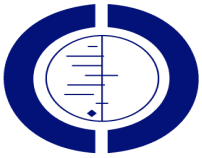


THE COCHRANE COLLABORATION®

Cochrane Health Promotion and Public Health Field

www.vichealth.vic.gov.au/cochrane

GUIDELINES

Systematic Reviews of Health Promotion and Public Health Interventions

Rebecca Armstrong, Elizabeth Waters on behalf of the *Guidelines for Systematic Reviews in Health Promotion and Public Health Taskforce.*

July 2007

Version 2

October 2007

These guidelines have been developed for the purpose of advising and supporting systematic reviews of health promotion and public health interventions within the Cochrane Collaboration.

Relevant publications, conference proceedings (papers, workshops, posters) relating to the guidelines:

1. Jackson N, Waters E, for the Guidelines for Systematic Reviews of Health Promotion and Public Health Interventions Taskforce. Guidelines for Cochrane systematic reviews of public health interventions. Health Promotion International 2005; 20 (4): 367-74. .
2. Jackson N, Waters E; Guidelines for Systematic Reviews of Health Promotion and Public Health Interventions Taskforce. The challenges of systematically reviewing public health interventions. J Public Health (Oxf). 2004 Sep;26(3):303-7.
3. Public Health Association 35th Annual Conference, September 29 to October 1, 2003, Brisbane, Australia. Proffered Paper: "Systematic reviews of health promotion and public health: recommendations of reviewers".
4. "Guidelines for reviewers of health promotion and public health interventions". XI Cochrane Colloquium, Barcelona, Spain. 26-31 October, 2003. (Workshop and poster)
5. "Guidelines for Systematic Reviews of Health Promotion and Public Health Interventions", The Fourth Annual Campbell Collaboration, Washington DC, February 18-20, 2004.(poster)
6. "Guidelines for Systematic Reviews of Health Promotion and Public Health Interventions", IUHPE Conference, Melbourne, Australia. 28th April 2004. (poster)
7. "Guidelines for Systematic Reviews of Health Promotion and Public Health Interventions", Australasian Contributors Meeting, Cochrane Collaboration, Sydney, Australia, 17-18 June, 2004. (Workshop)
8. "Guidelines for Systematic Reviews of Health Promotion and Public Health Interventions", 12th Cochrane Collaboration Colloquium, Ottawa, Canada, 1-6 October, 2004. (Workshop and poster)
9. Armstrong R, Waters E, Roberts H, Oliver S, Anderson L and Petticrew M. Systematic reviews of health promotion and public health interventions. In Heggenhougen K, (Ed), Encyclopedia of Public Health (*in press*). Elsevier Press.
10. Armstrong R and Waters E. Conducting systematic reviews of health promotion and public health interventions. Workshop presented at IUHPE 2007, Vancouver Canada.
11. Armstrong R, Waters E, Doyle J. Systematic reviews of health promotion and public health interventions. Worksop presented at the X111 Cochrane Colloquium, 22-26 October 2005, Melbourne.

How to cite this version of the guidelines:

Armstrong R, Waters E, Jackson N, Oliver S, Popay J, Shepherd J, Petticrew M, Anderson L, Bailie R, Brunton G, Hawe P, Kristjansson E, Naccarella L, Norris S, Pienaar E, Roberts H, Rogers W, Sowden A, Thomas H. Guidelines for Systematic reviews of health promotion and public health interventions. Version 2. Melbourne University: Australia.October 2007.

Feedback on the guidelines is warmly welcomed and can be forwarded to cochrane@vichealth.vic.gov.au

Summary of Changes to Cochrane HPPH Guidelines

Chapter 1: Planning the review	The EPPI Centre has devised some pertinent questions for review teams to consider with regards to planning the review.
Chapter 2: Study designs to include	Highlights some new research funded by the UK Methodology Programme. This work will explore the implications of the choice of study design when estimating the effects of policy interventions.
Chapter 3: Searching for health promotion and public health literature	Discusses some of the issues in searching for qualitative research. The use of searching filters is discussed. A series of changes have been made to the recommendations for handsearching.
Chapter 4: Quality assessment	Highlights www.health-evidence.ca as a source for assessing the quality of systematic reviews. Includes a link to the combined work of the Cochrane Qualitative Research Methods Group and Campbell Collaboration Process Implementation Methods Group.
Chapter 5: Theoretical framework	No major changes
Chapter 6: Integrity of intervention	No major changes
Chapter 7: Heterogeneity in public health and health promotion reviews	No major changes
Chapter 8: Integrating qualitative and quantitative studies	Recommendations for review synthesis have been revised. Some recommendations around the conduct of narrative synthesis have also been included. An introduction to the inclusion of qualitative studies in systematic reviews is also provided.
Chapter 9: Ethics and Inequalities	Significant re-writing to this chapter has occurred. Much of this includes the introduction of more up-to-date literature to explore concepts of ethics and inequalities. Work of the Campbell and Cochrane Health Equity Field is included. Recommendations remain essential the same.
Chapter 10: Sustainability	A list of issues to consider when reviewing sustainability of included studies has now been added to this chapter.
Chapter 11: Context	New papers exploring external validity have been included in this chapter.
Chapter 12: Applicability	Some questions for assessing the applicability and transferability of interventions have been included.

Table of contents

Acknowledgements

We are grateful to the *Guidelines for Systematic Reviews in Health Promotion and Public Health Taskforce* for the formulation and writing of sections within these Guidelines.

Taskforce members:

Rebecca Armstrong, Cochrane HPPH Group, University of Melbourne, Australia;
Dr. Laurie Anderson, Centers for Disease Control and Prevention, USA, and Affiliate Assistant Professor, Department of Epidemiology, School of Public Health, University of Washington, USA;
Associate Professor Ross Bailie, NHMRC Senior Research Fellow, Menzies School of Health Research, Australia;
Ginny Brunton, Evidence for Policy and Practice Information and Co-ordinating Centre, Social Science Research Unit, University of London, UK;
Professor Penny Hawe, Markin Chair in Health and Society, Department of Community Health Sciences, University of Calgary, Canada;
Nicki Jackson, Auckland University of Technology
Assistant Professor Elizabeth Kristjansson, School of Psychology and Institute of Population Health, University of Ottawa, Canada;
Dr. Lucio Naccarella, Department of Public Health and Department of General Practice, University of Melbourne, Australia;
Dr. Susan Norris, Centers for Disease Control and Prevention, USA;
Dr. Sandy Oliver, Reader in Public Policy, Evidence for Policy and Practice Information and Co-ordinating Centre, Social Science Research Unit, University of London, UK;
Professor Mark Petticrew, Associate Director MRC Social and Public Health Sciences Unit, UK;
Elizabeth Pienaar, South African Cochrane Centre;
Professor Jennie Popay, Professor of Sociology and Public Health, Lancaster University, UK;
Professor Helen Roberts, City University, UK;
Associate Professor Wendy Rogers, Flinders University, Australia;
Jonathan Shepherd, Wessex Institute of Health Research and Development, University of Southampton, UK;
Dr Amanda Sowden, Centre for Reviews and Dissemination, University of York, UK;
Associate Professor Helen Thomas, School of Nursing, McMaster University, Canada, and Clinical Consultant, Hamilton Department of Public Health and Community Services, Public Health Education and Development Program, Canada;
Professor Elizabeth Waters, Professorial Fellow, Public Health and Health Equity, University of Melbourne, Australia.

Foreword

These guidelines complement Sections 3 to 11 of the Cochrane Reviewers' Handbook (http://www.cochrane.org/resources/handbook/index.htm). The content has been prepared by health promotion and public health researchers, decision-makers and practitioners experienced with both the use and conduct of systematic reviews. While these guidelines were originally developed to support the conduct of systematic reviews, they are also important for the conduct of primary research and for more informal reviews of research evidence. Many of the topics may not be unique to health promotion and public health reviews, but they are issues that are important in enabling research to be used in public health policy and practice decision making.

Conducting systematic reviews of complex health promotion and public health interventions can be methodologically challenging due to a number of reasons.[1] Systematic reviews should aim to address two questions; 1) does the intervention work (effectiveness), and 2) why does it work (including how does it work)? The guidelines present recommendations to enable authors to address the above questions.

The guidelines contain twelve topic areas. Each topic is designed with two sections: background and recommendations. The guidelines suggest that review authors consider and address each of the recommendations in their review. Whilst it may not be possible to elicit some of the recommended information from the studies, for example, intervention context, theoretical frameworks, and process data, it is important to report when this data is not available. The gaps in reporting should stimulate further improvements in public health and health promotion research, and its publication.

These guidelines are intended to be iterative and will be updated every eighteen months with substantive updates every five years. We encourage those who use these guidelines to provide us with feedback. An evaluation sheet is available at the end of this document and via the Cochrane HPPH Group website (http://ph.cochrane.org). Please send any feedback to cochrane.vichealth.vic.gov.au

The Cochrane HPPH Field is in the process of transitioning from Field to Review Group. This new group will have an increased focus on upstream interventions, and recommendations and methods relevant to this particular application will be developed during 2008.

1. Planning the review

This section describes the process of deciding the scope of your review. Clinical and public health authors often diverge on the scope of reviews (broad versus narrow questions) which often reflects the population versus individual perspective of both disciplines. The recommendations below should enhance the review process and minimise any potential concerns.

Advisory Groups

Systematic reviews are likely to be more relevant to the end user and of higher quality if they are informed by advice from people with a range of experiences, in terms of both the topic and the methodology[2-4] and the application to decision making. The decisions made in the early stages of the review process will influence the content of the protocol and the subsequent review. The priorities of decision-makers and consumers may be different to the priorities of the author. Authors should address questions which are important to stakeholders and include relevant interventions, outcomes and populations. Therefore, the first step in any review should be to form an Advisory Group of people, including consumers, with relevant interests, skills and commitment. The principal author will need to coordinate their input to inform key review decisions.

A number of factors beyond the research evidence (e.g. resource constraints, values, timing, politics) affect decision-making.[5] Therefore, decision makers in different countries may ask different questions even when the content area is similar. The Cochrane Collaboration emphasises the importance of incorporating the needs of resource poor countries in each review process, and to ensure that the perspective of vulnerable and marginalised people are represented in the Advisory Group.[6] This should ensure that the conclusions regarding the value of the interventions are well informed and applicable to all groups in society.

Scope of the review

There are a variety of factors that will affect the scope of your review question. 'Lumping' or 'splitting' have become terms that reflect whether you are taking a broad perspective that could be comprised of more than one review, or splitting, which conveys that you are interested in a narrower question.

Lumping the review question, i.e. addressing a wide range of interventions, is likely to be time-consuming because of the searching and selecting processes. However, a lumped review question will better inform decisions about *which* interventions to implement when there may be a range of options. E.g., 'prevention of injuries in children' is a lumped review, but could be addressed by a range of individual review questions including 'prevention of drowning in toddlers'. These can then form the basis of an overview of reviews; a combination of a range of smaller reviews to address a broad question.

Splitting the review, or addressing a narrow range of interventions, may be less time-consuming. However, these reviews can only inform decisions about whether or not to implement narrowly focused interventions.

Split reviews may be more likely to inform immediate decisions of policy relevance. However, reviews that seek to answer broad questions may ultimately be of more use to policy makers. Consumers have been critical of reviews that emphasise single and/or narrow outcomes.[7] Policy makers are also more interested in finding out about interventions which affect multiple outcomes. Many social interventions are explicitly targeted at a range of outcomes (eg. public or social policy), and other interventions may result in affecting a range of outcomes.[8]

Recommendations

This section provides advice to authors on Advisory Groups and the key components of a public health or health promotion review question. It supplements Chapter 4 of the Handbook.

The following references provide advice on:
Identifying the need for a review[4];
Formulating the research question[9];
Technical and resource planning[10];
Involving consumers in scoping reviews, making them accessible and getting them used[11].; and
Planning for and conducting systematic reviews in the social sciences[12].

We complement these resources with suggestions of how to plan a review that focuses on health promotion and public health interventions.

Forming an Advisory Group

Choosing which interventions to review requires knowledge of current policy, practice and the views of the consumers or the population targeted by the interventions. Authors should aim to include a broad spectrum of lay and scientific people on their Advisory Group.

- Establish an Advisory Group. The members should be familiar with the topic and may include policy, funders, practitioners and potential recipients/consumers. Methodologists may also be included to assist in methodological questions. The broader the review, the broader the experience required of Advisory Group members. To identify priority topics for reviews, authors should also consult health professionals in developing countries.[13] Although we do not recommend a specific number of members for the Advisory Group, it is likely that an approximate size of six would be able to cover relevant dimensions and is manageable.

- Develop Terms of Reference for your Advisory Group to ensure there is clarity about the task(s) required. Tasks may include:
  - making and refining decisions about the scope of the review (populations, interventions, outcomes and possibly sub-group analyses of interest)
  - providing or suggesting important background material that describes the issues from different perspectives
  - helping to interpret the findings of the review
  - designing a dissemination plan and assisting with dissemination to relevant groups

- The Cochrane HPPH Field has developed some guidance for establishing an advisory group. This document can be accessed at: http://www.ph.cochrane.org/en/authors.html

- Develop job descriptions for consumers and other advisors to clarify expectations. Examples are provided in briefing notes for researchers.[14] Further information is also available at www.invo.org.uk.

Key review decisions

Key decisions about which interventions, populations, settings and outcomes to address will need to be made.

a. Scope of the review

- The scope of the review should be based on how the results of the review will be used. For example, the users of the review may be more interested in focusing on interventions or a specific intervention, populations, or outcomes.

- The scope of the review will also depend on how much time is available and the likely volume of research on the topic. Once you have identified your scope, you will need to estimate the number of citations you may have to screen. This can be done during the protocol stage by developing and testing your search strategy.

b. Populations, interventions and outcomes

Section 4.2.1 of the Handbook outlines the steps to consider when deciding which participants or populations, interventions and outcomes to include in a review.

- Consider whether there is value in limiting the population (eg. street youth, problem drinkers). These groups are often under-studied and may be different in a number of important respects from study populations usually included in reviews. *See Section 10 Ethics and inequalities for more information*.

- Qualitative research can contribute to framing the review question (eg. selecting interventions and outcomes of interest to participants).

- Determine if proximal/immediate, intermediate or distal outcomes are to be measured. If only intermediate outcomes are measured (eg. blood sugar levels in persons with diabetes) authors need to determine how strong the linkage is to more distal outcomes (eg. cardiovascular disease).

An example of the benefits of using an Advisory Group in the planning process
A review of HIV prevention for men who have sex with men (MSM) (Rees 2004b) employed explicit consensus methods to shape the review.
An Advisory Group was convened consisting of:
- research/academic organisations;
- policy representatives;
- service organisations; and
- representatives from charities and organisations that have emerged from and speak on behalf of people living with, or affected by, HIV/AIDS.

The group met three times over the course of the review. The group was presented with background information about the proposed review; its scope, conceptual basis, aims, research questions, stages, methods. Discussion focused on the policy relevance and political background/context to the review. The group made decisions on the inclusion criteria, dissemination strategies and timescales. Two rounds of voting identified and prioritised outcomes for analysis. Open discussion identified sub-groups of vulnerable MSM. A framework for characterising interventions of interest was also refined through Advisory Group discussions. The review followed all guidance provided by the Advisory Group.
The final review included synthesised evidence directly relating to health inequalities.

Practicalities
Conducting a review requires team work. Review author teams for health promotion or public health reviews may need to be large to deal with the breadth of scope and diverse sources. For this reason it helps to think about managing people, managing time and managing information in advance.

The EPPI-Centre has devised some pertinent questions for review teams to ask themselves about each stage of the review:

Planning the review:
1. Who has overall responsibility for the review and will act as guarantor for any publications?
2. Who is co-ordinating the review? Who is the primary contact person for people in the advisory group to contact?
3. Who is able to fulfil the co-ordinating role if the person(s) above is unavailable?

Writing the protocol
4. Who is going to be involved in writing the protocol?
5. In what ways are different user perspectives going to included in practice?
6. Who will be co-ordinating this work?
7. How will the group ensure that the methodological and conceptual issues are understood and agreed upon across the review team in time for the protocol to be written?

Setting up a system to manage reports
8. Who is responsible for setting up and maintaining the reference management system?
9. Where will the system be located and what backup procedures are in place in case of hardware failure? (A computer with specialist software and an internet connection will be needed as well as space to store a large number of documents)
10. Which computer software are you going to use to manage citations?
11. What will happen if this person is unable to fulfil this role? Who will take over and where will the data be stored?
12. Who is involved in the work to develop the research tools: inclusion and exclusion criteria, coding schemes to describe studies?

13. How are these tools going to be tested?
14. How will the group ensure that reviewers apply the tools as they are intended?

Searching for studies

15. Who will develop the search strategies to be used in this review?
16. How will the group ensure that studies not listed on electronic databases are found?
17. How will the bibliographic information and tracking of research papers in the review be stored and managed?
18. Where and how will hard and/or electronic copies of the research reports be stored?
19. Have the costs associated with reference retrieval and photocopying been budgeted for?

Synthesis and writing the report

20. Who is going to write the report, or sections of the report?
21. If the report is being written by different people, possibly in different locations, how will you ensure that you don't create multiple copies of the report, each with slightly different content?
22. Who are the named authors going to be?
23. Does the report need to be submitted to an advisory group? If so, how is feedback going to be managed?
24. Who is going to solicit user summaries for the final report? How is this process going to be managed?

2. Study designs to include

Health promotion and public health are broadly-defined activities that are evaluated using a wide variety of approaches and designs. No single method can be used to answer all relevant questions about all public health and health promotion problems and interventions. If the review question has been clearly written then knowledge of the types of study designs needed to answer it should automatically follow.[15] A preliminary scoping search will also help you to be familiar with the types of study designs that may have been used to study the intervention. The criteria used to select studies should primarily reflect the question/s being answered in the review, rather than any predetermined hierarchy.[16] The decisions about which type(s) of study design to include will influence subsequent phases of the review, particularly searching, quality assessment, and analysis (especially for meta-analyses).

Randomised controlled trials

Evidence from evaluated interventions, particularly where there is capacity to randomly allocate participants to an intervention and compare outcomes with those who have not received the intervention, provides a useful source of evidence of effectiveness.[17] For many health promotion and public health interventions these randomised controlled trial designs (RCTs) may not be available, due to issues such as feasibility and ethics. Cluster-RCTs are increasingly used within the field of public health; where the interventions are applied and analysed at the cluster level.[18] These trials can contribute valuable evidence if a sufficient number of cluster units (eg. schools, communities, health services, government areas) are randomised to ensure even distribution of potential confounders among groups.

Other effectiveness studies

Non-randomised controlled trials may also represent the best available evidence (of effectiveness). Non-RCT evidence can give an estimate of the nature, direction and size of effects. Demonstrating the patterns of evidence drawn from different study designs may lead to the development of subsequent study designs to test the intervention. This is particularly important where greater confidence in minimisation of bias is required. Studies generating quantitative data may also be relevant to other kinds of questions beyond effectiveness questions. For example, data may be gathered on the preferences of the likely recipients of the interventions and/or the factors that constrain and/or facilitate the successful outcome of particular interventions.

Research into the differences in effects yielded between randomised and non-randomised studies is currently in progress. The UK Methodology Programme has funded the EPPI-Centre (London) in collaboration with the NHS Centre for Reviews and Dissemination (York, UK), Wessex Institute for Health Research and Development (Southampton, UK) and the MRC Biostatistics Unit (Cambridge, UK) to conduct a study entitled 'RCTs for policy interventions? a review of reviews and meta-regression. The project, which is due to complete in 2007, is investigating the implications of the choice of study design when estimating the effect of policy interventions. Sets of similar interventions evaluated with different study designs will be sought from systematic reviews of policy interventions. Part one of the project will seek exhaustively for systematic reviews where these have described: (a) factors affecting results of systematic reviews of policy interventions that have included both randomised and non-randomised designs to estimate the effect of an intervention; (b) methods used by reviewers to try to identify and/or address the bias and confounding that are uncontrolled for in non-randomised designs.

Part two will focus on a smaller set of systematic reviews and will employ multi-variate regression to examine whether randomised studies produce significantly different results when compared with non-randomised studies and whether such heterogeneity in findings, where it is found, can be explained by factors other than research design (e.g. different

populations). Both parts of the study will suggest how to minimise the potential bias introduced by study design selection when evaluating the effects of policy interventions.

The role of qualitative research in effectiveness reviews

The argument that qualitative research should have a role in systematic reviews has been made previously.[19, 20] The aim of qualitative research is to "provide an in-depth understanding of people's experiences, perspectives and histories in the context of their personal circumstances and settings".[21] It attempts to understand why people behave the way they do[4] in order for us to better understand social action or behaviour.

Qualitative studies can contribute to reviews of effectiveness in a number of ways, including[4]:

- Helping to frame the review question (eg. selecting interventions and outcomes of interest to participants). Systematic reviews have been criticised for reflecting professionals' rather than consumers' priorities.[22] Qualitative research can be used to ensure that reviews do meet consumer needs. *See also Section 1 Planning the review*
- Identifying factors that enable/impede the implementation of the intervention (eg. human factors, contextual factors) – *see also Section 6 Integrity of Interventions, Section 12 Context*
- Describing the experience of the participants receiving the intervention
- Providing participants' subjective evaluations of outcomes
- Helping to understand the diversity of effects across studies, settings and groups – *see also Section 7 Heterogeneity*
- Providing a means of exploring the 'fit' between subjective needs of participants and evaluated interventions. This can be used to inform the development of new interventions or refining existing ones.

Methods commonly used in qualitative studies may include one or a number of the following; interviews (structured around respondents priorities/interests), focus groups, participant and/or non participant observation, conversation (discourse and narrative analysis), and documentary and video analysis. The unit of analysis within qualitative studies may include individuals or single cases, communities, populations or organisations. Anthropological research, which may involve some or all of these methods in the context of wide ranging 'fieldwork' can also be a valuable source of qualitative evidence.

Recommendations

This section provides advice on the types of study designs to consider including in a review. It supplements Chapter 4.2.4 of the Handbook which outlines the study designs to include when asking an answerable question.

Effectiveness studies

- Where RCTs (including cluster RCTs) are available to answer *questions of effectiveness or efficacy* they should be included in your review. This type of study design has the greatest potential for maximising internal validity.

- RCTs may not be available, and in these circumstances, non-RCTs are likely to represent the best *available* evidence and should be included.

Other relevant study designs for consideration include

- Controlled before and after studies (CBA): CBA designs measure the outcome variable in participants who have received the intervention and in those who have not, before and after the intervention is implemented.[23]

- Interrupted time series (ITS): ITS designs are those which take multiple observations over time which is 'interrupted' by the intervention. Mass media campaigns are one example of interventions that may be evaluated using ITS designs. The guidance produced by the Cochrane Effective Practice and Organisation of Care (EPOC) group on the review of time series studies is very useful for their appraisal and synthesis (EPOC 2004) – see http://www.epoc.uottawa.ca/inttime.pdf.

- Comparisons with historical controls or national trends: these may be the only evidence available about the effects/effectiveness of some policies and they should be considered for inclusion. They have been included in reviews (eg. [24, 25]) of the Cochrane Tobacco Group.

Integrating qualitative studies into effectiveness reviews

Authors should also refer to the NHS CRD Report Number 4 (2nd Edition)[4] for information on the role of qualitative studies in systematic reviews. The Cochrane Qualitative Research Methods Group (CQRMG) will be developing more detailed guidance in the coming year. In addition the CQRMG is developing a database of methodological texts focusing particularly on the systematic review of qualitative studies.

A number of methodological projects relating to qualitative research have recently been published. Details of relevant projects are highlighted in the report by Popay and Roen.[26] A useful bibliography of methodological studies relating to qualitative research is available on the Qualitative Research Methods Group website (http://www.joannabriggs.edu.au/cqrmg/index.html). The EPPI Centre is continuing their programme of methodological development by testing and refining their methods for integrating different types of studies in reviews. Two papers of particular interest are Oliver et al 2005[27] and Harden et al 2004.[28] Dixon-Woods et al 2006[29] has also recently published a paper outlining how qualitative research can be included in systematic reviews.

Note: You and your Advisory Group will know which study designs are relevant to answer your question. However, it is important to note that each Cochrane Review Group has its own policies regarding the inclusion of study designs other than RCTs. You will need to communicate with your Review Group about the rationale for including study designs other than RCTs.

Available resources for study design selection

Within the Cochrane Collaboration there are a number of resources to help authors working with study designs other than RCTs:

- The Cochrane Non-Randomised Studies Methods Group has produced draft guidelines on the inclusion of non-randomised studies in Cochrane reviews (http://www.cochrane.dk/nrsmg/).
- Some Cochrane Review Groups have produced guidance for authors about their policy of including study designs other than RCTs. The Cochrane HIV/AIDS Review Group held a workshop at the XI Cochrane Colloquium in Barcelona 2003 on 'the inclusion of observational studies in Cochrane reviews'. The group's editorial policy recognises that many important topics in HIV/AIDS prevention and treatment do not lend themselves to being evaluated through RCTs or have not yet been evaluated with RCTs. An 'interactive road map' describing the steps involved in conducting a systematic review with links to sections where further information (e.g. checklists/scales for appraising observational studies) is available on their website (see http://www.igh.org/Cochrane/).
- The Cochrane Qualitative Research Methods Group has developed guidance on the inclusion of qualitative research in systematic reviews http://www.joannabriggs.edu.au/cqrmg/index.html

3. Searching for health promotion and public health literature

Finding studies on public health and health promotion interventions is much more complicated than retrieving medical studies due to literature being widely scattered.[30] The multi-disciplinary nature of health promotion and public health means that studies can be found in a number of different areas and through a wide range of electronic databases .[31, 32] Difficulties also arise because terminology is imprecise and constantly changing.[32] Therefore, searching for public health and health promotion literature can be a very complex task, and requires authors to use retrieval methods other than database searching to retrieve studies. In a review of crime prevention interventions the authors identified more than fifty percent of the included studies from grey literature.[33] A review promoting a shift from cars to walking or cycling found only four of 69 relevant papers in major health databases such as Medline.[34]

Reports of qualitative studies may also be widely dispersed and may be catalogued on databases less familiar to medical researchers.[35] To be able to effectively locate qualitative studies, improvements are required in the indexing of studies and study filters for electronic databases. At present Medline has only one qualitative index (Qualitative Research, indexed 2003), whereas CINAHL utilises a number of methodological indexing terms that accurately describe the qualitative study design. [36]

To overcome some of the difficulties in identifying qualitative research described, current best practice requires the researcher to conduct comprehensive searches (e.g. sensitive searches of multiple sources). However, this approach, which attempts to maximise the number of relevant records identified, results in the retrieval of high numbers of records, many of which will not be relevant.[37] Due to inadequate indexing terms for qualitative research in bibliographic databases, we do not currently recommend that study design filters should be applied. We recognise that often pragmatic decisions may need to be taken when balancing the time and other resources required in conducting comprehensive searches against the ratio of relevant to non relevant studies identified. Researchers may decide that they need to apply study design filters and if so, they need to report this when describing their search strategies to make the potential limitations of the searches clear.

A debate is emerging about the best methods to identify qualitative studies for including in systematic reviews. Booth[38] argues that rather than adopting a 'trials-type search' authors should use a 'theory' driven approach. This is similar to the 'diversity' or 'saturation' approaches to population sampling sometimes used in primary qualitative studies. There is a need for further empirical testing of these different approaches to searching before conclusive recommendations about their utility for systematic reviews can be given.

Recommendations

This section provides advice on developing search strategies, identifying relevant databases to search, and locating grey health promotion and public health literature. It supplements Chapter 5 of the Handbook (Locating and Selecting Studies for Reviews).

Search strategies for identifying public health and health promotion studies

- Use sensitive searches which combine text words with indexing terms (terminology varies between databases and indexing terms are also called e.g. subject headings, descriptors, controlled terms, keywords, thesaurus terms). Indexing terms will also vary between databases (e.g. Homosexuality male/ in Medline and Homosexuals-Male in Cinahl).

- Use text words and synonyms freely as there may be few, or no indexing terms related to your topic, and because terminology varies historically and culturally. Text words are the words used by the authors in the title and abstract, for example a reviewer wishing to identify studies relevant to a review of sexual health promotion with men who have sex with men may need to use a range of terms to describe the population (e.g. Homosexual males, gay men).
- Allocate sufficient time (may take up to five days) to develop, test and re-test the search strategy to make sure it captures relevant studies.
- Use a skilled librarian to assist with search strategies, databases and relevant journals, where available. The review Advisory Group can also assist with determining keywords for electronic database and internet searching.
- Chapter 4 of "Using Research for Effective Health Promotion"[39] outlines further information on searching for health promotion studies.

The Cochrane HPPH Field has developed a training tool for identifying both published and unpublished health promotion and public health literature.[40]

Relevant electronic databases

- Search electronic databases which cover the range of disciplines relevant to your review topic. A list of free public health databases can be found at http://library.umassmed.edu/ebpph/dblist.cfm.
- Due to the delay between publication of a journal and its appearance on an electronic database it is important to check the content pages of recent journal publications to ensure your search is up-to-date. Amedeo (http://amedeo.com/medicine/smo.htm) provides weekly literature overviews in some topic areas/ journals related to HP&PH.

Electronic databases relevant to public health and health promotion include (websites listed for databases available freely via the internet):

Psychology:	PsycINFO/PscyLIT
Biomedical:	CINAHL, LILACS (Latin American Caribbean Health Sciences Literature) http://www.bireme.br/bvs/I/ibd.htm, Web of Science, Medline, EMBASE, CENTRAL, Combined Health Information Database (CHID) http://chid.nih.gov/, Chronic Disease Prevention Database (CDP) http://www.cdc.gov/cdp/ , SCOPUS
Sociology:	Sociofile, Sociological Abstracts, Social Science Citation Index, Social Policy and Practice
Education:	ERIC (Educational Resources Information Center), C2-SPECTR (Campbell Collaboration Social, Psychological, Educational and Criminological Trials Register) http://www.campbellcollaboration.org, REEL (Research Evidence in Education Library, EPPI-Centre) http://eppi.ioe.ac.uk
Transport:	NTIS (National Technical Information Service), TRIS (Transport Research Information Service) http://ntl.bts.gov/tris, IRRD (International Road Research Documentation), TRANSDOC (from ECMT (European Conference of Ministers of Transport)
Physical activity:	SportsDiscus
HP/PH:	BiblioMap, TRoPHI (Trials Register of Promoting Health Interventions) and DoPHER (Database of Promoting Health

	Effectiveness Reviews) (EPPI-Centre) http://eppi.ioe.ac.uk, Public Health electronic Library (National Institute for Health and Clinical Excellence) http://www.phel.gov.uk/, Global Health
Other:	Popline (population health, family planning) http://db.jhuccp.org/popinform/basic.html, Enviroline (environmental health) – available on Dialog, Toxfile (toxicology) – available on Dialog, Econlit (economics), NGC (National Guideline Clearinghouse) http://www.guideline.gov/
Qualitative:	ESRC Qualitative Data Archival Resource Centre (QUALIDATA) (http://www.qualidata.essex.ac.uk), Database of Interviews on Patient Experience (DIPEX) (http://www.dipex.org)

Locating studies by applying study design filters

Authors may or may not wish to apply a study design filter to their search strategy. A search strategy without a filter is likely to result in more work at the inclusion/exclusion stage of the review. However, the search will be more sensitive to identifying relevant studies. There is a need for more empirical search in this area, but authors will make a pragmatic decision based upon time and resources. Researchers may decide that they need to apply study design filters and if so, they need to report this when describing their search strategies in order for the reader to appreciate the potential limitations of the searches.

- RCTs: Refer to Appendix 5b of the Handbook: MEDLINE highly sensitive search strategy for b.1) SilverPlatter-MEDLINE, b.2) OVID-MEDLINE, and b.3) PubMed.
- Non-RCTs: At present, there are no validated filters available to identify non-RCTs. Some review groups recommend search filters for non-RCTs (including ITS designs). However, these filters may not be appropriate for public health and health promotion studies which are often indexed using a variety of study design terms.
- Further examples of search filters are available from the InterTASC Information Specialists' Sub-Group (http://www.york.ac.uk/inst/crd/intertasc/) and from the Hedges Project at McMaster University (http://hiru.mcmaster.ca/hedges/).
- Qualitative research: The Edward G Miner Library has developed a qualitative search filter using CINAHL and is available from: (http://www.urmc.rochester.edu/hslt/miner/digital_library/tip_sheets/Cinahl_eb_filters.pdf). Until developments in search filters for other databases and better indexing of qualitative studies occur, authors will be unable to limit their search by study design on databases other than CINAHL.
- Registers once held by the UK Health Development Agency (HDA) have now moved to the National Institute for Health and Clinical Excellence (NICE) (http://www.publichealth.nice.org.uk/page.aspx?o=home) website. A collaboration with the HDA, Health Evidence Bulletins Wales and the Centre for Reviews and Dissemination explored the indexing of study designs in different public health databases. Results can be found at http://www.publichealth.nice.org.uk/page.aspx?o=516406. The EPPI-Centre has carried out further testing of these terms when searching for public health reviews in social science databases (Brunton 2005) and found that whilst reviews were not consistently tagged, ASSIA proved to be the source with the highest yield of reviews.

Other sources of public health and health promotion grey literature

Internet:

- Use the internet to find suitable organisations that may hold studies.

- Google Scholar (http://scholar.google.com) is a useful search engine to locate studies.
- Government reports may contain relevant studies or references to them.

Handsearching:

- Search generalist public health and health promotion journals or journals relevant to the topic of the review.
- Browse the two lists (not exhaustive lists) of public health journals which have been produced to locate relevant journals to handsearch. Currently, there is no empirical evidence to be able to recommend which journals may be more likely to produce the greatest number of high quality studies.
  - The Lamar Soutter Library list[41] of public health journals compiled by the University of Massachusetts Medical School contains 710 journals.
  - The Core Public Health Journals list[42] compiled by Yale University contains 644 journals.
- Consider open access resources – the Directory of Open Access Journals lists 74 Public Health journals (http://www.doaj.org/ljbs?cpid=21). Projects like the Scientific Electronic Library Online (www.scielo.org) provide free access to scientific journal articles from Brazil, Chile, Cuba, and Spain and may also be a useful resource.
- The Cochrane HPPH Group is developing strategies to identify health promotion and public health trials and reviews to add to databases such as CENTRAL, Bibliomap, TRoPHI and DoPHER. The Cochrane HPPH Group ran a handsearching pilot to identify studies in indexed journals.[43] This work is now being extended to explore the potential for handsearching in non-indexed journals and for unpublished studies.
- The Effective Public Health Practice Project in Canada, in the conduct of their reviews, has found that the most productive journals to handsearch to locate public health and health promotion articles are: *American Journal of Health Promotion*, *American Journal of Preventive Medicine*, *American Journal of Public Health*, *Canadian Journal of Public Health*, *BMJ*. Other useful journals include *Annual Review of Public Health*, *Health Education and Behavior (formerly Health Education Quarterly)*, *Health Education Research*, *JAMA*, *Preventive Medicine*, *Public Health Reports*, *Social Science and Medicine*.
-

4. Quality Assessment

Assessing the quality of public health and health promotion studies may be difficult, partly due to the wide variety of study designs used.

Methodological research has been conducted to identify the dimensions of quality of RCTs most associated with bias.[44-47] Three key attributes were identified: concealment of allocation, blinding of participants, and procedures for dealing with withdrawals from the study and loss to follow-up.

The relevance of these criteria for quality assessment of health promotion and public health needs further consideration. Where RCTs are available, biases due to improper allocation concealment and attrition are relevant. Blinding of participants however is difficult, particularly in educational interventions. Other sources of bias include the potential for the control/comparison group to become 'contaminated' (e.g. within schools where participants in the intervention and control groups are highly likely to come into contact with each other). Contamination may be minimised through the use of cluster RCTs. Potential sources of bias also include inadequate validity and reliability of data collection methods, particularly where outcomes are subjective (e.g. reported behaviour). Blinding of assessors is equally important and should also be considered.

Identifying the strengths and weaknesses of included studies
Attention should also be paid to how *the results of a systematic review can reflect the strengths and weaknesses*, i.e. quality, of the included studies. Methods may include:

- 'quality weighting' (allocating more weight to studies of higher methodological quality enabling the poorer studies to exert less influence on the results);
- 'quality thresholds' (in which a subset of studies of higher quality inform the results of a review); and
- 'sensitivity analysis' (which explores the effects of the addition/removal of lower quality studies on the results and conclusions of a review).

These methods are further explained in Section 6.8 of the Handbook.

Qualitative research
There is a great deal of controversy surrounding the appraisal of the quality of qualitative studies.[19] Much of the debate centres on whether the concepts of quality used should be roughly the same as, parallel to, or quite different from those used to assess quantitative research.[21] Given the number of different types of qualitative study designs it would be very difficult to identify what may constitute a fatal flaw in study quality.[48] Different quality appraisal checklists may be required for each type of study design.

Recommendations

This section provides advice on quality criteria and quality checklists relevant to qualitative studies. It supplements Chapter 6 of the Handbook.

If studies of varying levels of quality are summarised together to estimate an effect size the results of the review may be biased. Authors need to be explicit about the approach used to assess quality and draw conclusions about the strength of evidence. The Centre for Reviews and Dissemination Report Number 4[4] may also be useful to authors, as the report, for example, lists quality criteria for both experimental and observational studies.

In addition to standard quality assessment authors may also choose to assess whether the intervention meets quality standards. For example, in a review of educational interventions for chronic conditions.[49] the authors rated studies according to both methodological strength and adherence to educational principles in their intervention (adapted from an educational principles rating scheme).

Quality assessment criteria

Authors need to consider the criteria to be used to assess quality at the planning stage of the review. Appraisal criteria will depend on the type of study included in the review.

a) Systematic reviews

- health-evidence.ca provides a useful tool for assessing the quality of systematic reviews (www.http://health-evidence.ca/Judge.aspx)
- The Critical Appraisal Skills program has developed a series of 10 questions to help make sense of reviews: http://www.phru.nhs.uk/Doc_Links/S.Reviews%20Appraisal%20Tool.pdf

a) RCTs for assessing effectiveness:

- A number of tools are available to assess the quality of RCTs.[50, 51] The Quality Assessment Tool for Quantitative Studies is strongly recommended (http://www.city.hamilton.on.ca/PHCS/EPHPP/). This tool was developed by the Effective Public Health Practice Project, Canada, and covers any quantitative study design and takes between 10-15 minutes to complete. A comprehensive dictionary for the assessment tool is also published on the website. This tool includes components of intervention integrity and was judged to be suitable to use in systematic reviews of effectiveness in the review by Deeks et al.[23]
- Refer to section 6 of the Handbook for more information on RCTs. Details relating to cluster RCTs are provided in section 8.11.2 of the Handbook.

b) Non-randomised studies of effectiveness:

- Authors should use the Quality Assessment Tool for Quantitative Studies described above.
- Refer to The Cochrane Non-Randomised Studies Methods Group guidelines on the inclusion of non-randomised studies in Cochrane reviews (http://www.cochrane.dk/nrsmg/).
- Methods papers are available from the Cochrane Effective Practice and Organisation of Care Group (EPOC) on interrupted time series and controlled before and after studies. (http://www.epoc.uottawa.ca/).
- The results of uncontrolled studies (also called before-and-after studies without a control group), should be treated with caution. The absence of a comparison group makes it impossible to know what would have happened without the intervention. Some of the particular problems with interpreting data from uncontrolled studies include susceptibility to problems with confounding (including seasonality) and regression to the mean.

Further sources of information on non-randomised studies include:

- Deeks et al.[23] systematically reviewed 193 quality assessment tools used to assess the quality of non-randomised studies. Six tools were judged to be potentially useful for systematic reviews. The full report can be downloaded free from the UK Health Technology Assessment Programme homepage (http://www.ncchta.org/)
- The US Agency for Healthcare Research and Quality (AHRQ)[52] systematically searched for tools and found 82 quality assessment instruments. Forty-nine of the instruments were suitable for RCTs, and 19 for observational studies (e.g. cohort studies, case control studies), although some were suitable for both.
- The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) has conducted a number of health promotion systematic using a set of structured data extraction and critical appraisal guidelines for both RCTs and non-RCTs.[53-57] Authors can code and classify interventions in terms of their characteristics (e.g.

provider, setting, media), and their evaluation (e.g. equivalence of study groups, attrition, validity and reliability of data collection and analysis).[58] The guidelines are particularly useful for authors whose conception of quality goes beyond internal validity to incorporate issues of external validity, such as generalisability, and replicability. They are available on request from the EPPI-Centre and can also be used by people writing reviews with EPPI-Centre support via specially written software (see http://eppi.ioe.ac.uk).

- Chapter 6 of 'Using Research for Effective Health Promotion'[59] describes critical appraisal methods for different research designs (experimental, observational, and qualitative).
- The Berkeley Systematic Reviews Group website has many links to quality assessment resources, including non-randomised studies (http://www.medepi.org/meta/) (http://www.medepi.net/meta/)

c) Qualitative studies:

Currently, there is no single validated checklist to use for all types of qualitative studies. Where possible, authors should use the checklists and report on their usefulness so methodological developments can occur. A set of *prima facie* criteria for assessing quality common to all qualitative research includes:

1. Method appropriate to research question
2. An explicit link to theory
3. Clearly stated aims and objectives
4. A clear description of context
5. A clear description of sample
6. A clear description of fieldwork methods
7. Some validation of data analysis
8. Inclusion of sufficient data to support interpretation

- The NHS CRD Report Number 4[4] describes a number of checklists to assist reviewers in the assessment of qualitative studies.
- A framework[21] consisting of 18 appraisal questions, is also available from (http://www.strategy.gov.uk/files/pdf/Quality_framework.pdf.
- The EPPI-Centre has developed a 12-question checklist for process evaluations (http://eppi.ioe.ac.uk/EPPIWeb/home.aspx?page=/hp/reports/phase/phase_process.htm). The EPPI Centre has also developed a 12-question checklist for appraising studies which examine people's perspectives and experiences of particular health issues or interventions at http://eppi.ioe.ac.uk/EPPIWebContent/hp/reports/health_eating02/Final_Report_web.pdf for an example of this appraisal tool.
- The Qualitative Research Methods Group and Campbell Collaboration Process Implementation Methods Group has posted a bibliography of methodological studies and discussion papers on qualitative literature, including critical appraisal of qualitative literature, on their website (http://mysite.freeserve.com/Cochrane_Qual_Method/index.htm).
- In addition, research is being undertaken to determine whether different guidelines would produce different decisions about inclusion and exclusion of papers, compared with intuitive expert judgement of quality (Economic and Social Research Council Research Methods project http://www.ccsr.ac.uk/methods/projects/posters/systematicreviews.shtml).

Accounting for the strengths and weaknesses of included studies

The Cochrane Non-Randomised Studies Methods Group suggests that quality weighting should be avoided. This is due to the lack of empirical evidence about the relative

importance of various quality dimensions of non-randomised studies. This may change with advances in methodological research in this area.

5. Theoretical framework

In the course of generating evidence about the effects of interventions theory can inform the development of interventions and frame the analysis for synthesising evidence. This section addresses the use of theory for both of these stages.

Interventions are usually based on a particular theory. However controversy remains about whether or not theory makes a difference to intervention effectiveness. As Oakley (1999) points out "the importance or unimportance of theory is unlikely to emerge unless review activity is structured to cross problem/outcome areas, and allow for the classification of interventions according to their theoretical base."

Theories relevant to health promotion seek to explain:
- individual behaviour (Stages of Change model[60], Health Belief Model[61], Theory of Reasoned Action[62] ;
- interpersonal influences (Social Learning Theory[63]; and
- activities throughout communities (community organisation theories, Organisational Change Theory, Diffusion of Innovations Theory).

These theories have been presented in an accessible framework by the Health Promotion Agency in Northern Ireland[64], and in a useful text.[65] Modern health promotion theory relates to categories of intervention or barriers and facilitators.[66, 67] These theories suggest that effectiveness is more likely to result from interventions that are multifaceted and target barriers and facilitators. Interventions which operate at the three interrelated levels: the individual (e.g. knowledge, attitudes, self-esteem); the community (e.g. social support networks, family relationships); or the wider society (e.g. social class, access to resources and services) are also more likely to be effective.

As interventions become more multifaceted, and thus more complex, it is important to reflect on the role theory has played. Parts of theories may be used in conjunction with others, a dominant theory may be used with or without others or a range of theories may be used to explain different intervention components. The impact of theory on intervention processes and impacts may not be formally collected. However, authors may reflect on the impact of theory.

In addition to examining included studies for their explicit use of theory in the development and delivery of interventions, systematic reviews can employ theoretical frameworks to analyse the active components of effective interventions. Theoretical frameworks can be used to characterise interventions in order to make them comparable in a review. More novel is the use of grounded theory applied to qualitative data to determine the sensitivity analyses appropriate to test in a meta-analysis. Examples of examining theory underpinning interventions in primary studies, and using theory for review level analysis appear in the box below.

Recommendations

This section provides advice on the roles of substantive theories relevant to public health and health promotion in the preparation of systematic reviews.

Authors should consider the following to make reviews more user-friendly:

- Background section of the review: This section should describe why, theoretically, the interventions under review might be effective. This may come from empirical evidence which may show: 1) similar interventions having an impact, or 2) identical interventions having an impact on a different population. Authors may also refer to a body of theoretical literature that justifies the possibility of effectiveness.

- In the results section authors can group interventions by their theoretical basis, eg. group all interventions based on the Stages of Change model. The studies, according to different theories, may be tabulated, combined narratively, or statistically combined. Alternatively, authors may group interventions depending on whether they seek to influence individual behaviour, interpersonal relationships, or community or structural factors. An innovative approach is to examine qualitative studies of people's views about what they think influences their lives and use these results as theories to group interventions – see box.
    - Note: Authors of primary studies may differ in how they describe the theoretical basis for their interventions. A review of stage based interventions to promote smoking cessation. Riemsma[68] found that many studies lacked information on the content of the intervention. This made it difficult to determine if, how and to what extent the Stages of Change theory was used in tailoring the intervention. Authors of reviews must choose whether they rely on what the studies say the intervention is about or whether to classify studies based on the main components of the known theory.
- It may also be useful for authors to assess whether interventions have used a Program Logic or Program Theory of Action approach[69-71] to developing, implementing and evaluating the effects of the intervention.
- Reviewers should refer to discussion and limitations sections of included studies to identify reflections of authors on impact of theory on intervention success or failure. It may also be necessary for reviewers to reflect on the impact of theory across included studies. These reflections can be reported narratively and qualifications about the rigor of these reflections should be included.
- Systematic reviews would be greatly enhanced if in the discussion attention was paid to the gaps in the theoretical coverage of interventions. For example, a large number of interventions seek to change the choices people make by focusing on single level changes (knowledge, attitude, behaviour, etc). The gaps lie in the lack of studies seeking to change the environment within which people make their choices.

Example of exploring the role of theory in effectiveness of interventions
Two systematic reviews[68, 72] of smoking cessation (one in pregnancy) combined the findings of studies where interventions were based on the Transtheoretical Model[73] of stages of change in readiness to stop smoking (pre-contemplation, contemplation, preparation and action). They both concluded that stage-based interventions are no more effective in general than interventions which do not tailor the intervention according to the stage of change.

Examples of use of theory to structure narrative reviews
A structured narrative review scoped the literature in order to structure a framework that combined Arnstein's ladder of participation[74] with other dimensions (e.g. individuals/ groups, range of forums, and single/ repeated interactions and theoretical basis of participation methods) to characterise and synthesise what has been learnt about public involvement in setting research agendas.[75]

A series of reviews that stratified studies according to the level of the intervention (individual, interpersonal, community and structural) was able to compare the effectiveness of these different approaches.[27] The effective interventions were those targeting interpersonal and community factors.

Example of use of theory for determining sensitivity analyses in meta-analysis
In a review of children and healthy eating, the meta-analysis was informed by the children's theories of what presented barriers or facilitators to their eating healthily.[57] The interventions were grouped for meta-analysis according to whether they addressed factors the children thought were important influences. This is similar to the use of grounded theory in qualitative research, whereby the theory is constructed from the empirical data.

6. Integrity of intervention

The integrity (or fidelity) of the intervention is the degree to which the intervention is implemented as planned. Authors need to ask "Were all of the specified procedures/components of the intervention actually carried out?"

Assessing the integrity of implementation is important for a variety of reasons. If integrity data are not collected, it is difficult to determine whether non-significant results are due to a poorly designed intervention or to an incomplete delivery of the specified components (Dane 1998). Authors who do not consider integrity data when claiming that an intervention is ineffective are said to me making a Type III error.[76, 77] Integrity information is particularly important in preventive interventions, which are often implemented in dynamic and complex conditions that present numerous obstacles to complete intervention delivery.[78] Integrity information will also help to determine why an intervention succeeded or failed. The key components of an intervention that were effective or ineffective will be identified through integrity data. Furthermore, integrity data reveals important information about the feasibility of an intervention in real life settings– how likely it is that the intervention can and will be implemented with integrity. If studies show that it is difficult to achieve integrity of implementation, then the program is likely to have low feasibility in practice.[79]

Five aspects of integrity/fidelity described by Dane and Schneider[78] are:

1. Adherence: the extent to which specified components of the intervention were delivered as prescribed.

2. Exposure: an index that may include any of the following: (a) the number of sessions implemented; (b) the length of each session; or (c) the frequency with which intervention components were implemented.

3. Quality of delivery: a measure of qualitative aspects of delivery that are not directly related to the implementation of the content of the intervention. This includes implementer enthusiasm, leader preparedness and training, global estimates of session effectiveness, and leader attitude towards the program.

4. Participant responsiveness: a measure of participant response to components of the intervention, which may include indicators such as levels of participation and enthusiasm.

5. Program differentiation: a manipulation check that is performed to ensure that the participants in each experimental group received only the planned interventions. Contamination may be a problem within many public health and health promotion studies where intervention and control groups often come into contact with each other.

Dane and Schneider[78] strongly recommend that all five dimensions should be measured in order to provide a comprehensive picture of intervention integrity. As Dumas et al.[80] point out policy makers and prevention researchers regularly confront the realisation that the effectiveness of different interventions is difficult, if not impossible, to compare. Therefore, policy-makers, practitioners or other groups have not always found systematic reviews useful[81-84] when intervention details are not provided in the review. Therefore, it is important in the systematic review to describe in detail what the interventions involved and what factors affected implementation.

Unfortunately, this information is not always available in reports of studies. It has been reported that only a limited number of studies disentangle the factors that ensure successful

outcomes, characterise the failure to achieve success, or attempt to document the steps involved in achieving successful implementation of complex interventions.[85, 86] Furthermore, there is debate about the factors that relate to integrity. For example, it is possible that an intervention that is well-liked by participants may not always result in effective outcomes.

Process evaluations focus on the way in which interventions are implemented. These studies may be reported within quantitative studies of effectiveness, or they may be published separately. Process evaluations examine the dynamics of the intervention and the actual implementation of intervention components. This information allows us to understand the strengths and weaknesses of an intervention. Frequently in health promotion, process evaluations are carried out throughout the intervention to provide feedback during the intervention process. This information contributes to the further development of an intervention as it evolves. To manage this dynamism, process evaluation studies utilise a flexible approach to data collection and often a variety of methods which produce a range of different types of data. Significantly, they encompass both quantitative and qualitative methods. These may include:

- the collation of routine data;
- surveys generating quantitative data;
- in-depth qualitative interviews producing narrative data;
- participant and non-participant observation;
- focus groups; and
- ethnographic fieldwork.

There is a tendency for process evaluations to focus on collecting quantitative data such as recruitment numbers or intervention dose received.[87] However, decision-making theory highlights the value of practitioners engaging in reflective practice in order to highlight critical success factors and potential reasons for intervention failure, and to contextualise research results.[88-90] This information may not be reported in primary studies. Where it is reported, it might be present in discussion or limitations sections. Reviewers are therefore encouraged to check these sections for reflections on critical success and failure factors. Collecting as much information about integrity of interventions is particularly essential for complex interventions and multi-strategic interventions.

Inevitably, given this range of data types there are also diverse approaches to analysis in process evaluations. These include the use of statistical packages as well as various approaches to the analysis of qualitative data including grounded theory approaches and software packages such as Nudist and Ethnograph.

Recommendations

This section provides advice on the collection, assessment, and synthesis of information relating to the integrity of the intervention.

Information should be elicited from process evaluations regarding factors which influence the effectiveness of interventions. The data abstraction form used in your review should contain information relating to the assessment of integrity of the intervention.

Describing studies:
In describing interventions, authors should consider ways in which they can identify what made a difference.

- Did the intervention measure key process factors of implementation? Or did the trial illuminate what the key process factors were after the intervention (as a result of the process evaluation)?
- Did interventions with high integrity thoroughly show a greater impact (as listed by Dane and Schneider[78])?

- Did the authors identify critical success or failure factors? If so, is there any indication of how these reflections were generated (i.e. qualitative data collection, group-based reflection, author reflection based on theory). This will help explore they key question 'what made a difference'.
- The critical appraisal and data extraction checklists developed by the Effective Public Health Practice Project (http://www.hamilton.ca/PHCS/EPHPP/) requires reviewers to report on many of the factors relating to intervention integrity. In addition, the EPPI-Centre has produced a checklist for reviewers to assess the coverage of a process evaluation (http://eppi.ioe.ac.uk/EPPIWeb/home.aspx?page=/hp/reports/phase/phase_process.htm).

Authors may also wish to gather evidence on the views of potential recipients of public health and health promotion interventions. These data can be gathered and analysed either quantitatively or qualitatively and the findings used to critique the included trials – *see Section 8 – Integrating qualitative and quantitative data.*

Discuss whether:

- Results of primary intervention studies correspond with results from studies investigating the recipients' views of the types of interventions studied? An example of addressing this question is highlighted in Thomas.[91]
- Potential recipients suggested factors that could be addressed by future interventions?

Examples of reviews which evaluated intervention integrity

A review of peer delivered health promotion interventions for young people[92] provides an example of data synthesis of process evaluations. Including process evaluations identified factors influencing implementation: young people viewing peer leaders as 'teachers' rather than peers, and adults undermining peer leaders' control over the content and/or organisation of their sessions. Further information on this process can be found in the NHS CRD report[4] and the original work is available on the EPPI-Centre website (http://eppi.ioe.ac.uk/EPPIWeb/home.aspx?page=/hp/reports/peer_health/peer-delivered_health_promotion_intro.htm).

A review of smoking cessation in pregnancy[72] assessed the implementation of interventions. This approach found that process evaluation of the intervention occurred in only some trials and in some of these the implementation was less than ideal. As these trials were among the largest published trials it showed that it may be inappropriate, when interventions are complex, for trial size to be taken as a surrogate measure of trial quality. It also highlighted how the transfer of an intervention from one setting to another may reduce its effectiveness if elements are changed or aspects of the materials are culturally inappropriate.

7. Heterogeneity in public health and health promotion reviews

Public health and health promotion interventions are often complex interventions. Complexity is usually due to the characteristics of the interventions, study population/s, outcomes, or other methodological issues.[93] Furthermore, complexity is introduced because the effectiveness of the interventions may be modified by the context in which it operates.[93, 94] Because of these variations, authors may expect considerable heterogeneity (i.e. differences in results) across studies and need to consider this when synthesising results.

Many of the concepts in Section 8.7 and 8.8 of the Handbook generally apply as much to public health and health promotion interventions as to the evaluation of clinical interventions. However, the differences in the former studies are likely to be greater and incorporate not only clinical differences, but historical, cultural, spatial and other differences that may affect both the delivery and impact of the interventions being reviewed.

Variability in study populations, interventions and settings
Variability in participants among studies may be particularly great as the populations in public health and health promotion studies will often not be selected according to the same pre-defined criteria. Intervention variability may also be considerable and this should be examined and described in detail. This information is likely to be found in process evaluations. The content of complex interventions may also vary among specific settings or populations. This variability may be intentional as interventions are tailored to local needs (including population determinants of health such as race, gender, socio-economic position).

Variability in outcomes
In clinical interventions variations in outcomes is termed clinical heterogeneity. This also exists in public health research. However, given the longer causal chains for many public health interventions, proximal/immediate, intermediate, and distal outcomes may be reported.

Variability in study designs
Methodological diversity may be even more common in public health and health promotion reviews than in clinical reviews. In some areas of public health RCTs or outcome evaluations may be uncommon. In a study of articles published in three leading health promotion journals the number of published RCTs was 10% or less.[95]

In clinical interventions the main potential sources of variation may be known. In that case heterogeneity can be explored by subgroup analysis, based on existing theories about how the intervention works for various groups. By comparison, theories about mechanisms and interactions may be less fully developed for many social and public health interventions, and so the exploring and interpreting heterogeneity may be much more complex. It may, for example, be more difficult to anticipate the main sources of heterogeneity *a priori*.

Analysis
Variation in intervention effects may be detected statistically using a heterogeneity test. If the test reveals no statistical heterogeneity it should not be assumed that a meta-analysis is appropriate. This is because of the number of potential sources of variation (already described). Similar effect sizes may be obtained from studies which are conceptually very different. In reviews of health promotion and public health intervention the author needs to make the case for meta-analysis before proceeding.

Subgroup analyses

The same cautions about subgroup analyses apply to public health and health promotion interventions as apply to interventions in clinical populations. However, there is probably greater interest among users of public health and health promotion systematic reviews in the range of effects in subgroups. Users are often interested, for example, in examining variation in effects among different social (e.g., ethnic, demographic, and educational) groups, or among different contexts within which the intervention has been delivered.[94] While such sources of variation are clearly of relevance to practitioners and policy makers, the interpretation of such findings requires caution. Evidence from such analyses may provide only indicative evidence of effectiveness, and provide an indication of where further primary research may be targeted.

Recommendations

This section provides information to reviewers on the potential sources of heterogeneity within their reviews.

Authors should consider the likely sources of heterogeneity as described above, and consider these as they synthesise and analyse the results, either narratively or in meta-analysis.

8. Integrating qualitative and quantitative studies

The value of including both qualitative and quantitative studies within a public health or health promotion review has been outlined in previous sections. This section provides recommendations to authors on methods to integrate both types of data.

Recommendations
This section provides advice on synthesising qualitative and quantitative studies.

Synthesis

- Several invaluable sources describing a range of approaches to synthesising qualitative and quantitative evidence are available: including: "Integrative approaches to qualitative and quantitative evidence"[48]; "Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field"[96]; and "Systematic Reviews in the Social Sciences: A Practical Guide".[12]
- The EPPI-Centre has developed synthesis methods to combine the findings of intervention effectiveness studies with studies of people's views about what influences their behaviour and health. This necessarily involved combining the findings from studies with diverse designs presenting qualitative and quantitative data.
- The first approach has been employed in a series of reviews addressing the barriers and facilitators to young people's healthy behaviour.[27] Dahlgren and Whitehead's[97] multilayered model of social determinants of health provided a framework for juxtaposing the findings of intervention evaluations and views studies addressing influences at the level of the individual, family and friends, community and wider society. This arrangement facilitated the comparison of people's views about what influences their behaviour and health with the components of complex interventions seeking to influence their behaviour and health, whatever the study designs and type of data. This comparison led to conclusions about felt needs, promising interventions (matched by people's views, but not adequately evaluated for their effects), and evidence of effectiveness and appropriateness (from good quality trials and people's views respectively).
- The second approach[91] involved conducting three types of syntheses in the same review:

    1) a statistical meta-analysis to pool trials of interventions tackling a particular problem (or a narrative synthesis when meta-analysis is not appropriate or possible);
    2) a synthesis of studies examining people's perspectives or experiences of that problem using qualitative analysis ('views' studies); and
    3) a 'mixed methods' synthesis bringing the products of 1) and 2) together.

  These developments have been driven by particular review questions rather than methodology. 'Users' of reviews want to know about the effects of interventions **and** which interventions will be most appropriate and relevant to people. The methods for each of the three syntheses are described in brief below, with full details reported elsewhere.[28, 91]

  Synthesis 1) Effectiveness synthesis for trials
  Effect sizes from good quality trials are extracted and, if appropriate, pooled using statistical meta-analysis. Heterogeneity is explored statistically by carrying out sub-group analyses on a range of categories specified in advance (e.g. study quality, study design, setting and type of intervention).

  Synthesis 2) Qualitative synthesis for 'views' studies

The textual findings from 'views' studies are copied verbatim and entered into a software package to aid qualitative analysis. Two or more authors undertake a thematic analysis on this data. Themes are descriptive and stay close to the data, building up a picture of the range and depth of people's perspectives and experiences of a particular health issue. The content of the descriptive themes are then considered in the light of the relevant review question (e.g. what helps and what stops children eating fruit and vegetables?). The results can be used to generate implications for intervention development. The products of this kind of synthesis can also be considered as 'theories' about which interventions might work. These theories are grounded in people's own understandings about their lives and health. This method has much in common with the work of others who have emphasised the theory building potential of synthesis.[28]

Synthesis 3) A 'mixed methods' synthesis
The implications for interventions produced in Synthesis 2 are put together with the interventions which have been evaluated in Synthesis 1. This will identify any matches, miss-matches and gaps. Gaps are used for recommending what kinds of interventions need to be developed and tested. The effect sizes from interventions which matched implications for interventions derived from people's views can be compared to those which do not, using sub-group analysis. This provides a way to highlight which types of interventions are both effective and appropriate. These methods integrate 'quantitative' estimates of benefit and harm with 'qualitative' understanding from people's lives, whilst preserving the unique contribution of each.[91] This method is unlike Bayesian methods which is another approach to combining 'qualitative' and 'quantitative' studies within systematic reviews which translates textual data into numerical data.

- Guidance on the conduct of narrative synthesis – an approach that can be used to integrate qualitative and quantitative data - has been produced by a UK research team with funding from the ESRC.[98] The guidance, which is based on a systematic search of methodological literature, is not prescriptive. It is provides a general framework for the conduct of narrative synthesis consisting of four elements: 1. Developing a theoretical model of how the interventions work, why and for whom; 2.Developing a preliminary synthesis; 3. Exploring relationships in the data; and 4. Assessing the robustness of the synthesis product. The guidance also describes a number of specific tools and techniques identified during the methodological review that could be used in each element of the narrative synthesis process. The results of using the guidance in two demonstration reviews are provided: one focusing on evidence on effects the other on evidence of factors impacting on implementation.

- The Cochrane Qualitative Research Methods Group is producing preliminary guidance and this will evolve as results of methodological research become available. Methodological approaches to the systematic review of qualitative research findings a currently the subject of debate and there are a number of approaches that may be considered appropriate.[26] Authors aiming to include qualitative studies in their systematic review should familiarise themselves with existing methodological work in this area (for example Nobblit and Hare[99]; Popay and Roen[26]; Popay[100]; Harden[28] 2004; Pearson[101]; Oliver et al.[27]; Pearson et al.[102]; Petticrew and Roberts[12]; Britten et al.[103]) and should do so in a way that will contribute to the methodological agenda, eg. comparing the results of an exhaustive versus theoretical approach to searching, utilising different frameworks for appraisal, synthesis, etc.

9. Ethics, equity and inequalities

Public health and health promotion interventions have the potential to improve the health of populations. Systematic reviews can determine the effectiveness of these interventions in achieving their desired outcomes. There are some specific ethical considerations that should be taken into account in reviewing the effectiveness of HP and PH interventions. Standardly, effectiveness is measured in terms of the total number/population who benefit from the intervention. This consequentialist approach takes no account of the distribution of benefits[104], and therefore does not address issues of health equity. Overall improvements in health behaviours or health outcomes may actually mask the differences in health outcomes between groups.[105] Interventions that work for those in the middle and upper socio-economic positions may not be as effective for those who are disadvantaged. Even well-intentioned interventions may actually increase inequalities. Health differentials that exist between groups may be due to complex interactions between many of the factors relating to disadvantage.[106]

Systematic reviews of HPPH interventions have the potential to investigate differential outcomes for groups with varying levels of disadvantage. This can inform strategies aimed at reducing health inequalities and health inequities. Health inequalities are "differences, variations, and disparities in the health achievements of individuals and groups".[107]

These variations may not always reflect inequity in health. For example, young adults are expected to be healthier than the elderly population. Health equity is an ethical concept referring to the fairness or unfairness of particular health inequalities. The International Society for Equity in Health defines equity in health as: "the absence of potentially remediable, systematic differences in one or more aspects of health status across socially, economically, demographically, or geographically defined populations or subgroups".[108] Turning this around, health inequities are those health inequalities that are unfair or unjust, or stem from some kind of injustice.[107] Reviews of effectiveness of HP and PH interventions can provide information about the effects of interventions on health inequalities. This information can then be used to address health inequities.

Systematic reviews rely upon there being sufficient detail in trial data to allow for identification of relevant subgroups for analysis in relation to health inequalities. This requires attention not only to levels of benefit or harm, but also looking at the distributions of these; who is benefiting, who is harmed, who is excluded?

Reviews of the effectiveness of interventions in relation to health inequalities require three components for calculation of effectiveness for reducing health inequalities:

- a valid measure of health status (or change in health status);
- a measure of socio-economic position (or disadvantage); and
- a statistical method for summarising the magnitude of health differences between people in different groups.

To date, very few systematic reviews have focused on the effect of interventions on inequalities in health.[109] Historically Cochrane reviews have paid poor attention to equity issues[110], although this is now starting to change, with the establishment of a Campbell and Cochrane Health Equity Field. The Field aims to identify interventions that improve health status of the poor and reduce health inequalities. They will be working specifically on systematic reviews of effective interventions and in doing so, develop methods for applying an equity lens. A recent study[111] found that Cochrane reviews (relating to tobacco control) rarely presented information on outcomes stratified by socioeconomic position and that the differences between groups were often not considered in

the design of the review. In a recent Lancet editorial, Tugwell and his colleagues[112] reported on a similar study of a random sample of 95 Cochrane Reviews and their constituent primary studies. Here, only 1 Cochrane review out of the 95 sampled reported differences in treatment effect by socioeconomic status (SES).[112] This also highlights the fact that most primary studies have not reported, or sought to establish, how the effects of interventions are distributed between groups.[105, 106, 111-113] To confound matters, studies are often underpowered to examine the existence of differences between groups.[105]

Conducting reviews addressing inequalities is complicated not only by limited collection of information about differences between groups, but also by the fact that .disadvantaged groups are often excluded from the commissioning and design of research, and from participation in trials. This means that the evidence generated may be limited in its applicability to those groups.[114] Britton [115] (1998) found that participants in four RCTs aimed at promoting health or preventing disease were more likely to be younger, of higher social status (in terms of income, housing, education and car ownership), and to believe in and adopt a 'healthy lifestyle' (eg. not smoking, taking regular exercise) than non-participants. A further problem relates to the outcomes chosen in the trial as markers of effectiveness, as these may not include outcomes that are of significance to disadvantaged populations. Assessments of effectiveness often include implicit evaluations, such as estimates of the worth of some of the factors such as likely side effects, consequences of refusal, the availability of other treatments or prevention options, or economic factors.[116] These assessments may be made with little consideration for the implications for disadvantaged groups. Finally, there can be procedural problems. McGowan[117] for example, found that the terms used to describe inequalities in the indexing of the selected databases was limiting.

Despite these barriers, systematic reviews can play an important role in raising awareness of health inequalities. The first Cochrane review to specifically address inequalities was by Kristjansson.[112] This now updated review examined the effectiveness of school feeding programs for improving the physical, psychological, and social health of disadvantaged children and for reducing socio-economic inequalities in health. Reviews serve an important function in drawing attention to gaps in knowledge and practice and by making recommendations for improved practice.[106] For example, the 2005 review of interventions to increase participation in sport by Jackson et al[118] identified a lack of controlled studies assessing the effects of interventions to increase participation in sport. This review recommended specifically that interventions should include socioeconomic differentials to address the known lower levels of participation in sport of people from lower socioeconomic groups. Two recent Cochrane protocols investigating peer based interventions for HIV-infected women, and peer support for people with chronic disease both have objectives that explicitly address health inequalities.[119, 120] The Cochrane Health Equity Field have identified a number of equity relevant reviews:

equity.cochrane.org/Files/Cochrane%20Health%20Equity%20Field%20and%20Campbell%20Equity%20Methods%20Group%20Newsletter%20-%20Vol%201,%20Issue%201.pdf

Recommendations

This section provides advice on the conduct of systematic reviews examining inequalities as an outcome.

- Consider the ethical implications of every decision made throughout the review process. Ethical issues may arise during many stages of the review, including decisions relating to: the topic of the review; who is involved throughout the review; and which interventions and outcomes to include.

Defining inequalities

Decide which indicators of disadvantage or status are relevant to the review. Disadvantage may be considered in terms of place of residence, race/ethnicity, occupation, gender, religion, education, socio-economic position (SES) and social capital, known by the PROGRESS acronym[121] (Evans 2003). Authors should carefully consider which of these are relevant to their population of interest; data will then be extracted by these factors. The Cochrane Health Equity Field are working on definitions of equity as relevant to Cochrane reviews: http://www.equity.cochrane.org/en/index.html

Defining effectiveness

Kristjansson[122] describes an effective intervention for reducing health inequalities as one which is:

- more effective for people in lower SES;

A 'potentially' effective intervention is one which is:

- equally effective across the socioeconomic spectrum and may therefore reduce health inequalities due to the prevalence of health problems among the disadvantaged being greater.

The judgement becomes more difficult when the intervention is targeted only at lower SES groups. In the review of school feeding problems[122], effective interventions aimed solely at disadvantaged children were labelled as 'potentially' effective in reducing socio-economic inequalities in health.

It is impossible to determine differential effectiveness if studies comprise mixed SES groups but do not include results that can be broken down by SES grouping.

Finding studies which examine inequalities

To locate studies examining inequalities:

- conduct broad searches for studies; and
- contact authors for further information regarding socio-economic data.

This latter task may be necessary because primary studies often fail to present information on the socio-economic composition of participants.[106, 111, 113] The school feeding review found that less than half of the included studies came from formal literature searches. The remaining studies came from handsearching journals or references in reviews and other primary studies. Furthermore, in the same review, when study participants came from mixed SES backgrounds, data was rarely reported by SES and some authors actually controlled for SES in their analyses.

Analysis issues

- determine the number of subgroup analyses at the beginning of the review. Further information on subgroup analyses can be found in section 8.0 of the Handbook.

10. Sustainability

Sustainability refers to the continuation of an intervention or program and/or its outcomes.[123, 124] Sustainability should be an important consideration in systematic reviews. Attention to the long-term viability of health interventions is likely to increase as policy makers, practitioners and funders become increasingly concerned with allocating scarce resources effectively and efficiently.[123] Users of reviews are often interested in knowing whether the health benefits (e.g. reductions in specific diseases or improvements in health) from an intervention are going to be sustained beyond the life of the intervention.

Public health and health promotion interventions are usually complex and changes in health behaviours are often slow and difficult. Changes in morbidity and mortality occur over an even longer time period and generally require interventions to be effective over an extended time.[123] This means that interventions (eg. educational messages) may need to remain in place for new generations of individuals to be exposed to them.[123] Swerissen and Crisp[124] believe that health promotion interventions are more likely to produce sustainable outcomes if they address appropriate levels of social organisation in seeking to achieve positive outcomes.

However, many routine program evaluations have an emphasis on process with little or no focus on outcomes.[123] Even where there are data on outcomes there is often little information on the extent to which the intervention is sustained. The follow-up period in studies also limits the extent to which long term outcomes can be assessed. Careful consideration in reviews of how previous studies have addressed (or failed to address) issues of sustainability will increase our understanding in this area. It should hopefully stimulate improved design for assessment of sustainability in future studies.

It should be noted that a sustained or sustainable program does not necessarily result in sustained outcomes and that not all interventions need to be sustained in order to be useful or effective.[123]

Shediac-Rizkallah and Bone present a useful framework for addressing sustainability.[123] In this framework key aspects of program sustainability are defined as 1) maintenance of health benefits from the program; 2) institutionalisation of a program within an organisation; and 3) capacity building in the community. Key factors influencing sustainability are defined as 1) factors in the broader environment; 2) factors within the organisational setting; and 3) project design and implementation factors.

Recommendations

This section provides advice to reviewers when considering issues of sustainability.

- Consider whether the sustainability of intervention outcomes is relevant to the objectives of the intervention. If this is the case, consider what outcomes have (or should have) been measured, over what period, and what is the pattern of outcomes over time.

- Information should be sought on both contextual factors and project characteristics that may explain the extent of sustainability. Where sustainability of outcomes has not been measured, explore the potential of the intervention outcomes to be sustained. Three frameworks have been identified to assess the sustainability. Review these in relation to the interests and focus of the review:

    1. The following five factors identified by Bossert[125]:

- The economic and political variables surrounding the implementation and evaluation of the intervention;
- The strength of the institution implementing the intervention;
- The full integration of activities into existing programs/services/curriculum/etc;
- Whether the program includes a strong training component (capacity building); and
- Community involvement/participation in the program.

2. The framework developed by Swerissen and Crisp[124] guides decisions about the likely sustainability of programs and effects at different levels of social organisation. This framework outlines the relationships between intervention level, strategies and the likely sustainability of programs and effects.

3. The Centre for Health Promotion, University of Toronto, has also produced a document outlining the four integrated components of sustainability.[126]

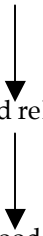
- In examining the quality of studies included in a review some of the particular features relevant to sustainability that should be considered (in addition to general aspects of study quality) are:
  - a broad conceptualisation of potential influences on sustainability (preferably using a standard framework[127] such as that proposed by Shediak-Rizkallah and Bone[123];
  - the extent to which data on program activity, program outcomes and on factors that might influence program sustainability are collected prospectively over the course of the program[128];
  - the extent to which the program is (perhaps appropriately) modified over time, and the extent to which the program may be considered a new or different program as a result of modification[128];
  - the extent to which quality as well as quantity of program activity is sustained during the process of institutionalisation[128];
  - the potential effect of natural progression of disease and aging of cohorts on the potential for maintenance of health benefits.[128]

11. Context

The type of interventions implemented, and their subsequent success or failure are highly dependent on the social, economic and political context in which they are developed and implemented.[129]

Example[129]:

Media-based intervention to promote the consumption of fruit and vegetables (F & V)



Dependent on the following contextual factors:

the availability and relative price of F & V

Dependent on the following contextual factors:

geographic factors, food distribution systems and retail prices.

A problem in reviewing public health and health promotion interventions is how to disentangle "intervention" effects from effects that should be more appropriately called *program x context interactions*.[130] Traditionally, outcomes have been attributed to the intervention. However, the outcomes noted in studies may in fact be due to pre-existing factors of the context into which the intervention was introduced. Hence, context should be considered and measured as an effect modifier in trials.[130, 131] Such contextual factors might relate to aspects of the program's "*host organisation*". Broader aspects of context might include aspects of the *system within which the host organisation operates*. Some investigators would also argue that context factors also pertain to *the characteristics of the target group or population*. For many years these aspects have been acknowledged (but not clearly specified) when decision makers have argued that results of evidence reviews from other countries do not apply in their own country.

Use of the term "context evaluation" became more prevalent in health promotion after the review by Israel and colleagues.[132]. However the systematic investigation of context level interactions as part of the design of RCTs of community or organisational-level interventions is almost unknown.[130, 131] Instead, aspects of context have been explored as part of the more developed field of sustainability research or research on program instutionalisation – *see Section 11 Sustainability*.[123, 133-137] A related and growing multidisciplinary research field is the implementation and integration sciences which are leading researchers more into the complexity of the change processes that interventions represent – *see Integrity of Interventions section*.[138-140] At the present time, quantitative studies lag behind qualitative analyses of context.

Systematically disentangling context effects from intervention effects in anything other than a study set up for this purpose is extremely difficult. Whilst some programs have been transferred from one context to another and benefits have been observed[141] others have not[72]. Cluster randomised designs may be expected (in theory) to even out important aspects of context, provided that the sample size is sufficient. However, few investigators at present measure or report on any aspect of context that might be important to our assessment.

We also note recent calls for a greater focus on external validity.[142, 143] Working together, journal editors and researchers are encouraging more examination of and reporting on aspects of intervention context. This will affect how intervention studies are interpreted (appreciation of interaction effects) and how findings are generalised.

Recommendations

This section provides some guidance to reviewers for considering issues of context in the assessment of effectiveness.

- Take caution when making generalisations from one context to another.

- Report on the presence (or otherwise) of context-related information in intervention studies, where this information is available:[130]
    - aspects of the host organisation and staff, eg, number, experience, morale, expertise of staff, competing priorities to the staff's attention, the organisation's history of innovation, size of the organisation, the status of the program in the organisation, the resources made available to the program;
    - aspects of the system, eg, payment and fee structures for services, reward structures, degrees of specialisation in service delivery; and
    - characteristics of the target population (eg, socioeconomic, cultural, literacy levels, place of residence).
    - Some researchers of program-context interaction are extending their investigation to the inter-organisational network within which the host organisation operates.[130]

This will start to chart the systematic gaps in studies and draw attention to critical factors that should be reported. It may alert investigators to the need to qualify their statements about "intervention" effects. It may also spawn more combined methods research (qualitative and quantitative) that may alleviate this problem in future.

12. Applicability

Applicability needs to be considered when deciding how to translate the findings of a given study or review to a specific population, intervention, or setting. Applicability is essentially synonymous with *external validity* and *generalisability*. Applicability is preferred to other terms because the central issue is how to *apply* the results of a study or review to another situation. *Transferability* or the *potential for translation*, are similar and appropriate terms. Applicability is closely related to integrity, context, and sustainability as discussed in previous sections.

Applicability is different from efficacy or effectiveness. Efficacy studies provide information about whether an intervention yields a beneficial outcome under ideal conditions. Effectiveness is determined under more common (real-world) conditions, where it may be more difficult to apply the intervention. The applicability of both efficacy and effectiveness studies must be considered when translating study results to a new setting.

Internal validity, or the degree to which the effects of an intervention can be attributed to the intervention being assessed, must be considered separately from issues of applicability. Internal validity applies only to the validity within the specific situation and confines of a specific study. When internal validity is high, one can be confident that the results presented for that study are due to the intervention and not to other (confounding) factors. Applicability relates to how one might extrapolate the results of a specific study or group of studies to different situations. Thus, a study may have high internal validity, but may not be applicable to the needs of the user.

Applicability is a key part of the process of summarising evidence, since the goal is to identify interventions that are likely to be effective in different settings. Summarizing the results of multiple studies, conducted among different populations and in different settings, is, in itself, a test of the applicability of findings. When study findings are replicated in a variety of circumstances, confidence that the results are transferable is reinforced. On the other hand, when effects vary by population, setting, or intervention features, the spectrum of circumstances to which the evidence is likely to be applicable (or not) is better understood.

Systematic reviews of public health and health promotion interventions encompass several issues that make the process of determining applicability even more complex than in the clinical trials literature. First, a number of public health interventions do not involve randomization. Although not an inherent characteristic of non-randomized designs, these studies may have less well-defined inclusion criteria, settings, and interventions, making determinations of applicability more difficult. Then again, results from randomized controlled trials may be less generalisable due to unrepresentative providers of the intervention or study participants not being typical of the target group.[17] Second, public health and health promotion interventions tend to have multiple components. This makes it difficult to 1) determine what specific intervention component had the noted effect or 2) assess the synergy between components. Third, in community interventions, implementation and adherence may be much more difficult to achieve and to measure. This also makes it harder to interpret and apply the findings. Fourth, in public health and health promotion interventions the underlying socio-cultural characteristics of communities are complex and difficult to measure. Thus it is difficult to define to whom and to what degree the intervention was applied, complicating determinations of applicability. On the other hand, this heterogeneity may increase applicability, as the original populations, settings, and interventions may be quite diverse, and increasing the likelihood that the evidence can be applied broadly.

Authors are ideally positioned to summarise the various aspects of the evidence that are relevant to potential users. This enables users to compare their situation or setting to that presented in the review and note the similarities and differences. Users can then be explicit about the relationship between the body of evidence and their specific situation.

Recommendations

This section provides guidance to reviewers for considering the applicability of the results of their systematic review.

Key characteristics should be summarised for individual studies as well as for the body of evidence. Table 1. provides a list of characteristics to include. This list of characteristics can be tailored to the focus of the specific review.

Table 1. Characteristics for individual studies to be included in reviews

Applicability	What is the spectrum of circumstances -- population, intervention, and setting -- to which the evidence is pertinent, and what important variations in effect exist across different circumstances?
Relevance	Are the outcomes noted in the review relevant to the user of the study results? In some cases the outcomes may be too proximal (e.g., intermediate outcomes such as changes in self-reported dietary intake) and the linkages to distal health and quality of life outcomes (e.g., morbidity and mortality) unclear. If the study involves a comparison group, how does the nature of the comparison condition apply to the current circumstance?
Appropriateness	This encompasses value judgments. An intervention might be potentially applicable, relevant, and feasible, but the values of the community may not support the intervention.
Feasibility	Can the intervention can be replicated in a given setting. This includes cost as well as such non-monetary resources as expertise, training required for intervention staff, interest, and cultural considerations.
Adverse effects	Attention must be given to the balance of positive and negative (adverse) effects, and consider opportunity costs in choosing one course of action rather than another.
Equitability	Does the intervention distribute resources fairly and does it reduce health inequalities?
Sustainability	Is infrastructure in place, along with ongoing resources and incentives, to maintain an intervention? This highlights the importance of considering the short-term versus long-term benefits of an intervention.

The following questions may also be useful in assessing the applicability and transferability of interventions into policy and practice contexts[144]:

Applicability

- Does the **political environment** of the local society allow this intervention to be implemented?
- Is there any political barrier to implementing this intervention?
- Would the general public and the targeted (sub) population accept this intervention? Does any aspect of the intervention go against local **social norms**? Is it ethically acceptable?
- Can the contents of the intervention be tailored to suit the local culture?
- Are the essential **resources** for implementing this intervention available in the local setting? (a list of essential resources may help to answer this question)
- Does the target population in the local setting have a sufficient **educational** level to comprehend the contents of the intervention?

- Which organisation will be responsible for the provision of this intervention in the local setting?
- Is there any possible barrier to implementing this intervention due to the **structure of that organisation**?
- Does the provider of the intervention in the local setting have the **skill** to deliver this intervention? If not will training be available?

Transferability

- What is the **baseline prevalence** of the health problem of interest in the local setting? What us the difference in prevalence between the study setting and the local setting?
- Are the **characteristics of the target population** comparable between the study setting and the local setting? With regards to the particular aspects that will be addressed in the intervention is it possible that the characteristics of the target population, such as ethnicity, socioeconomic status, educational level etc will have an impact on the effectiveness of the intervention?
- Is the **capacity to implement** the intervention comparable between the study setting in such matters as political environment, social acceptability, resources, organisational structure and the skills of the local providers?

When a body of evidence is synthesised in total for a review, a summary table can be developed with the relevant characteristics relating to applicability. The user can compare their situation to the profile of individual studies or the body of evidence, facilitating conclusions about potential applicability. The user can also select interventions that match most closely the user's values and interests with respect to these characteristics (e.g., minimize potential for adverse effects, maximise equity in a population, etc.). .

References

1. Jackson, N. and E. Waters, *The challenges of systematically reviewing public health interventions.* J Public Health (Oxf), 2004. **26**(3): p. 303-7.
2. Rees, R., et al., *Use of an advisory group to ensure relevance: reflections on participation of stakeholders in a review of sexual health promotion for men who have sex with men (MSM)*, in *XII Cochrane Colloquium*. 2004: Ottawa, Canada.
3. Thomas, B.H., et al., *A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions.* Worldviews Evid Based Nurs, 2004. **1**(3): p. 176-84.
4. National Health Service: Centre for Reviews and Dissemination. *Undertaking systematic reviews of research on effectiveness. CRD's guidance for those carrying out or commissioning reviews. CRD Report Number 4 (2nd Edition)*. 2001 [cited 2007 19 September]; Available from: http://www.york.ac.uk/inst/crd/report4.htm.
5. Brownson, R., et al., *Evidence-based public health*. 2003, USA: Oxford University Press.
6. Steel, R. *Involving marginalised and vulnerable groups in research: a discussion document. Consumers in NHS Research.* 2001 [cited 2007 19 September]; Available from: http://www.invo.org.uk/pdfs/Involving%20Marginalised%20and%20Vull Groups%20in%20Researchver2.pdf.
7. Oliver, S., *Exploring lay perspectives on questions of effectievenss*, in *Non-random reflections on health services research*, A. Maynard and I. Chalmers, Editors. 1997, BMJ Publishing Group: London. p. 272-291.
8. Oakley, A. *An infrastructure for assessing social and educational interventions: the same or different?* . Background paper for the meeting at The School of Public Policy 15-16 July 1999 [cited 2007 19 September]; Available from: http://www.ucl.ac.uk/spp/download/publications/Annexe4.pdf.
9. Booth, A. and A. Fry-Smith, *Developing the research question*, in *etext on Health Technology Assessment (HTA) information resources*, Topfer LA and Auston I, Editors. 2004, National Information Center on Health Services Research and Health Care Technology USA.
10. Cochrane Open Learning Material. *Cochrane collaboration open learning material for reviewers. Version 1.1*. 2002 [cited 2006 28 February]; Available from: www.cochrane-net.org/openlearning.
11. The Cochrane Collaboration. *The Cochrane Collaboration Manual, Issue 2 2005*. 2005 [cited 23 February 2005]; Available from: www.cochrane.org/admin/manual.htm.
12. Petticrew, M. and H. Roberts, *Systematic reviews in the social sciences: a practical guide*. 2005, Oxford: Blackwell Publishing.
13. Richards, T. and J. Tumwine, *Poor countries make the best teachers: discuss.* Bmj, 2004. **329**(7475): p. 1113-4.
14. Hanley, B., et al., *Involving consumers in research and development in the NHS: briefing notes for researchers*. 2000, Help for Health Trust: Winchester.

15. Petticrew, M. and H. Roberts, *Evidence, hierarchies, and typologies: horses for courses.* J Epidemiol Community Health, 2003. **57**(7): p. 527-9.
16. Glasziou, P., J.P. Vandenbroucke, and I. Chalmers, *Assessing the quality of research.* Bmj, 2004. **328**(7430): p. 39-41.
17. Black, N., *Why we need observational studies to evaluate the effectiveness of health care.* Bmj, 1996. **312**(7040): p. 1215-8.
18. Donner, A. and N. Klar, *Pitfalls of and controversies in cluster randomization trials.* Am J Public Health, 2004. **94**(3): p. 416-22.
19. Popay, J., A. Rogers, and G. Williams, *Rationale and standards for the systematic review of qualitative literature in health services research.* Qual Health Res, 1998. **8**(3): p. 341-51.
20. Dixon-Woods, M. and R. Fitzpatrick, *Qualitative research in systematic reviews. Has established a place for itself.* Bmj, 2001. **323**(7316): p. 765-6.
21. Spencer, L., et al., *Quality in qualitative evaluation: a framework for assessing research evidence*. 2003, Government Chief Social Researcher's Office. Crown Copyright: UK.
22. Berkwits, M., *From practice to research: the case for criticism in an age of evidence.* Soc Sci Med, 1998. **47**(10): p. 1539-45.
23. Deeks, J.J., et al., *Evaluating non-randomised intervention studies.* Health Technol Assess, 2003. **7**(27): p. iii-x, 1-173.
24. Serra, C., et al., *Interventions for preventing tobacco smoking in public places.* Cochrane Database Syst Rev, 2000(3): p. CD001294.
25. Stead, L.F. and T. Lancaster, *Interventions for preventing tobacco sales to minors.* Cochrane Database Syst Rev, 2005(1): p. CD001497.
26. Popay, J. and K. Roen, *Synthesis of evidence from research using diverse study designs.* Health Care Reports, 2003. **1**(1): p. 1-25.
27. Oliver, S., et al., *An emerging framework for including different types of evidence in systematic reviews for public policy.* Evaluation, the International Journal of Theory, Research and Practice, 2005. **11**(4): p. 428-446.
28. Harden, A., et al., *Applying systematic review methods to studies of people's views: an example from public health research.* J Epidemiol Community Health, 2004. **58**(9): p. 794-800.
29. Dixon-Woods, M., et al., *How can systematic reviews incorporate qualitative research? A critical perspective.* Qualitative Research, 2006. **6**: p. 27 - 44.
30. Peersman, G. and A. Oakley, *Learning from research*, in *Using research for effective health promotion*, S. Oliver and G. Peersman, Editors. 2001, Open University Press: Buckingham.
31. Beahler, C.C., J.J. Sundheim, and N.I. Trapp, *Information retrieval in systematic reviews: challenges in the public health arena.* Am J Prev Med, 2000. **18**(4 Suppl): p. 6-10.
32. Grayson, L. and A. Gomersall, *Working paper 19*. 2003, EXRC UK Centre for Evidence Based Policy and Practice.
33. Casteel, C. and C. Peek-Asa, *Effectiveness of crime prevention through environmental design (CPTED) in reducing robberies.* Am J Prev Med, 2000. **18**(4 Suppl): p. 99-115.

34. Ogilvie, D., et al., *Systematic reviews of health effects of social interventions: 1. Finding the evidence: how far should you go?* J Epidemiol Community Health, 2005. **59**(9): p. 804-8.
35. Dixon-Woods, M., R. Fitzpatrick, and K. Roberts, *Including qualitative research in systematic reviews: opportunities and problems.* J Eval Clin Pract, 2001. **7**(2): p. 125-33.
36. Evans, D., *Database searches for qualitative research.* J Med Libr Assoc, 2002. **90**: p. 290-3.
37. Shaw, R., et al., *Finding qualitative research: an evaluation of search strategies.* BMC Medical Research Methodology, 2004. **4**: p. 5.
38. Booth, A. *Cochrane or cock-eyed? How should we conduct systematic reviews of qualitative research?* in *Qualitative Evidence-based Practice Conference: Taking a Critical Stance*. 2001. Coventry University, UK.
39. Harden, A., *Finding research evidence: systematic searching*, in *Using research for effective health promotion*, S. Oliver and G. Peersman, Editors. 2001, Open University Press: Buckingham. p. 47-68.
40. Cochrane Health Promotion and Public Health Field. *Handsearchers' training materials*. 2005 [cited 2007; Available from: http://www.ph.cochrane.org/Files/Website%20Documents/Cochrane_HPPHField_Handsearching_Materials.pdf.
41. Lamar Soutter Library. *The Lamar Soutter Library*. 2004 [cited 2005 March]; Available from: http://library.umassed.edu/ebpph.
42. Core Public Health Journal Project. *Core Public Health Journal Project (Ver 0.9)*. 2004 [cited 2005 18 March]; Available from: www.med.yale.edu/eph/library/phjournals.
43. Armstrong, R., et al., *It's in your hands: the value of handsearching in conducting systematic reviews of public health interventions.* J Public Health (Oxf), 2005. **27**(4): p. 388-91.
44. Chalmers, T.C., et al., *Bias in treatment assignment in controlled clinical trials.* N Engl J Med, 1983. **309**(22): p. 1358-61.
45. Juni, P., D.G. Altman, and M. Egger, *Systematic reviews in health care: Assessing the quality of controlled clinical trials.* Bmj, 2001. **323**(7303): p. 42-6.
46. Moher, D., et al., *Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses?* Lancet, 1998. **352**(9128): p. 609-13.
47. Schulz, K.F., et al., *Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials.* Jama, 1995. **273**(5): p. 408-12.
48. Dixon-Woods, M., et al., *Integrative approaches to qualitative and quantitative evidence*. 2004, Health Develoment Agency: London.
49. Mullen, P.D., L.W. Green, and G.S. Persinger, *Clinical trials of patient education for chronic conditions: a comparative meta-analysis of intervention types.* Prev Med, 1985. **14**(6): p. 753-81.
50. Moher, D., et al., *Assessing the quality of reports of randomised trials: implications for the conduct of meta-analyses.* Health Technol Assess, 1999. **3**(12): p. i-iv, 1-98.
51. Moher, D., et al., *Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists.* Control Clin Trials, 1995. **16**(1): p. 62-73.

52. West, S.K., V., et al., *Systems to rate the strength of scientific evidence. Evidence report/technology assessment 47(Prepared by the Research Triagle Institute-university of North Carolina Evidence-based Practice Centre under Contract No. 290-92-0011)*. 2002, Agency for Healthcare Research and Quality: Rockville, MD.
53. Brunton, G., et al., *Children and physical activity: a systematic review of barriers and facilitators*. 2003, EPPI-Centre, Social Science Research Unit, Institute of Education, University of London: London.
54. Harden, A., et al., *Young people and mental health: a systematic review of barriers and facilitatiors*. 2001, EPPI-Centre, Social Science Research Unit, Institute of Education, University of London: London.
55. Rees, R., et al., *Young people and phsyical activity: a systematic review of research on barriers and facilitators*. 2001, EPPI-Centre, Social Science Research Unit, Institute of Education, University of London: London.
56. Shepherd, J., et al., *Young people and healthy eating: a systematic review of research on barriers and facilitators*. 2001, EPPI-Centre, Social Science Research Unit, Institute of Education, University of London: London.
57. Thomas, J., et al. *Children and healthy eating: a systematic review of barriers and facilitators*. 2003 [cited.
58. Peersman, G., S. Oliver, and A. Oakley, *EPPI-Centre Review Guidelines*. 1997, EPPI-Centre, Social Science Research Unit, Institute of Education, University of London: London.
59. Oliver, S. and G. Peersman, *Critical appraisal of research evidence: finding useful and reliable answers*, in *Using research for effective health promotion*, S. Oliver and G. Peersman, Editors. 2001, Open University press: Buckingham. p. 82-95.
60. Prochaska, J.O. and C.C. DiClemente, *Transtheoretical therapy toward a more integrative model of change.* Psychotherapy: Theory, Research and Practice, 1982. **19**(3): p. 276-287.
61. Rosenstock, I., *Historical origins of the health belief model.* Health Education Monographs, 1974. **2**(4).
62. Ajzen, I. and M. Fishbein, *Understanding attitudes and predicting social behaviour*. 1980, Englewood Cliffs, New Jersey: Prentice Hall.
63. Bandura, A., *Social learning theory*. 1977, Englewood Cliffs, New Jersey: Prentice-Hall.
64. Health Promotion Agency for Northern Ireland. *Health promotion theories and models*. 2004 [cited 2006 28 February].
65. Nutbeam, D. and E. Harris, *Theory in a nutshell. A practical guide to health promotion theories*. 2nd ed. 2004, Sydney: McGraw-Hill.
66. Green, L. and M. Kreuter, *Health promotion planning: an educational and ecological approach*. 3rd Edition ed. 1999, Mountain View, California: Mayfield.
67. Tones, K. and S. Tilford, *Health education: effectiveness, efficiency and equity*. 1994, London: Chapman & Hall.
68. Riemsma, R.P., et al., *Systematic review of the effectiveness of stage based interventions to promote smoking cessation.* Bmj, 2003. **326**(7400): p. 1175-7.

69. Patton, M., *The program's theory of action: conceptualizing casual linkages*, in *Utilization-focussed Evaluation*, M. Patton, Editor. 1986, Sage Publications: Thousand Oaks.
70. Wholey, J., *Evaluability assessment: Developing program theory*, in *Using Theory in Evaluation, New Directions in Evaluation Series*, L. Bickman, Editor. 1987, Jossey-Bass Publishers: San Francisco, CA Wholey, J. p. 77-92.
71. Funnell, S., *Program Logic: An adaptable tool for designing and evaluating programs.* Evaluation News and Comment, 1997. **July**: p. 5-17.
72. Lumley, J., et al., *Interventions for promoting smoking cessation during pregnancy.* Cochrane Database Syst Rev, 2004(4): p. CD001055.
73. Prochaska, J.O., C.C. DiClemente, and J.C. Norcross, *In search of how people change. Applications to addictive behaviors.* Am Psychol, 1992. **47**(9): p. 1102-14.
74. Arnstein, S., *Ladder of citizen participation.* American Institute of Planners Journal, 1969. **35**: p. 216-224.
75. Oliver, S., et al., *Involving consumers in research and development agenda setting for the NHS: developing an evidence-based approach.* Health Technol Assess, 2004. **8**(15): p. 1-148, III-IV.
76. Scanlon, J.W., et al., *Evaluability assessment: Avoiding Type III and IV errors*, in *Evaluation management: A source book of readings*, G.R. Gilbert and P.J. Conklin, Editors. 1977, U.S. Civil Service Commission: Charlottesville, VA.
77. Dobson, L. and T. Cook, *Avoiding type III error in program evaluation: results from a field experiment.* Evaluation and Program Planning, 1980. **3**: p. 269-276.
78. Dane, A.V. and B.H. Schneider, *Program integrity in primary and early secondary prevention: are implementation effects out of control?* Clin Psychol Rev, 1998. **18**(1): p. 23-45.
79. Dusenbury, L., et al., *A review of research on fidelity of implementation: implications for drug abuse prevention in school settings.* Health Educ Res, 2003. **18**(2): p. 237-56.
80. Dumas, J.E., et al., *Promoting intervention fidelity. Conceptual issues, methods, and preliminary results from the EARLY ALLIANCE prevention trial.* Am J Prev Med, 2001. **20**(1 Suppl): p. 38-47.
81. Greenhalgh, T., et al., *Transferability of principles of evidence based medicine to improve educational quality: systematic review and case study of an online course in primary health care.* Bmj, 2003. **326**(7381): p. 142-5.
82. McDonald, G., *Social work: beyond control?*, in *Non-random reflections on health services research*, A. Maynard and I. Chalmers, Editors. 1997, BMJ Publishing Group: London.
83. Macintyre, S., *Good intentions and received wisdom are not enough*, in *Evidence into practice: challenges and opportunities for UK public health*. 2001, The Royal College of Physicians: London.
84. Oliver, S., *Making research more useful: integrating different perspectives and different methods*, in *Using research for effective health promotion*, S. Oliver and G. Peersman, Editors. 2001, Open University Press: Buckingham.

85. Steckler, A. and L. Linnan, eds. *Process evaluation for public health interventions and research*. 2002, Jossey-Bass: USA.
86. Green, J. and K. Tones, *Towards a secure evidence base for health promotion.* J Public Health Med, 1999. **21**(2): p. 133-9.
87. Murta, S.G., K. Sanderson, and B. Oldenburg, *Process evaluation in occupational stress management programs: a systematic review.* Am J Health Promot, 2007. **21**(4): p. 248-54.
88. Lockyer, J., S.T. Gondocz, and R.L. Thivierge, *Knowledge translation: the role and place of practice reflection.* J Contin Educ Health Prof, 2004. **24**(1): p. 50-6.
89. Mantzoukas, S., *A review of evidence-based practice, nursing research and reflection: levelling the hierarchy.* J Clin Nurs, 2007.
90. Rycroft-Malone, J., et al., *What counts as evidence in evidence-based practice?* J Adv Nurs, 2004. **47**(1): p. 81-90.
91. Thomas, J., et al., *Integrating qualitative research with trials in systematic reviews.* Bmj, 2004. **328**(7446): p. 1010-2.
92. Harden, A., A. Oakley, and S. Oliver, *Peer-delivered health promotion for young people: a systematic review of different study designs.* Health Education Journal, 2001. **60**(4): p. 339-353.
93. Grimshaw, J., et al., *Complexity and systematic reviews: report to the US Congress Office of Technology Assessment*. 1995, Office of Technology Assessment: Washington, DC.
94. Rychetnik, L., et al., *Criteria for evaluating evidence on public health interventions.* J Epidemiol Community Health, 2002. **56**(2): p. 119-27.
95. Rychetnik, L., D. Nutbeam, and P. Hawe, *Lessons from a review of publications in three health promotion journals from 1989 to 1994.* Health Educ Res, 1997. **12**(4): p. 491-504.
96. Mays, N., C. Pope, and J. Popay, *Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field.* J Health Serv Res Policy, 2005. **10 Suppl 1**: p. 6-20.
97. Dahlgren, G. and M. Whitehead, *Policies and strategies to promote social equity in health*. 1991, Institute of Futures Studies: Stockholm.
98. Popay, J., et al., *Guidance on the conduct of narrative synthesis in systematic reviews. This guidance is available from j.popay@lancaster.ac.uk*. 2006, Lancaster University: UK.
99. Noblitt, G. and R. Hare, eds. *Meta-ethnography: synthesising qualitative studies*. 1988, Sage Publications: Newbury Park, CA.
100. Popay, J., ed. *Moving beyond effectivebess: methodological issues in the synthesis of diverse sources if evidence.* 2006, National Institute for Health and Clinical Excellence: London.
101. Pearson, A., *Balancing the evidence: incorporating the synthesis of qualitative data into systematic reviews.* JBI Reports, 2004. **2**(2): p. 45-64.
102. Pearson, A., et al., *The JBI model of evidence-based healthcare.* Journal of Evidence-Based Healthcare, 2005. **3**(8): p. 207-216.
103. Britten, N., et al., *Using meta ethnography to synthesise qualitative research: a worked example.* J Health Serv Res Policy, 2002. **7**(4): p. 209-15.

104. Hawe, P. and A. Shiell, *Preserving innovation under increasing accountability pressures: the health promotion investment portfolio approach.* Health Promotion Journal of Australia, 1995. **5**(2): p. 4-9.
105. Macintyre, S., *Evaluating the evidence on measures to reduce inequalities in health*, in *Health inequalities: evidence, policy and implementation. Proceedings from a meeting of the Health Equity Network.*, A. Oliver and M. Exworthy, Editors. 2003, The Nuffield Trust: London.
106. National Health and Medical Research Council, *Using socioeconomic evidence in clinical practice guidelines*. 2003, Commonwealth of Australia: Canberra.
107. Kawachi, I., S.V. Subramanian, and N. Almeida-Filho, *A glossary for health inequalities.* J Epidemiol Community Health, 2002. **56**(9): p. 647-52.
108. Macino, J.A. and B. Starfield, *Annotated bibliography on equity in health, 1980-2001.* International Journal for Equity in Health, 2002. **1**(1): p. 1-20.
109. Millward, L., M. Kelly, and D. Nutbeam, *Public health interventions research: the evidence*. 2001, Health Development Agency: London.
110. Maynard, A., *Evidence based medicine. Cost effectiveness and equity are ignored.* Bmj, 1996. **313**(7050): p. 170-1.
111. Ogilvie, D. and M. Petticrew, *Reducing social inequalities in smoking: can evidence inform policy? A pilot study.* Tob Control, 2004. **13**(2): p. 129-31.
112. Tugwell, P., et al., *Cochrane and Campbell Collaborations, and health equity.* Lancet, 2006. **367**(9517): p. 1128-30.
113. Oakley, A., G. Peersman, and S. oliver, *Social characteristics of participants in health promotion effectiveness research: trial and error?* Education for Health, 1998. **11**(3): p. 305-317.
114. Rogers, W.A., *Evidence based medicine and justice: a framework for looking at the impact of EBM upon vulnerable or disadvantaged groups.* J Med Ethics, 2004. **30**(2): p. 141-5.
115. Britton, A., et al., *Choosing between randomised and non-randomised studies: a systematic review.* Health Technol Assess, 1998. **2**(13): p. i-iv, 1-124.
116. Rogers, W.A., *Is there a tension between doctors' duty of care and evidence-based medicine?* Health Care Analysis, 2002. **10**: p. 277-287.
117. McGowan, J., et al. *Identifying studies that include issues of equity for inclusion in Cochrane reviews*. in *Cochrane Colloquium*. 2003. Barcelona, Spain.
118. Jackson, N.W., et al., *Interventions implemented through sporting organisations for increasing participation in sport.* Cochrane Database Syst Rev, 2005(2): p. CD004812.
119. Doull, M., et al., *Peer support strategies for improving the health and wellbeing of individuals with chronic diseases (protocol).* The Cochrane Database of Systematic Reviews, 2005. **Issue 3**(Art. No.: CD 005352.DOI:10.1002/14651858.CD005352).
120. Doull, M., et al., *Peer-based interventions for reducing morbidity and mortality in HIV-infected women (protocol).* The Cochrane Database of

Systematic Reviews, 2004. **Issue 2**(Art. No.: CD004774. DOI: 10.1001/14651858.CD004774).
121. Evans, T. and H. Brown, *Road traffic crashes: operationalizing equity in the context of health sector reform.* Inj Control Saf Promot, 2003. **10**(1-2): p. 11-2.
122. Kristjansson, E.A., et al., *School feeding for improving the physical and psychosocial health of disadvantaged elementary school children.* Cochrane Database Syst Rev, 2007(1): p. CD004676.
123. Shediac-Rizkallah, M.C. and L.R. Bone, *Planning for the sustainability of community-based health programs: conceptual frameworks and future directions for research, practice and policy.* Health Educ Res, 1998. **13**(1): p. 87-108.
124. Swerissen, H. and B.R. Crisp, *The sustainability of health promotion interventions for different levels of social organization.* Health Promot Int, 2004. **19**(1): p. 123-30.
125. Bossert, T.J., *Can they get along without us? Sustainability of donor-supported health projects in Central America and Africa.* Soc Sci Med, 1990. **30**(9): p. 1015-23.
126. The Health Communication Unit. *Overview of sustainability*. 2001 [cited 2005.
127. Scheirer, M.A., *Is sustainability possible? A review and commentary on empirical studies of program sustainability.* American Journal of Evaluation, 2005. **26**(3): p. 320-347.
128. Bailie, R.S., et al., *Investigating the sustainability of outcomes in a chronic disease treatment programme.* Soc Sci Med, 2006. **63**(6): p. 1661-70.
129. Frommer, M. and L. Rychetnik, *From evidence-based medicine to evidence-based public health*, in *Evidence-based health policy: problems and possibilities*, V. Lin and B. Gibson, Editors. 2003, Oxford University Press: Melbourne.
130. Hawe, P., et al., *Methods for exploring implementation variation and local context within a cluster randomised community intervention trial.* J Epidemiol Community Health, 2004. **58**(9): p. 788-93.
131. Eccles, M., et al., *Research designs for studies evaluating the effectiveness of change and improvement strategies.* Qual Saf Health Care, 2003. **12**(1): p. 47-52.
132. Israel, B.A., et al., *Evaluation of health education programs: current assessment and future directions.* Health Educ Q, 1995. **22**(3): p. 364-89.
133. Bracht, N., et al., *Community ownership and program continuation following a health demonstration project.* Health Educ Res, 1994. **9**(2): p. 243-55.
134. Goodman, R. and A. Steckler, *A model of institutionalisation of health promotion programs.* Family and Community Health, 1987. **11**: p. 63-78.
135. Green, L., *Is institutionalisation the proper goal of grant making?* Family and Community Health, 1987. **11**: p. 79.
136. Evashwick, C. and M. Ory, *Organizational characteristics of successful innovative health care programs sustained over time.* Fam Community Health, 2003. **26**(3): p. 177-93.

137. Bammer, G. *Integration and implementation sciences: will developing a new specialisation improve our effectiveness in tacking complex issues?* 2003 [cited 2007 September]; Available from: http://www.ksg.harvard.edu/sed/docs/sdsem/bammer_paper030919.pdf.
138. Bauman, L.J., R.E. Stein, and H.T. Ireys, *Reinventing fidelity: the transfer of social technology among settings.* Am J Community Psychol, 1991. **19**(4): p. 619-39.
139. Ottoson, J. and L. Green, *Reconciling concept and context: theory of implementation.* Advances in health Education and Promotion, 1987. **2**: p. 353-382.
140. Scheirer, M.A., *Designing and using process evaluations*, in *Handbook of Practical Program Evaluation*, J.S. Wholey, H.P. Hartry, and K.E. Newcomer, Editors. 1994, Jossey Bass: San Francisco.
141. Resnicow, K.A., D. Cross, and E.L. Wynder, *The know your body program: a review of evaluation studies.* Bulletin of the New York Academy of Medicine, 1993. **70**(3): p. 188-207.
142. Green, L.W. and R.E. Glasgow, *Evaluating the relevance, generalization, and applicability of research: issues in external validation and translation methodology.* Eval Health Prof, 2006. **29**(1): p. 126-53.
143. Glasgow, R.E., et al., *External validity: we need to do more.* Ann Behav Med, 2006. **31**(2): p. 105-8.
144. Wang, S., J.R. Moss, and J.E. Hiller, *Applicability and transferability of interventions in evidence-based public health.* Health Promot Int, 2006. **21**(1): p. 76-83.