Systematic review – a method for nursing research

This article by Arja Holopainen, Tuovi Hakulinen-Viitanen and Kerttu Tossavainen explains the five stages of the systematic review process and describes how this method was applied to an analysis of studies dealing with nurse ‘teacherhood’. The authors argue that systematic review is an excellent method for summarising research knowledge and for highlighting evidence significant for nursing, nursing education and nursing research.

Introduction

The continuous growth of research knowledge and the demands for evidence-based practice have created a need to gather, analyse and synthesise previous research knowledge (Evans and Pearson 2001, Magarey 2001, Evans 2002, Jones 2004). The method of reviewing previous research can be referred to as ‘research of research’ (Droogan and Song 1996, Droogan and Cullum 1998, Whittemore 2005). Various methods can be used in a research review, depending on, for example, what the aim is of gathering and evaluating existing research knowledge, what kind of studies are included in the review and how they are analysed. The methods of reviewing existing research knowledge are integrative review, meta-analysis, systematic review and methods that combine qualitative research (for example, metasummary, metasynthesis, formal grounded theory and metastudy). What these different methods have in common is that they generally follow the strict methodological demands applied to research (Whittemore 2005).
In a systematic review, the main focus is often on quantitative research because it is generally used, in addition to producing a meta-analysis, to gather the best possible research evidence to develop the evidence-based practice (Greener and Grimshaw 1996, Magarey 2001, Whittemore 2005). Systematic review can further be used to gather existing research knowledge regardless of the research methods used in the primary studies, in which case quantitative and qualitative studies can be included. In that case, too, either statistical or qualitative methods can be used as the methods of analysis (Evans 2001, Stevens 2001, Jones 2004, Duffy 2005).

This article concentrates on how to accomplish a systematic review. We conducted a systematic review concentrating on nurse ‘teacherhood’ studies to gather and evaluate research knowledge concerning the topic. The concept of nurse ‘teacherhood’ included nurse teachers’ tasks, their different multidimensional roles and also nurse teachers’ individual experiences of being a nurse teacher. In this article, that review will be used as an example to describe the process of carrying out a systematic review. This systematic review is also a part of the first author’s doctoral dissertation.

**Systematic review**

The existing research knowledge of a clearly defined topic is evaluated in a systematic review because the aim is usually to gather the best possible research evidence to develop evidence-based practice (Droogan and Cullum 1998, Evans 2001, Hewitt-Taylor 2002). Another aim of a systematic review can be to combine previous research results as a basis of scientific research (Conn et al 2003a) or to identify gaps in research knowledge (Hewitt-Taylor 2002). The purpose of the systematic review reported in this article was to gather previous research knowledge concerning nurse ‘teacherhood’ to use as the basis of an ongoing study of Finnish nurse teachers’ experiences of their nurse ‘teacherhood’ in polytechnics.

Systematic review helps to clarify how this topic has been examined and what kind of research evidence has been gained. It differs from a traditional literature review in that a more rigorous approach to research is applied. Special attention is focused on, for example, the exhaustiveness of the literature search and the analysis of the obtained material. In a traditional literature review, the
purpose is to formulate a general idea about the research results on the topic of interest. The purpose of the systematic review is to compile research knowledge on the topic as exhaustively as possible (Droogan and Song 1996, Hek et al 2000, Magarey 2001, Conn et al 2003a, Whittemore 2005).

**Stages of systematic review**
The stages of systematic review are: definition of the purpose of the systematic review and formulation of a research question; search of the literature; data evaluation and analysis; and presentation of the results. These stages should be documented so explicitly that the reader can estimate the realisation and validity of each stage (Droogan and Song 1996, Greener and Grimshaw 1996, Magarey 2001, Whittemore 2005). Systematic review is usually initiated by describing briefly the object of research (Lipp 2003). For example, in the present systematic review of nurse ‘teacherhood’, a short description of the topic and its contents was given.

**Purpose of the review and formulation of a research question**
In a systematic review, the purpose of the review and the research problem or question concerning the chosen topic are presented (Droogan and Song 1996, Magarey 2001, Whittemore 2005). In the example presented in this article, the purpose of the systematic review was to describe Finnish and international nursing science studies discussing nurse ‘teacherhood’ between January 1990 and April 2004. The following research questions were addressed:
- What were the topics of nurse ‘teacherhood’ studies during this period?
- How have the topics of nurse ‘teacherhood’ studies changed?
- What were the key results of the nurse ‘teacherhood’ studies conducted?

**Searching the literature**
The literature search is a time-consuming and difficult stage. It is, however, extremely important. If the literature search is deficient, or if inadequate databases have been used, the whole systematic review is defective and may give rise to faulty conclusions (Droogan and Song 1996, Magarey 2001, Whittemore 2005). Several methods are usually used when searching the literature. Published studies are searched by using electronic databases relevant
to the topic. The literature search is usually supplemented with a manual search of, for example, journals or research registers that specify the electronic search. There also exists ‘grey literature’: unpublished studies searched for especially when the best possible research evidence is gathered to develop evidence-based practice. The literature search can focus on, for example, the internet, conference proceedings or contacts with other investigators (Droogan and Song 1996, Greener and Grimshaw 1996, Magarey 2001, Conn et al 2003a, 2003b). The purpose of the systematic review determines how expansive a literature search is needed (Conn et al 2003a). When conducting a literature search, the researcher should bear in mind that not all published studies are necessarily of high quality, nor are all unpublished studies of poor quality (Hek et al 2000, Conn et al 2003b).

In addition to the electronic databases, the search terms used in the literature search, the search limits and the search results must be described. The inclusion criteria may contain the time limits of the review and limitations relating to the language and the type of article (Jones 2004, Whittemore 2005). The following electronic databases were searched for the systematic review discussing nurse ‘teacherhood’: Cinahl, Medline, Medic (Finnish Medicine and Health Sciences Database) and Arto (Reference Database of Finnish Articles). Table 1 presents the search terms, limits and results of the electronic search of studies discussing nurse ‘teacherhood’.

The literature search was verified manually by searching Finnish academic dissertations discussing nurse ‘teacherhood’ both from universities’ abstract books and websites. Similarly, articles discussing the topic were searched for manually by going through the issues of one Finnish journal (Journal of Nursing Science) and one Nordic journal (Nursing Science and Research in the Nordic Countries). The aim was to find the studies relevant to the topic that could not be found by using electronic databases. Altogether 2,524 studies were obtained by using electronic and manual searches.

Data evaluation and analysis
After the literature search, the suitability of the material is evaluated to find out how relevant it is to the research problems or questions. Pre-defined, specific inclusion and exclusion criteria are used to assist the evaluation process (Jones
Table 1. Electronic data search for studies discussing nurse ‘teacherhood’

<table>
<thead>
<tr>
<th>Databases</th>
<th>Search terms (brackets show the translation of the Finnish search)</th>
<th>Inclusion criteria</th>
<th>Number of hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cinahl</td>
<td>nurse teacher?.mp. &lt;br&gt; nurse educator?.mp. &lt;br&gt; nurse lecturer?.mp. &lt;br&gt; nurse faculty?.mp. &lt;br&gt; nurse tutor?.mp.</td>
<td>Publication year: 1990-2004 &lt;br&gt; Research &lt;br&gt; English language</td>
<td>1,628</td>
</tr>
<tr>
<td>Medline</td>
<td>nurse teacher?.mp. &lt;br&gt; nurse educator?.mp. &lt;br&gt; nurse lecturer?.mp. &lt;br&gt; nurse faculty?.mp. &lt;br&gt; nurse tutor?.mp.</td>
<td>Publication year: 1990-2004 &lt;br&gt; Review articles &lt;br&gt; English language</td>
<td>788</td>
</tr>
<tr>
<td>Medic</td>
<td>hoitotyön opettaj* (nurse teacher*) &lt;br&gt; terveysalan opettaj* &lt;br&gt; (teacher in health care*) &lt;br&gt; terveydenhuollon opettaj* &lt;br&gt; (teacher in health care*) &lt;br&gt; terveysiestiteiden opettaj* &lt;br&gt; (teacher in health care sciences*) &lt;br&gt; sairaanhoidon opettaj* &lt;br&gt; (nurse teacher*)</td>
<td>Publication year: 1990-2004</td>
<td>63</td>
</tr>
<tr>
<td>Arto</td>
<td>“hoitotyön opettaj?” &lt;br&gt; (nurse teacher*) &lt;br&gt; “terveysalan opettaj?” &lt;br&gt; (teacher in health care*) &lt;br&gt; “terveydenhuollon opettaj?” &lt;br&gt; (teacher in health care*) &lt;br&gt; “terveysiestiteiden opettaj?” &lt;br&gt; (teacher in health care sciences*) &lt;br&gt; “sairaanhoidon opettaj?” &lt;br&gt; (nurse teacher*)</td>
<td>Publication year: 1990-2004 &lt;br&gt; Word search (relevance)</td>
<td>18</td>
</tr>
</tbody>
</table>

mp=title, original title, abstract, name of substance, mesh subject heading  
* or ? = symbol showing word’s truncation aka ‘a wildcard’
The inclusion and exclusion criteria of the systematic review discussing nurse ‘teacherhood’ are listed in Table 2.

The strength of evidence of the primary studies is often used as an inclusion criterion in a systematic review conducted to develop evidence-based practice, since the aim is to include studies with a high level of evidence. Studies using qualitative research methods are then excluded from the review, mainly because the evidence produced by randomised controlled trials is considered more reliable than qualitative evidence (Droogan and Song 1996, Meade and Richardson 1997, Evans 2002). However, according to Evans and Pearson (2001), problems can occur if only studies compatible with experimental research designs are accepted in the development of evidence-based nursing. With the help of qualitative research, it is possible to gather information about such aspects as people’s impressions and experiences concerning their health or illness. Thus, evidence produced by qualitative studies is also needed in evidence-based nursing (Goding and Edwards 2002, Hewitt-Taylor 2002).
The systematic review discussing nurse ‘teacherhood’ presented in this article was conducted to summarise previous research knowledge concerning nurse ‘teacherhood’. The scholars wanted to get as exhaustive a picture as possible about what perspectives of nurse ‘teacherhood’ had been studied between January 1990 and April 2004. Since the aim was not to gather the best possible research evidence, all studies discussing the topic were included in the review regardless of the research methods used or the strength of evidence. Of the 207 studies included in the analysis, 92 were quantitative, 75 were qualitative and 40 were studies where both quantitative and qualitative methods were used.

After the literature search, the material is read through and evaluated to clarify whether the studies meet the criteria set for the material (Whittemore 2005). In the systematic review discussing nurse ‘teacherhood’, the studies obtained via the literature search \((n=2,524)\) were first evaluated according to the title and then according to the abstract to find out whether they met the pre-defined criteria of systematic review (Table 2). On the basis of the title, 2,155 articles were rejected where the title was sufficient to indicate that they did not meet the inclusion criteria (for example, the study did not concern nurse teachers). Of the 369 remaining articles, 161 were rejected after a closer examination of the abstract. Next, either the whole research article or the primary study was surveyed. At this point, 11 more studies were rejected. The reasons that led to the rejection of the studies based on the abstract or the whole article are presented in Table 3. The final material included 207 studies.

In the systematic review, the selected material is carefully read through and the results of the primary studies are condensed for further analysis (Cavanagh 1997, Hewitt-Taylor 2002, Whittemore 2005). In the analysis, the researcher organises, categorises and combines the material of the primary studies to find answers to the research problems or questions. The methods of analysis of the systematic review can be either statistical or qualitative, depending on the purpose and the material of the review. The analysis should meet the demands generally imposed on research analysis, and its different stages should be described in sufficient detail. It is recommended that two scholars conduct the analysis independently (Droogan and Song 1996, Greener and Grimshaw...
To increase the validity of the analysis, the scholars’ consensus concerning the results of the analysis can be calculated (Greener and Grimshaw 1996, Whittemore 2005). In the systematic review discussing nurse ‘teacherhood’, the results were tabulated for a condensed presentation, which included the researcher, the year and country of publication, the topic and purpose of research and the collection of research material, the sample and the main results (a total of 87 pages). Due to the purpose of the systematic review and the quality of the material, content analysis was used in the review. This method is highly suitable for the analysis of articles and other written material (Cavanagh 1997, Graneheim and Lundman 2004). Two scholars conducted the process of selection and analysis independently. The consensus concerning the selection of studies was 90 per cent. By using the table of the primary studies that the researchers had compiled, they divided the studies into three thematic categories – expansion of nurse ‘teacherhood’; skills of nurse ‘teacherhood’ and their development; and nurse ‘teacherhood’ and membership of a work-

<table>
<thead>
<tr>
<th>Reason for rejection</th>
<th>Studies n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The focus of the study was not on nurse ‘teacherhood’, but on</td>
<td></td>
</tr>
<tr>
<td>the learning and actions of a nursing student</td>
<td>35</td>
</tr>
<tr>
<td>the development of nurse education, teaching methods or curriculum</td>
<td>24</td>
</tr>
<tr>
<td>the actions of nurses supervising clinical training but mainly operating in nursing practice</td>
<td>30</td>
</tr>
<tr>
<td>the functioning of the administration or organisation of nurse education</td>
<td>5</td>
</tr>
<tr>
<td>nursing practice and its development</td>
<td>25</td>
</tr>
<tr>
<td>2. The research article was not available in Finland</td>
<td>29</td>
</tr>
<tr>
<td>3. Some other reason (such as not a study, study already included)</td>
<td>24</td>
</tr>
<tr>
<td>Totals</td>
<td>172</td>
</tr>
</tbody>
</table>

ing community – on the basis of their purpose and key results. Each of the studies was categorised according to only one thematic category. The analysis was resumed by using content analysis. First, the condensed meaning units were formulated by means of expressions describing the contents, which helped in the further categorisation of the material. Next, the analysis was continued by combining the condensed meaning units into subcategories and the subcategories into main categories. The categories were named based on content (Graneheim and Lundman 2004). A more detailed description and the condensed meaning units are published in another article (Holopainen et al 2007) describing the results of the analysis.

Presentation of results
The results of the systematic review can be presented as conclusions, analysis of the results or synthesis. Conclusions represent the results of the primary studies as different categories or themes. The analysis of the results includes the conclusions and their descriptions, as well as the criticism of the methods, the results and their applicability. The synthesis of the results is at a higher level of abstraction than the conclusions or the analysis, and it includes a new model or framework for the topic of interest of the systematic review (Whittemore 2005). In the example, the results were presented as a synthesis, because the results concerning the topics of nurse ‘teacherhood’ studies also included the changes in the topics and the results.

Validity of systematic review
The gathering and examination of previous research knowledge is a demanding process because the possibility of errors is significant in a process where the researcher has to make decisions and choices at every stage. For example, a literature search may fail if valid studies are rejected, or if it is conducted too concisely (Hek et al 2000, Duffy 2005, Whittemore 2005). This article describes the databases, search terms and search limits used in the literature search of the systematic review discussing nurse ‘teacherhood’ (Table 1), so that the reader can evaluate the validity of the literature search. The inclusion and exclusion criteria of the studies included in the systematic review are also presented (Table 2) to help the reader to see the grounds on which the studies
were included in the review. Validity is weakened by the fact that 29 studies were excluded from the material due to their limited availability.

The participation of several scholars in the selection and analysis of studies increases the validity of the systematic review (Droogan and Song 1996, Magarey 2001). In the present case, the selection and analysis of the material was conducted by two scholars. In unclear cases, both scholars familiarised themselves with the research, negotiating consent for the decision to include or exclude the study. Pre-defined inclusion criteria (Table 2) guided the selection of the studies. Similarly, in unclear cases at the stage of analysis, the researchers decided together on how to classify the studies. To evaluate and increase validity, each scholar also analysed ten studies that had already been analysed by the other. The consensus concerning the categorisation of the studies was 100 per cent. Validity was also increased by the fact that the scholars were previously familiar with content analysis (Whittemore 2005).

Generally, in a systematic review, the condensed results of the primary studies are also presented as a table to ensure the validity of the study (Droogan and Song 1996, Whittemore 2005). In the present case, due to the large number of primary studies \( n = 207 \), the table was too long (87 pages) to be included in the article reporting the results. For this reason, the validity of the results was increased by referring to some of the primary studies in connection with the results. Condensing the primary studies into tables helped the analysis because it enabled the management of the large material (Rutledge et al 2004). In unclear cases, the competency of the conclusions made in the analysis was verified from the primary studies to guarantee validity.

Generally, to increase the validity of a systematic review, the credibility of the results of the primary studies is also examined to evaluate their strength of evidence (Whittemore 2005). In our case, the validity of the results of the primary studies was not evaluated because the purpose of the systematic review discussing nurse ‘teacherhood’ was not to solve a clinical problem on the basis of research evidence. The purpose was to describe nursing science studies discussing nurse ‘teacherhood’ during a set period and to gather research knowledge concerning the topic regardless of the strength of evidence.
Discussion
Systematic review is a time-consuming process when carried out carefully. A sufficient amount of time has to be invested, especially in the literature search, because an exhaustive literature search and the retrieval of research articles is a surprisingly difficult process (Lipp 2003, Jones 2004). The systematic review discussing nurse ‘teacherhood’ took almost a year. The search of the literature for relevant research articles took a long time because a literature search requires diligence and searching the articles requires patience. Some articles were unreasonably difficult to obtain, which is why some of the articles that met the inclusion criteria (n=29) were excluded from the systematic review. Generally, systematic reviews are rarely successful in searching and obtaining all the relevant articles discussing the selected topic, although this goal should always be pursued (Evans 2001, 2002). The most important thing is to describe the literature search and the selection of material in sufficient detail.

Similarly, the analysis of the material proved very time-consuming. A thorough reading and analysis of extensive research material (207 studies) took two researchers more than six months. In addition, the recording of the results also took time because the material produced a great number of results.

Systematic review proved to be a demanding but interesting process. The review helped us to get a good impression of what aspects had been examined in the studies discussing nurse ‘teacherhood’ during the past 15 years. The research questions of this systematic review were intended to be extensive since the purpose was to gather all studies discussing nurse ‘teacherhood’ in the period January 1990 to April 2004.

Systematic review is often used in developing evidence-based practice to evaluate the research evidence of some clearly defined topic. It can also be used for more varied purposes when there is existing knowledge on a given research topic, such as nurse teacherhood, or when approaching the research topic as a target of development (Hemsley-Brown and Sharp 2003). Systematic review is an excellent method to summarise research knowledge and to present evidence for nursing, nursing education and nursing research.

Arja Holopainen PhD is a research director at the Nursing Research Foundation, Helsinki, Finland
Tuovi Hakulinen-Viitanen PhD, is a research director at the Development and Research Centre for Maternity and Child Health Care, Stakes – National Research and Development Centre for Welfare and Health, Helsinki, Finland

Kerttu Tossavainen PhD is a professor at the Department of Nursing Science, University of Kuopio, Kuopio, Finland

This article has been subject to double-blind review


