/or other (e.g., institutionalization, health care expenditures); and (d) incorporated a control group or a pretest-posttest design with a sample size greater than one. Both authors rated studies for relevance. When all four of the relevance criteria were met the study was then included in the validity appraisal.

The first criterion included both published and unpublished studies in an attempt to reduce publication bias. The second criterion identified if the study assessed an intervention for caregivers of persons with dementia. This eliminated studies that did not report data specific to caregivers of persons with dementia. The third criterion identified the outcomes of the studies to be included in the review. This criterion was broad so as to include as many studies as possible that addressed the salient outcomes for caregivers of persons with dementia. The fourth and final criterion addressed the types of study designs that were to be included. Descriptive case studies were not considered; rather studies needed to compare and assess the effectiveness of the intervention under examination. The review was not limited to randomized controlled trials (RCTs), as this could have excluded studies that may assess effectiveness utilizing an alternate research design, for example a pretest-posttest design. The relevance tool was pre-tested and revised to meet the purposes of the caregiver review.

Validity tool. The validity tool is necessary to assess the quality of included studies, limit bias in the systematic review, and guide interpretation of findings (Clarke & Oxman, 2000). Validity is assessed by considering potential sources of bias and error within a study (Forbes, 2003). Potential sources of bias in healthcare studies include: selection bias, performance bias, attrition bias, and detection bias (Clarke & Oxman). Selection bias may be avoided with true randomization of participants to groups or controlling for relevant confounders; performance bias is reduced by blinding both participants and investigators; attrition bias is a result of withdrawal of subjects from a study; and finally detection bias may be avoided by blinding of data collectors completing outcome measures (Clarke & Oxman; Forbes).

The criteria of the validity tool in the caregiver review attempted to address these potential biases. The first criterion addressed the type of design and process of allocation to the intervention or control groups as a means to assess selection bias (with randomization being the gold standard design). However, the

caregiver review was not limited to RCTs, as other methods are often used in nursing research (e.g., pretest-posttest, cohort designs) because the randomization of subjects is not always possible for ethical reasons. The second criterion was related to attrition and assessed the rate of withdrawal over the entire length of the study, which varied greatly. The third criterion considered the control for potential confounders in an effort to limit selection bias, particularly in designs other than RCTs. The fourth criterion assessed detection bias, that is, how accurately were the data measured and collected (e.g., were instruments pre-tested and data collectors blinded?). Lastly, the level of statistical analysis was assessed as studies that employed multivariate statistics could control for confounders, thus reducing bias. The validity tool was used to rate studies on the extent to which each source of systematic bias was addressed. To limit rater bias, two reviewers independently assessed relevant studies for validity. The validity tool, rating scale and dictionary were pre-tested and revised as necessary.

Literature Search Strategies

Literature searches for systematic reviews must be as comprehensive as possible (Evans, 2001). This process is one of the rigorous steps that separate systematic reviews from traditional narrative reviews. Both published and unpublished research studies should be sought. A systematic review, unlike a narrative review, attempts to include all of the relevant literature in the area of investigation. Unpublished research is vital as it may systematically differ from what is published, possibly because the results are non-significant and not due to methodological rigor (Forbes, 2003).

Utilizing the assistance of experienced librarians for literature searching is a must. Their expertise will make searches of various databases relevant and thorough. A consistent appropriate list of subject terms and key words should be used with as many relevant databases as possible. Sufficient detail about the search strategy should be reported to allow for replication (Counsell, 1997). As well, reference lists of all retrieved studies should be reviewed for further studies.

The search strategies for the caregiver review included on-line computer searches, hand searches of selected relevant journals, and searches of individual reference lists of all retrieved studies. On-line searches of CINAHL, PubMed, and PsycInfo were conducted for the years between 1992 to April 2002 with the assistance of an experienced librarian. The key words that were utilized included caregiver or carer, dementia or Alzheimer, burden, depression, strain, stress, support, respite, education, intervention, effective, assess, evaluate, and measure.

DOI: 10.2202/1548-923X.1042

All inter-library loan requests were received. On-line CISTI Source was used to aid in hand searching the table of contents for the *Gerontologist, Journal of Gerontological Nursing,* and *Journal of the American Geriatrics Society.* Retrieved studies' reference lists were also searched for relevant studies, which were retrieved and subsequently reviewed.

Unpublished dissertations were included in the librarian searches; none were relevant to the caregiver review. Searches completed by the librarian were catalogued using the reference management program of EndNote (Institute for Scientific Information, 2000). The primary authors of some articles were contacted for clarification of their studies and for additional information.

Assessment of Studies Utilizing Relevance and Validity Tools

All retrieved studies should be assessed for relevance to determine if they are to be included in the review. It may be important to have two raters that have varying knowledge levels in the subject area and methodology (Forbes, 2003). Two raters should independently evaluate studies for relevance and subsequent validity in order to reduce bias as much as possible (Forbes). A sample of 20 studies should be reviewed and level of agreement determined using a kappa rating. If the level of agreement is high (>.8) subsequent studies may be reviewed by only one rater.

Relevance tool. Of the 92 retrieved studies, 36 met all four relevance criteria. The first 19 studies were reviewed by both authors. A high level of agreement by the two readers was reached (kappa=.8); the remaining studies were assessed independently by the primary author, with any subsequent concerns discussed and a consensus reached between the two raters.

Validity tool. The next phase of the review involved rating the 36 relevant studies for validity. The first 12 relevant studies were rated independently again by two readers and 100% agreement was reached. Subsequent studies were rated by the primary author, with any concerns discussed and consensus reached. Of the 36 studies reviewed, 11 were judged to be strong, 11 were moderate, 13 were weak, and only 1 was judged poor. Descriptive analyses were completed for the 11 strong and 11 moderate studies in the areas of: methodological weaknesses, country in which study was conducted, interventions, outcomes, and study design.

Data Extraction

Consistent, uniform data extraction is required to obtain essential information about the studies (Forbes, 2003). To minimize bias in extracting information from the articles, a data extraction tool should be developed to include general information and specific study characteristics as reported by the researchers (e.g., the country in which the study was conducted; methods employed; participant characteristics; type, intensity, and duration of the intervention; outcomes measured; and instruments used to assess caregivers and care receivers).

The data extraction tool for the caregiver review was modified from the one developed by Forbes and Strang (1997). It was pre-tested and revised appropriately to reflect the changes in the validity tool. The tool was used to extract significant data from the strong and moderate studies. It was completed independently by the primary author, with concerns discussed with the second author until consensus was reached.

Many of the studies included in the caregiver review employed randomization of participants to either an intervention or control group (*n*=15); the remainder utilized a pretest-posttest design (*n*=7). The majority of studies were conducted in the United States (*n*=15), one in Finland, while the remaining were in Canada, United Kingdom/Ireland and Australia (each *n*=2). The studies were categorized according to type of intervention. The most common intervention was education (*n*=8), followed by case management (*n*=4), psychotherapy (*n*=3), respite (*n*=3), technology (*n*=2), assessment clinic (*n*=1), and home care (*n*=1). The most common measured outcomes were found to be depression in caregivers (*n*=9), institutionalization of care receivers (*n*=8), caregiver strain (*n*=7), caregiver coping or appraisal of situation (*n*=7), stress or anxiety (*n*=6), quality of life or health of caregiver (*n*=6), death of care receiver (*n*=4), use of formal services (*n*=4), and caregiver knowledge of dementia (*n*=3).

Data Synthesis

Data synthesis is utilized to summarize findings in an informative, concise manner. A comprehensive means to synthesize data is the use of summary tables and graphs that may include information collected about the characteristics of participants, interventions, outcomes, and study quality. If appropriate (i.e., the subjects, interventions, and outcome measures are homogeneous), statistical analysis (i.e., meta-analysis) can be used to pool data (Forbes, 2003).

DOI: 10.2202/1548-923X.1042

The caregiver review used a descriptive synthesis to summarize findings. Non-significant findings were most common. Significant findings to note include, respite offered relief to caregivers while the care receiver was out of the home (e.g., Larkin & Hopcroft, 1993); case management increased the likelihood for caregivers to use community services (e.g., Newcomer, Spitalny, Fox, & Yordi, 1999); and educational interventions were found to increase caregiver knowledge of dementia (e.g., Coen, O'Boyle, & Lawlor, 1998). Additional discussion, description, and summary of the included studies, can be found elsewhere (Peacock, 2003; Peacock & Forbes, 2003).

Meta-analysis is a statistical method that combines and synthesizes results of separate, similar studies, which can be used to evaluate the effectiveness of interventions (Jones, 1994; Moher et al., 1998). Due to the heterogeneity of the interventions and instruments, meta-analysis was not appropriate in the caregiver review.

Report-Writing and Dissemination

Upon completing a review, researchers have an obligation to share and present the review results with as many interested parties as possible, particularly to groups or individuals that aided in the formulation of the research question. This information must be presented in a way that is appropriate to their level of understanding.

The caregiver review was reviewed by a consumer (an unpaid caregiver of a parent with Alzheimer Disease) for clarity and meaningfulness. The caregiver review has also been disseminated in a variety of ways since completion of the study. The process of a systematic review and the findings were presented to peers in a graduate seminar. An oral presentation and a summary fact sheet were completed for the provincial Alzheimer Society whose members include consumers. A presentation was made at a scientific conference (Canadian Association on Gerontology); and study findings were published (Peacock & Forbes, 2003).

LIMITATIONS

Some may view the inclusion of non-RCTs in a systematic review as a limitation. Historically, systematic reviews have been primarily conducted and used by physicians with RCTs as the gold standard (Forbes, 2003). Nurses,

however, are beginning to utilize systematic reviews as a means to answer their questions related to the effectiveness of interventions and to guide their practice (Forbes & Clark, 2003). There is a continuing debate as to what constitutes quality research (i.e., should non-RCTs be included in a systematic review?). Most nurses would argue that other rigorously designed studies inform our practice (Evans & Pearson, 2001).

The caregiver review is limited by the articles retrieved. Due to cost, the European database Embase, was not searched. This database may have provided abstracts of additional foreign articles other than the ones included in the review. The databases that were searched, however, are considered thorough and reliable. As well, research may have been performed in this area but not published, limiting access to those findings; no researchers that were contacted shared information of other possible work in progress. Publication bias (i.e., not publishing studies that revealed only non-significant results) may also skew the findings of a review.

Carrying out the caregiver review was difficult due to the heterogeneous nature of the interventions. Yet, to single out a particular intervention would have resulted in a limited review. The variety of interventions and use of multiple instruments for measuring similar variables made it impossible to pool findings in order to conduct a meta-analysis. The findings of the caregiver review must be considered in light of the methodological limitations found in the included studies and in the conduct of the review.

IMPLICATIONS FOR NURSING EDUCATION

Conducting a systematic review has been a meaningful approach to partially fulfill the thesis requirement for a graduate degree. Instead of conducting an original research study that used a specific research design, the systematic review provided exposure to a variety of research methods and statistical analyses. The reviewer must become knowledgeable about these approaches and analyses in order to be able to critique the research study. As well, by reviewing all of the available research in an area, the reviewer becomes well-versed in the state of the science and research gaps in the topic area. This knowledge provides a solid foundation on which to build practice guidelines and/or further research studies.

The learning process of independently assessing studies, then coming together with the second reviewer to discuss concerns and discrepancies was invaluable. What constitutes rigorous research becomes abundantly clear. For example, how the researcher(s) addressed several types of bias (selection,

DOI: 10.2202/1548-923X.1042

performance, attrition, and detection) was evaluated in each study. Becoming familiar with the qualities that constitute rigorous research enhances effective critiquing of studies. Learning to effectively critique studies is a useful and worthwhile skill that all healthcare practitioners and decision-makers should acquire to ensure that their decisions and interventions are based on strong research evidence.

A challenge faced in conducting this review was largely in the area of reading and understanding published work that was not well-written. Not all primary researchers returned requests for clarification of findings or for additional information required in the review, which limited the comprehensiveness of the review. Through the process of conducting the systematic review the authors became acutely aware of what types of information are required in a report to adequately inform the readers. Health care practitioners, policy-makers and researchers need particular information about the studies to be able to make a decision about the usefulness of the research for their particular situation.

CONCLUSION

Systematic reviews are a valuable form of research that answer questions relevant to a variety of disciplines. They are an appropriate and useful method of research to undertake during a graduate program. The steps of a systematic review have been outlined to inform graduate students and their supervisors who are interested in conducting a systematic review. Since systematic reviews are rigorous assessments and synthesis of research studies that limit bias as much as possible, they are useful in guiding clinicians, policy-makers and consumers in making decisions about the effectiveness of interventions. Conducting a systematic review during a graduate program provides the opportunity to learn what constitutes a rigorous research study, to become familiar with the state of the science and research gaps in a topic area, and provides the foundation on which to build practice guidelines and/or further research.

REFERENCES

Ciliska, D., Hayward, S., Dobbins, M., Brunton, G., & Underwood, J. (1999). Transferring public-health nursing research to health-system planning: Assessing the relevance and accessibility of systematic reviews. *Canadian Journal of Nursing Research, 31*(1), 23-36.

Clarke, M., & Oxman, A. D. (2000). Cochrane reviewers' handbook 4.1 [updated June 2000]. In *Review Manager* (RevMan) [Computer Program]. Version 4.1. Oxford, England: The Cochrane Collaboration.

Coen, R., O'Boyle, C., & Lawlor, B. (1998). Dementia care education and patient behaviour disturbance. *International Journal of Geriatric Psychiatry, 14,* 302-306.

Counsell, C. (1997). Formulating questions and locating primary studies for inclusion in systematic reviews. *Annals of Internal Medicine, 127,* 380-387.

Evans, D. (2001). Systematic reviews of nursing research. *Intensive and Critical Care Nursing, 17,* 51-57.

Evans, D., & Pearson, A. (2001). Systematic reviews: Gatekeepers of nursing knowledge. *Journal of Clinical Nursing, 10,* 593-599.

Forbes, D. A. (1998). Strategies for managing behavioral symptomology associated with dementia of the Alzheimer type: A systematic overview. *Canadian Journal of Nursing Research, 30*(2), 67-86.

Forbes, D. A. (2003). An example of the use of systematic reviews to answer an effectiveness question. *Western Journal of Nursing Research, 25*(2), 179-192.

Forbes, D. A., & Clark, K. (2003). The Cochrane library can answer your nursing care effectiveness questions. *Canadian Journal of Nursing Research, 35*(3), 18-25.

Forbes, D. A., & Phillipchuk, D. (2001). The dissemination and use of nursing research. *Canadian Nurse, 97*(7), 18-22.

Forbes, D. A., & Strang, V. R. (1997). *Strategies to manage the behavioral symptomology associated with SDAT: A systematic overview*. Edmonton, AB: Alberta Association of Registered Nurses.

Institute for Scientific Information: Researchsoft. (2000). Endnote for students: Bibliographies made easy [Computer Program]. Berkeley, CA: Authors.

Jones, A. (1994). An introduction to meta-analysis. *Respiratory Care, 39,* 34-49.

Klassen, T., Jadad, A., & Moher, D. (1998). Guides for reading and interpreting systematic reviews: Getting started. *Archives of Pediatric Adolescent Medicine, 152,* 700-704.

DOI: 10.2202/1548-923X.1042

Larkin, J., & Hopcroft, B. (1993). In-hospital respite as a moderator of caregiver stress. *Health and Social Work, 18,* 132-138.

Moher, D., Jadad, A., & Klassen, T. (1998). Guides for reading and interpreting systematic reviews: How did the authors synthesize the data and make their conclusions? *Archives of Pediatric Adolescent Medicine, 152*, 915-920.

Newcomer, R., Spitalny, M., Fox, P., & Yordi, C. (1999). Effects of the medicare Alzheimer's disease demonstration on the use of community-based services. *Health Services Research, 34*, 645-667.

Peacock, S. C. (2003). *A systematic review of interventions for unpaid caregivers of persons with dementia.* Unpublished master's thesis, University of Saskatchewan, Saskatoon, SK.

Peacock, S. C., & Forbes, D. A. (2003). Interventions for caregivers of persons with dementia: A systematic review. *Canadian Journal of Nursing Research*, *35*(4), 88-107.