

How to conduct a systematic or evidence-based literature review

Article Sections

1. **Introduction**
2. [Defining search terms and identifying articles](#)
3. [Evaluation and synthesis](#)
4. [Summary](#)
5. [References](#)

By Margaret Adolphus

This article will describe the type of literature review which examines the literature on a subject with methods that are both explicit and transparent, and which follows a standard protocol or set series of stages.

The aim is to reduce bias and provide a comprehensive body of knowledge on a particular subject, and/or evidence for a particular intervention.

Such searches are variously termed *evidence-based*, *scoping*, or *systematic*.

As its name suggests, an evidence-based review seeks to comprehensively cover the field in order to find the best form of intervention. For example, what is the best treatment for an irregular heart beat (arrhythmia)?

The concept of evidence-based practice has its origins in, and is most common in, medicine, but is also used in other fields, particularly in nursing and allied health professions, psychology, educational research, and business and management.

A systematic review is described thus by Manning Fieden (2010, p. 386):

Systematic review offers a model for summarizing and critiquing the literature to improve future practice and possibly encourage higher levels of research methods. A systematic literature review of 30 years should reveal evidence toward a maturing research methodology.

The results of the systematic review are either summarized in a narrative, or, where there is much quantitative research, in statistical form.

A scoping review is very similar, and is a term used to provide a summary of a topic, conducted as a prelude to undertaking further research (see Hidalgo Landa et al. 2011).

A key aim of all these reviews is to avoid bias, which the researcher may unintentionally introduce through his or her own personal interest or preferences, or simply a lack of awareness of the breadth of literature on that topic.

Protocols for carrying out a systematic or evidence-based review

The main steps for carrying out a systematic review are drawn from healthcare and can be summarized as follows (Higgins and Green, 2004, quoted in Hidalgo Landa et al. 2011):

1. Define the search terms.
2. Identify the databases and search engines, as well as journals which may need to be hand searched, and query with the chosen search terms.
3. Decide on, and apply, filters for inclusion and exclusion.
4. Ensure that the resulting articles are representative, by repeating the filtering process.

Some users of systematic reviews adopt variations to these common protocols. For example, the Evidence for Policy and Practice Information and Co-ordinating Centre at London University's Institute of Education will consult with a range of users about subjects for review, to avoid policy-making becoming too political.

Whatever the protocol used, it needs to be carefully documented in such a way as to be transparent; other researchers should be able to follow the same procedures and get the same results. This will greatly assist its external validity.

It also removes bias, which may distort subsequent research.

Methodological checklists, with questions and templates, have been devised to aid the process of critical appraisal. The International Centre for Allied Health Evidence has provided a gateway to these checklists: see <http://w3.unisa.edu.au/cahe/resources/cat/default.asp> (Look under the heading, Systematic Reviews).

Doing any literature search is likely to give rise to a large number of articles, so it will be useful to have access to bibliographic software which can manage and help with the citation of references, such as EndNote (<http://www.endnote.com/>) or RefWorks (<http://www.refworks.com/>).

Defining search terms and identifying articles

The first step in any research is to determine the research question, which should be clearly formulated, and without any ambiguity. You should also avoid asking two questions in one.

One way of approaching question formulation is that proposed in the acronym CIMO (Denyer and Tranfield, 2009, quoted in Colicchia and Strozzi ,2012, p.7):

- Context: What is being studied – i.e. individuals, institutions etc.?
- Intervention: What event or action is being studied?
- Mechanisms: What mechanisms explain the relationship between the interventions and outcomes?
- Outcome: What are the effects of the intervention, and how will it be measured?

Search terms

Once you have a question or questions, you can start to list keywords, i.e. words and phrases with which you can interrogate search engines and databases. You can use either free text, or the terms identified by the software's thesaurus.

It is important to include variations in common terms.

Hotchkiss et al. (2012) researched the question, "How Can Routine Health Information Systems Improve Health Systems Functioning in Low- and Middle-Income Countries?" using keywords "routine health information systems", "health management information systems", "health information systems", "electronic health records", and "electronic medical records". They then used keywords "model", "conceptual", and "framework".

You can combine your keywords using Boolean operators:

Colicchia and Strozzi (2012) developed a new methodology for carrying out a systematic review of risk management in the supply chain. They brainstormed 20 keywords, then refined them with a team of three academics and two supply chain managers. They then combined the keywords in a series of strings using Boolean logic (again brainstormed and refined); they were particularly looking for the overlap between Risk Management and Supply Chain Management, but wanted to avoid results that were too generic. They ended up with 30 strings, for example "supply chain risk" AND (vulnerability OR complex*)

Querying search engines and databases

You will need to use search out and use all databases that are relevant to your subject area.

Scott et al. (2012) needed to consult peer reviewed medical and nursing literature for their research on inpatient handovers. They searched Pub Med, Cumulative Index to Nursing and Allied Health Literature (CINAHL), and the Cochrane Library, although they excluded grey literature.

Manning Fiegen (2010) conducted a review of business instruction literature, and used library and business education bibliographic databases for English Language publications between 1980 and 2009. Databases included EbscoHost Premier, Emerald Fulltext, Library Literature and Information Science, LISA, and the ISI's Web of Knowledge.

The automated searching of databases and search engines may be insufficient to uncover all articles, and may need to be supplemented. Journals and websites may need to be hand-searched, personal contacts approached, citations and references followed, etc.

Mullen and Singh (2010) researched breakfast consumption, carrying out electronic searches in Proquest5000, PsychINFO, Medline, and Web of Knowledge, with keywords breakfast, skip*, miss*, eat*, consum*, have, ingest, omit*. Additional papers were sourced from reference lists, and the OVID Dissertation and Thesis Database was searched for grey material.

Settles and Kuskova (2012) look at methodology for emerging market research, and identify the leading management journals through the list provided by Podsakoff, Mackenzie, Bachrach and Podsakoff (2005), excluding those aimed at a practitioner market, or carrying mostly conceptual articles. See their full list on p. 131.

Citation network analysis

Another technique, known as *Citation Network Analysis (CNA)*, is as its name suggests based on tracking citations. The advantage is that it can show the evolution of knowledge as opposed to merely providing a snapshot at one point in time.

The idea behind CNA is that citations establish a network of knowledge, on the assumption that an article which cites many other articles in the field will both be cited itself and also cause an increase in the articles it cites, thereby acting as a sort of knowledge junction (Colicchia and Strozzi, 2012, p. 6).

Analysis of the network of citations helps to show the dynamic behavior of the field (Colicchia and Strozzi, 2012, p.11). Knowledge exists not just in a formless mass of articles, but in a system of nodes and networks.

Evaluation and synthesis

Once you have carried out your preliminary search, you will undoubtedly have a large number of abstracts. These will probably be your main focus at this stage.

Evaluation

This stage involves establishing filters so that you weed out, using the abstracts, the articles that are not relevant.

As will all previous stages, you are aiming for a literature review that is free of bias. Also one which takes a feasible amount of time to complete.

There are no set criteria for what filters should be, but they must be explicit and open to scrutiny, so that people can see what evidence the authors are drawing on and what they have discarded and why; not always the case in non-systematic reviews (Evidence for Policy and Practice Information and Co-ordinating Centre, 2007).

Writing about the field of information systems, Hidalgo Landa et al. (2011, p.50) propose a highly systematized set of filters:

- Global filters which are metadata-based, i.e. which exclude or include based on author, years of publication etc.
- Semantic – i.e. exclude those articles which use the same terms but with a different meaning.
- Evidence-grade – based on the quality of the article.

Should the resulting set of articles be too large to be reviewed manually, they recommend using a simple randomized selection.

In actual fact, Emerald articles employ a range of filters. Relevance to the study's subject is the most common: cf. Manning Fiegen (2010: p.388), Scott et. al. (2012), Blank et al. (2010), Siha and Saad (2008, p.780) who used the filter of sufficient information on BPI implementation and rigorous coverage as a secondary review, Coward (2010), Culverwell et al. (2008: p.41), Hotchkiss et al. (2012: p.29), Yeager and Menachemi (2011: p.237), and Unruh et al. (2009: p.120).

Keywords can be used to search either the abstract or fulltext.

Coward (2010: p.710) used the keywords "educational quality assurance", "healthcare", and "clinical governance" to filter his collection of abstracts for relevance.

After subject, the second most common filter used by authors is quality/type of paper.

Scott et al. (2012) filtered papers according to whether or not they discussed handover methods and came to a conclusion.

Manning Fiegen (2010) excluded articles that were not peer reviewed, and where the author did not include one librarian.

Siha and Saad (2008: p.780) excluded rhetorical, opinion-based and theoretical articles, concentrating on case studies and research articles.

Parker and Lemak (2011: p.156) excluded commentaries and articles that were not research articles.

Year(s) of publication is an obvious filter: Culverwell et.al's 2008 dementia study concentrated on articles published in the previous ten years.

Geography can be a filter, but it can also be a result: the selected articles reveal the geographical spread of research and publication.

Braye et al. (2011: p.120), when surveying the literature on self-neglect, found that three-quarters of articles selected for full review came from North America.

Unruh et al. (2009), in their study of nursing staff measures, limited their research to the US.

Finally, there is the practical issue of accessibility: Yeager and Menachemi (2011: p.137) excluded articles which were not accessible as fulltext.

Synthesis

The resulting set of articles will need to be reviewed, and the results presented.

One good way of setting out the results is to have a table with the study references listed against the findings, as in this (partial) example from Yeager and Menachemi (2011: p.244):

Study Reference	Study Design	Population Studied (n)	Primary SMS Outcome Positive (Yes/No)	Findings
Franklin et al. (2006)	RCT	Diabetic children 8–18 years old (92)	No	SMS did not in itself improve glycemic control, but it helped support the introduction of intensive insulin therapy.
H. S. Kim (2007)	RCT	Diabetes patients (51)	Yes	Nurse-led patient education via SMS lowered blood glucose levels in patients.
C. Kim et al. (2007)	RCT	Diabetic adults (80)	Yes	Computer algorithms can be used to generate clinical recommendations via SMS for improvement of glucose control by providing medical advice
H. S. Kim and Jeong (2007)	RCT	Diabetic adults (51)	Yes	After 3 and 6 months glycosylated hemoglobin decreased, 2 hours post-meal glucose.
Franklin et al. (2008)	RCT	Diabetic children 8–18 years old (92)	Yes	Patients receiving intensive diabetes management, comprising intensive insulin therapy and daily text message intensive support, showed significant improvements in glycemic control compared to groups allocated to conventional insulin therapy.
Yoon and Kim (2008)	RCT	Diabetic adults (50)	Yes	SMS group had decreased HbA _{1C} and 2-hour post-meal glucose levels. No difference in total cholesterol triglycerides and HDL cholesterol.
Cho et al. (2009)	RCT	Diabetic adults (69)	Yes	Baseline HbA _{1C} levels were significantly decreased in both the Internet and SMS groups after 12 weeks' intervention.

Given the likely complexity of the resulting research, authors often use summary tables. For example, Yeager and Menachemi coded the articles they studied according to certain characteristics which they summarized in tabular form as follows (2011:p.241):

Table 1. Summary of Studies of the Impacts of SMS^a in Health Care.

Variable	Total <i>n</i> (%)
<i>Study design</i>	
<i>Experimental (n)</i>	31 (50.8%)
Randomized control trial (31)	
<i>Observational (n)</i>	30 (49.2%)
Cohort study (13)	
Case series (9)	
Pre/post test (5)	
Retrospective review (2)	
Case study (1)	
<i>Health care setting</i>	
Specialty care	34 (55.7%)
Primary care	15 (24.6%)
Other	12 (19.7%)
<i>Type of subject studied</i>	
Patients with diabetes	19 (31.1%)
Smokers	7 (11.5%)
Patients with mental health disorders	4 (6.5%)
Pregnant women	2 (3.2%)
Obese/overweight adults	2 (3.2%)
Some other study population	27 (44.5%)
<i>Continent of research</i>	
Europe	26 (42.6%)
Asia	18 (29.5%)
North America	9 (14.8%)
Australia/New Zealand	8 (13.1%)
<i>Direction of SMS communication</i>	
One-way to patient	38 (62.3%)
One-way to provider	4 (6.6%)
Two-way	19 (31.1%)
<i>Focus of article</i>	
Disease management	27 (44.3%)
Public health related	24 (39.3%)
Management/administration of health services	10 (16.4%)

^aShort message service.

For a full description of the results of the literature review, see pp. 240 ff.

Commenting on the paucity of comparative studies of emerging markets, Settle and Kuskova (2012: p.138ff.) discuss the various methods of analysis, and their pros and cons. They also provide a tabular summary:

Table 2. Methods of Assessing Group Variation at –a Glance.

Method	Requirements	Advantages	Limitations	Recommendations
Logistic regression	Binary or limited length ordinal outcome variable Relatively large sample size (50 data points per predictor to achieve stable results)	Robust to violations of many statistical assumptions (normality of observations, equal variance in groups, homogeneity of variance, normal distribution of effort terms) Can handle nonlinear effects, explicit interactions, and power terms Does not assume linear relationship between independent and dependent variables Provides rank-order of relative importance of independent variables Widely available instrument, easy to implement and interpret	Supports only one dependent variable Limited number of outcomes for the dependent variable (binomial or presented as classes)	Ideal instrument of choice when the question of interest is the contribution of independent variable to affiliation of a country to emergent or developed market group
Multiple group comparison using analysis of covariance modeling	Relatively large sample size	Tests several dependent relationships simultaneously Explicitly models measurement error No limitations on the type of the number of outcome variables Allows comparison between groups on every level, from individual variable to the entire model, thus accounting for heterogeneity in data	Inability to adjust for type I and type II errors Inability to perform post hoc analysis Technique is complex and may require training Cannot implicitly handle non-independent data; requires complex modeling by incorporating design effects	Ideal instrument of choice when two or more groups of data need to be compared on every level, from individual observation to complex relationships
Hierarchical linear modeling	Data taken on at least two levels.	Allows testing data crossing more than one level of analysis Controls for autocorrelation, heteroscedasticity, within-respondent variance, and group effects	Does not guarantee control for unobserved cross-level factors Does not work well with longitudinal data with short series length	Ideal instrument of choice when independence of observations is questionable – that is, data is nested

The object of synthesizing the results of the papers reviewed is just that: rather than just list individual findings, it should draw together common strands, show up apparent contradictions, etc.

According to EPPI, there are three types of synthesis (Evidence for Policy and Practice Information and Co-ordinating Centre, 2007):

- Statistical: the numerical results arising from the combined studies are subject to a set of statistical tests. This is known as meta-analysis.
- Narrative: the combined results are summarized verbally, for example, organized according to theme, study type etc.
- Conceptual: the different concepts of the studies are brought together and a new concept described.

When reviewing the literature, it is useful to have a system for recording the data, especially if more than one person is involved. For example, there could be a system for coding:

- general description;
- results; and
- methods.

In synthesizing the results of their study on inpatient handovers, Scott et al. (2012) used a range of techniques. Following a tabular analysis, they organized data according to various themes from papers they had classified as implementation studies, conceptual models, background papers and reviews. They used concept mapping to explore relationships between the data and indications as to what worked for whom. The background papers were used to derive contextual issues and dimensions of practice. See pp. 17ff. for a full discussion of the results.

Review of evidence

Seeing as many literature reviews are conducted with a view to determining the best form of evidence for a particular intervention, a systematic review needs to conclude by doing just that.

Scott et al.'s 2012 look at inpatient handover concludes with a look at the evidence for what works best, concluding that there are a number of recurring themes in the literature (p.21).

Siha and Saad (2008, p.791) provide a summary of empirical evidence on the success factors and causes of failure of each type of approach to business process improvement.

Summary

A systematic review collects and reviews the literature using explicit methods which can be independently verified.

It starts by using keywords and search strings to query all relevant databases, and may also carry out a manual search of known or recommended sources.

The result is a series of abstracts which again need to be filtered to produce an objective and scientific selection.

The resulting articles are then read, and summarized.

Yeager and Menachemi (2011) depict the process diagrammatically as follows:

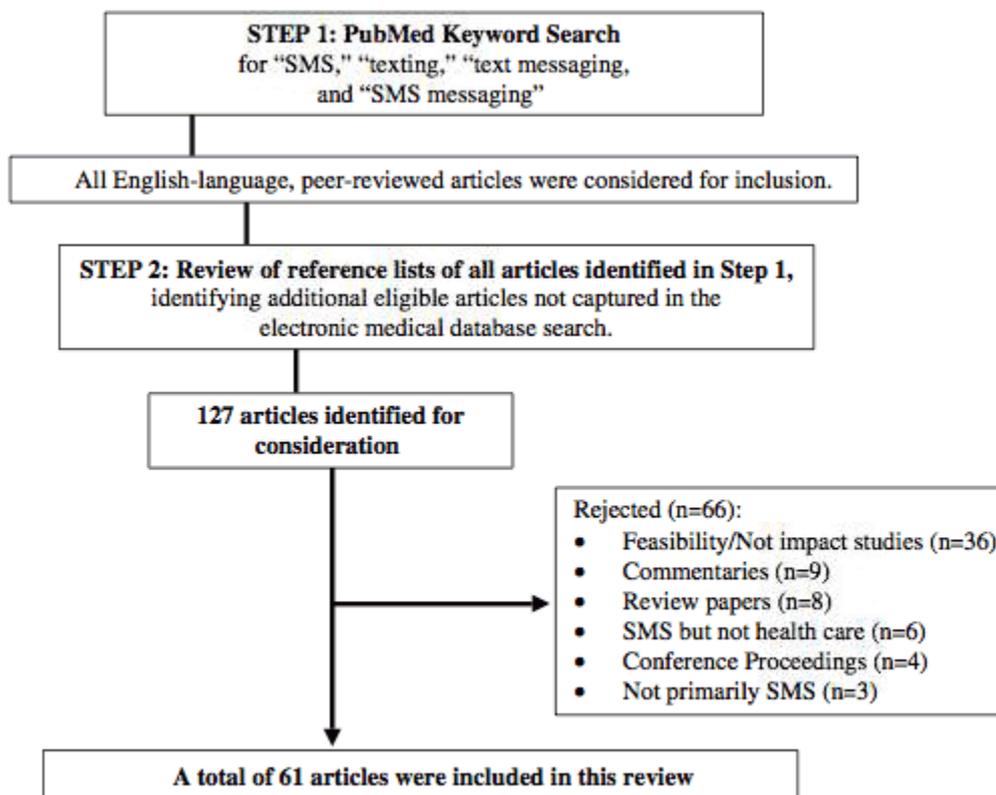


Fig. 1. Steps to Identify Articles for the Systematic Review.

The result is an authoritative body of knowledge which can serve as a basis for further research, or can help us know what particular intervention works best in what circumstances. Both are of considerable value, and repay the effort involved.

References

- Blank, L., Baxter, S., Goyder, E., Naylor, P., Guillaume, L., Wilkinson, A., Hummel, S., and Chilcott, J. (2010), "Promoting well-being by changing behaviour: a systematic review and narrative synthesis of the effectiveness of whole secondary school behavioural interventions", *Mental Health Review Journal*, Vol. 15, No. 2, pp.43-53, ISSN 1361 9322.
- Braye, S., Orr, D., and Preston-Shoot, M. (2011), "Conceptualising and responding to self-neglect: the challenges for adult safeguarding", *Journal of Adult Protection, The*, Vol. 13, No. 4, pp.182-193, ISSN 1466 8203.
- Colicchia, C., and Strozzi, F. (2012), "Supply chain risk management: a new methodology for a systematic review", *Supply Chain Management: An International Journal*, Vol. 17, No. 4, pp.403-418, ISSN 1359 8546. (Date online 15/5/2012) Downloaded on: 19-06-2012.
- Coward, R. (2010), "Educational governance in the NHS: a literature review", *International Journal of Health Care Quality Assurance*, Vol. 23, No. 8, pp.708-717, ISSN 0952 6862.
- Evidence for Policy and Practice Information and Co-ordinating Centre, (2007), EPPI-Centre Methods for Conducting Systematic Reviews, EPPI-centre, Social Science Research Unit, Institute of Education, University of London, available at <http://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=hQB8y4uVwI%3d&tabid=88>, accessed June 25th 2012.
- Hidalgo Landa, A., Szabo, I., Le Brun, L., Owen, I., and Fletcher, G. (2011), "Evidence Based Scoping Reviews", *The Electronic Journal Information Systems Evaluation*, Vol. 14, No. 1, pp. 46-52. Available online at www.ejise.com, accessed 26th June 2012.
- David R. Hotchkiss, Mark L. Diana, Karen G. Fleischman Foreit, (2012), "How Can Routine Health Information Systems Improve Health Systems Functioning in Low- and Middle-Income Countries? Assessing the Evidence Base", Nir Menachemi, Sanjay Singh, in (ed.) *Health Information Technology in the International Context (Advances in Health Care Management, Volume 12)*, Emerald Group Publishing Limited, pp. 25-58.
- Fiegen, A.M. (2010), "Systematic review of research methods: the case of business instruction", *Reference Services Review*, Vol. 38, No. 3, pp. 385-397, ISSN 0090 7324.
- Mullan, B.A., and Singh, M. (2010), "A systematic review of the quality, content, and context of breakfast consumption", *Nutrition & Food Science*, Vol. 40, No. 1, pp. 81-114, ISSN 0034 6659.
- Victoria A. Parker, Christy Harris Lemak, (2011), "Navigating Patient Navigation: Crossing Health Services Research and Clinical Boundaries", in John D. Blair, Myron D. Fottler (ed.) *Biennial Review of Health Care Management (Advances in Health Care Management, Volume 11)*, Emerald Group Publishing Limited, pp.149-183.
- Scott, P., Ross, P., and Prytherch, D. (2012), "Evidence-based inpatient handovers: a literature review and research agenda", *Clinical Governance: An International Journal*, Vol. 17, No. 1, pp. 14-27, ISSN 1477 7274.
- Alexander Settles, Valentina Kuskova, (2012), "When West Meets East: Methods of Assessing Group Variation in Comparative Emerging Markets Research", Catherine L. Wang, David J. Ketchen, Donald D. Bergh, in (ed.) *West Meets East: Toward Methodological Exchange*

(*Research Methodology in Strategy and Management*, Volume 7), Emerald Group Publishing Limited, pp. 127-160.

Siha, S.M., and Saad, G.H. (2008), "Business process improvement: empirical assessment and extensions", *Business Process Management Journal*, Vol. 14, No. 6, pp. 778-802, ISSN 1463 7154.

Culverwell, A., Milne, A., Guss, R., and Tuppen, J. (2008), "Screening for dementia in primary care: how is it measuring up?", *Quality in Ageing and Older Adults*, Vol. 9, No. 3, pp.39-44, ISSN 1471 7794.

Valerie A. Yeager, Nir Menachemi, (2011), "Text Messaging in Health Care: A Systematic Review of Impact Studies", John D. Blair, Myron D. Fottler, in (ed.) *Biennial Review of Health Care Management (Advances in Health Care Management*, Volume 11), Emerald Group Publishing Limited, pp. 235-261.

Lynn Unruh, C. Allison Russo, H. Joanna Jiang, Carol Stocks, (2009),"Measuring up: Are nurse staffing measures adequate for health services research?", Grant T. Savage, Myron D. Fottler, in (ed.) *Biennial Review of Health Care Management: Meso Perspective (Advances in Health Care Management*, Volume 8), Emerald Group Publishing Limited, pp. 115-154.

Printed from: http://www.emeraldinsight.com/authors/guides/write/evidence_based.htm?part=1
on Monday February 18th, 2013

© Emerald Group Publishing Limited