

A Schema for Evaluating Evidence on Public Health Interventions

Version 4



A Schema for Evaluating Evidence on Public Health Interventions

Version 4

National Public Health Partnership

April 2002

Lucie Rychetnik and Michael Frommer
University of Sydney

The authors

Dr Lucie Rychetnik and Associate Professor Michael Frommer are based at the School of Public Health, University of Sydney, NSW 2006.

Email: lucier@med.usyd.edu.au

Email: mfrommer@med.usyd.edu.au

Further copies

Printed copies of the Schema are available from the National Public Health Partnership Secretariat, Level 12, 589 Collins St, Melbourne, Victoria 3001, Australia.

Telephone: (61 3) 9616 1515

Facsimile: (61 3) 9616 1500

Email: nphp@dhs.vic.gov.au

Website: www.nphp.gov.au

Citation

Rychetnik L, Frommer M. *A Schema for Evaluating Evidence on Public Health Interventions; Version 4*. National Public Health Partnership, Melbourne, 2002.

ISBN: 0-9580326-7-X

Copyright: National Public Health Partnership, April 2002

This work is copyright. This work may be reproduced in whole or in part for research or training purposes subject to the inclusion of an acknowledgment of the source and provided no commercial usage or sale is to be made. Reproduction for purposes other than those indicated above requires prior written permission of the National Public Health Partnership, GPO Box 4057, Melbourne 3001, Victoria, Australia.

Acknowledgements

We would like to thank all those who have contributed to the development of the Schema. In particular, we would like to thank the members of the project steering group for their valuable guidance and constructive feedback.

Steering group for stage one (first draft of Schema)

Prof George Rubin (chair), Prof Sandy Gifford, Dr William Hart for Prof John Catford, Prof David Henry, Dr Shirley Hendy, Prof Vivian Lin, Ms Bebe Loff for Ms Meredith Carter, Prof John McNeil, Dr Cathy Mead, Ms Helen Moore, Prof Don Nutbeam, Prof Sally Redman, Dr Leonie Segal.

Steering group for stage two (testing and refining Schema)

Prof Vivian Lin (chair), Prof Robert Bush, Dr Neil Formica, Mr Brendan Gibson, Prof Sandy Gifford, Dr William Hart for Prof John Catford, Dr Shirley Hendy, Prof David Henry, Ms Bebe Loff for Ms Meredith Carter, Prof John McNeil, Dr Cathy Mead, Ms Helen Moore, Prof Don Nutbeam, Prof Aileen Plant, Prof Sally Redman, Dr Glenn Salkeld.

We also thank the researchers and practitioners who participated in consultations on the Schema and those who conducted case studies to trial drafts of the Schema. They all provided suggestions on how the Schema could be improved (see 'How the schema was developed').

Case studies testing Version Two of Schema

The first review team was led by Associate Professor James Harrison, Director of the National Injury Surveillance Unit, Flinders University, Adelaide. He was assisted by research colleagues Ms Emma Miller and Ms Su Gruszin.

The second review team was led by Professor Adrian Bauman from the University of NSW. He was assisted by research colleagues Mr Ben Smith from University of Sydney and Ms Dafna Merom from the Epidemiology Unit, South Western Area Health Service.

Case study testing Version 3.1 of Schema

The review team was based at Northern Rivers Public Health Unit and comprised Mr Mark Bartlett (Director of Public Health), Ms Janet Terry (Public Health Officer), and Ms Marianne Trent (Public Health Nurse).

Finally, special thanks to members of the National Public Health Partnership Secretariat who provided invaluable encouragement and project support, particularly Dr Cathy Mead, Mr Gianfranco Spinoso, Ms Karen Roger, Ms Natalie Dunham, Mr Darryl Kosch and Ms Necia Burford.

Editing and design by Aria Design, Geelong.

Contents

Introduction	1
User's Notes for Version 4	7
Section 1: The scope of your review	17
Section 2: The papers in the review	19
Section 3: Describing the results from the papers selected	29
Section 4: Interpreting each article	31
Section 5: Summarising the body of evidence	33
Appendix 1: Hierarchies of study design and designation of levels of evidence	37
Appendix 2: Supplementary Guides 1 to 5	39
Supplementary Guide 1: Appraising reviews	40
Supplementary Guide 2: Appraising randomised controlled trials	42
Supplementary Guide 3: Appraising observational studies	43
Supplementary Guide 4: Appraising economic evaluations	45
Supplementary Guide 5: Appraising qualitative studies	46
Appendix 3: A table for summarising all the papers reviewed	49
References	51

Introduction

THE SCHEMA

This Schema is intended to be used as guide in the appraisal of evidence on public health interventions. It is designed to be applied to *evidence* in the form of a collection of research papers or evaluation reports that examine and describe the effects (benefits and harms) of an intervention.

The Schema covers two stages of evidence appraisal. The first is the appraisal of individual papers or reports to determine whether they provide credible and useful information about an intervention(s). The second is the formulation of conclusions about the value of the available evidence, enabling the preparation of a summary statement on what is known, and what is not known, about a type of public health intervention(s).

BACKGROUND

The impetus for this project arose from discussion in the Australian National Health and Medical Research Council's Health Advisory Committee (HAC) about difficulties that can arise when established methods from evidence-based medicine are used to evaluate research on public health interventions. The HAC and the National Public Health Partnership (NPHP) sponsored two expert workshops (in April and May 1999) to consider these difficulties. Participants concluded that established methods of evidence appraisal provided a useful starting point for evaluating evidence on all types of health interventions. They recommended that these methods should be expanded or adapted to take account of the diversity of public health interventions, the different evaluations that are conducted in public health settings, and the importance of contextual factors in public health research and practice.

The NPHP Group subsequently commissioned a project to develop a Schema for Evaluating Evidence on Public Health Interventions (Stage One); and then to trial and revise the Schema in applied settings (Stage Two). The product of Stage One was a Discussion Paper (including the first draft of the Schema) published by the NPHP in May 2000.^{1*} The product of the Stage Two is Version 4 of the Schema. Further information on how the Schema was developed and refined is provided below.

RATIONALE FOR EVALUATING EVIDENCE

Prediction of the safety and effectiveness of specific public health interventions depends on a knowledge of what interventions have worked, or failed, in the past. Evidence-based practice relies on the findings of sound evaluation research to determine whether an intervention is likely to be effective. The principles of evidence-based practice, which are described in Table 1, underpinned the development of the Schema.

* Superscript numerals refer to the References on page 51.

Table 1: Principles of evidence based practice

- 1) It is important to know whether public health interventions are effective and do more good than harm.
- 2) The benefits and costs of public health interventions should be described and evaluated, so they can be weighed against other options for the use of resources.
- 3) People who make (or are affected by) evidence-based decisions about public health interventions should be aware of the strengths, limitations and gaps in the available evidence.

These principles imply a need to define criteria that can be used to judge the strengths, limitations and gaps in the available evidence. The criteria should enable us to determine whether a collection of research papers and/or evaluation reports provides evidence that is *good enough* to support a decision on the benefits and harms of a public health intervention.

CRITERIA FOR EVALUATING EVIDENCE

The criteria used to assess the available evidence can influence the conclusions that we may reach about possible interventions, and thus affect public health policy and practice. We acknowledge that the evaluation of evidence can often be contested, especially if an evidence review is invoked to make controversial recommendations about the allocation of funding to public health programs.

As far as possible, evidence evaluation criteria should reflect accepted standards of evaluation research, and should be aligned with the needs of public health practitioners, policy makers and recipients of public health programs. Our aim has been to reflect these standards and needs in the structure and content of the Schema.

The biggest challenge in developing the Schema was to balance a requirement for comprehensive appraisal of evidence with a requirement for brevity and ease of application. For example, many potential users of the Schema emphasised the importance of taking into account contextual social and political factors when judging the credibility and utility of papers or reports that were to be used as evidence about the value of an intervention. We received many suggestions for questions that should be added as successive drafts of the Schema were tested and revised. Conversely, others wanted to minimise the number of questions asked and preferred to focus on simple grading systems, such as the hierarchical ‘levels of evidence’ determined by study design (Appendix 1).

Development of the Schema was informed by a review of existing guides (checklists or criteria) for evaluating papers or reports on health care interventions.^{1,2} Existing guides addressed core aspects of critical appraisal, such as the scientific rigour of an evaluation, and several have been incorporated into this Schema. However, most of the existing guides did not ask reviewers to take into account the setting in which interventions had been implemented, or the social and political environment in which they were evaluated.

The Schema presents a comprehensive range of questions about evidence. For example, it raises questions about the *intervention* itself and the setting in which it was implemented; questions about the *evaluation* of the intervention (i.e. the context and methods of the evaluation); and questions about the available *reports*. The Schema also encourages users to give explicit consideration to, and interpret, the evidence within the context of their own clearly defined *review* setting. These dimensions of critical appraisal, and their relationships, are illustrated in Figure 1 below.

Figure 1: Examining the evidence about an intervention.



HOW THE SCHEMA WAS DEVELOPED

First draft of the Schema

The steps involved in developing the first draft of the Schema in 1999–2000 included:

- a literature review on evidence-based health care, contemporary methods of critical appraisal, and on standards for public health evaluation; and
- consultation with senior public health researchers and practitioners around Australia.

The findings of the literature review and consultations

The findings of the literature review and the consultations with public health professionals are detailed in the discussion paper accompanying the first draft of the Schema.¹ Key points distilled from those findings are as follows.

- Evaluating evidence on public health interventions should build on existing approaches for evaluating health research and support the principles of evidence-based practice. It must however reflect the complexity of public health practice and the diversity of evaluations that are conducted in public health settings. A Schema should help reviewers to examine heterogeneous collections of evaluation papers and reports.
- Many community interventions have not been subjected to evaluative research and their outcomes and harms are unknown or have not been quantified. A Schema for evaluating evidence should help to assess what is known about the benefits and harms of interventions and the degree of uncertainty about those effects. It should also guide reviewers to make explicit the gaps in the evidence.
- Public health interventions often involve manipulating environments or contextual variables, or interact with contextual variables. Also, many practitioners and policy makers take a social and historical view of public health and require a discussion of context, politics and economics as well as biomedical and technical factors. A Schema for evaluating evidence should take account of the social aspects of public health.

Expanding the criteria used in evidence-based medicine

In developing the Schema we sought to cover critical appraisal questions that would reflect 'best practice' in the evaluation of public health interventions. For example, the Schema includes questions that encourage reviewers to consider whether the evidence demonstrates that an intervention was adequately implemented in the evaluation setting(s), whether information is provided about the implementation context, and whether interactions that occur between public health interventions and their context were assessed and reported.

These contextual and applied issues were included to supplement the critical appraisal approach already established in evidence-based medicine, which emphasise the nature and quality of an evaluation study design and the study methods.

Revising drafts and preparing Version 4 of the Schema

Version 4 of the Schema is the result of an iterative development process in which earlier drafts were pilot-tested in a series of case studies and revised. The steps involved were as follows:

- Pilot test the first draft of the Schema in a case study conducted by the authors of the Schema.
- Identify technical difficulties and revise the Schema to prepare Version 2.
- Test Version 2 the Schema in a further two case studies (2 and 3) conducted by two separate and independent groups of researchers/ practitioners.
- Conduct a consultation on Version 2 of the Schema with community representatives.
- Review feedback from case studies 2 and 3 and community consultation and prepare Version 3 of the Schema.
- Test Version 3 of the Schema in a fourth case study conducted by a different independent group of practitioners based in a regional Public Health Unit.
- Conduct a consultation on Version 3 of the Schema with researchers and practitioners nominated by the NPHP Aboriginal and Torres Strait Islander Working Group (see below).
- Review feedback from case study 4 and consultations, and revise Version 3 to prepare Version 4 of the Schema for subsequent publication.

As with previous drafts, Version 4 of the Schema has been envisaged as 'work in progress'. Ideas about the critical appraisal of evidence and the role of evidence in public health policy and practice will continue to evolve. It is thus anticipated that, over time, modifications to the Schema may be required.

Indigenous health

Version 4 of the Schema is based on consultations with potential users of the Schema. As indicated above, these consultations included a one-day workshop with researchers, practitioners and policy makers selected by NPHP Aboriginal and Torres Strait Islander Working Group. The consultation participants were of the view that the Schema would assist users who have experience in Indigenous health to appraise research evidence in a way that is relevant and useful to Aboriginal and Torres Strait Islander communities. However, to reflect the Australian public health context explicitly, they recommended additional prompts that will encourage other users of the Schema to consider evaluation research against the needs of local Indigenous communities. Such prompts have been included in Version 4 of the Schema.

User's Notes for Version 4

WHAT THE SCHEMA DOES

This Schema is a guide to the critical appraisal stages of a literature review, and to the synthesis of findings from the critical appraisal.

The overall process of reviewing published and other literature on public health interventions, and using the findings from the review, involves at least six steps:

- Identify the purpose of the literature review and formulate the review question(s) to be addressed
- Find and collate studies to be reviewed (i.e. articles and evaluation reports)
- Appraise each article or evaluation report
- Formulate a statement on the body of evidence
- Publish review findings
- Apply review findings to inform decisions about public health policy or practice.

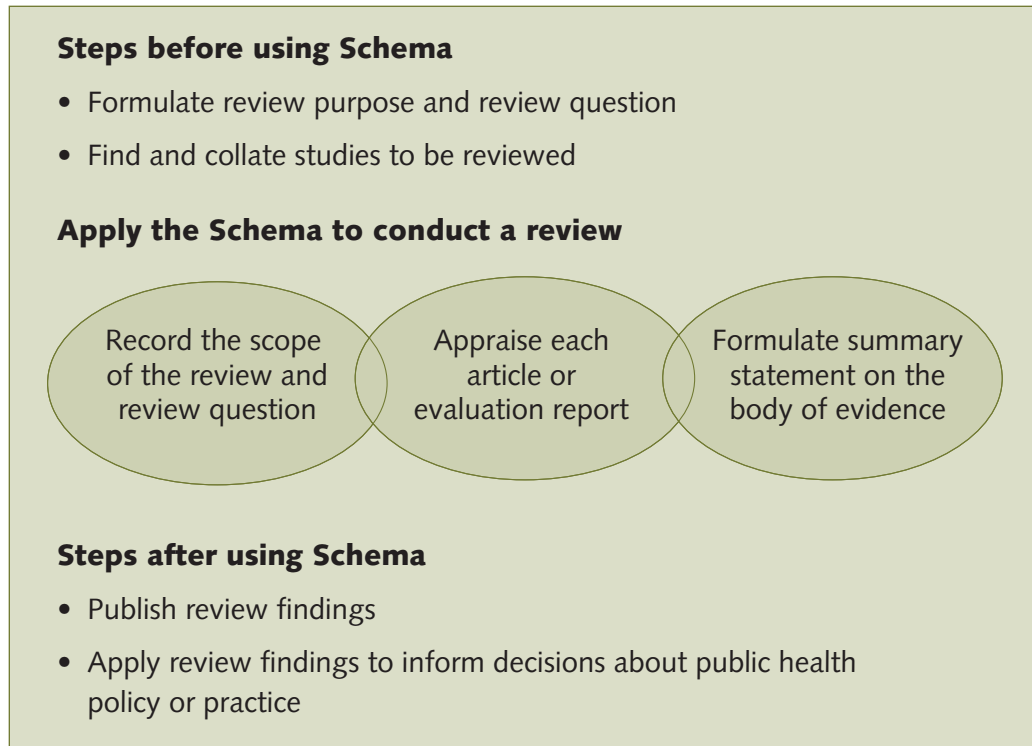
This Schema deals with the third and fourth steps. The Schema is used to **appraise individual papers** and to **formulate a summary statement** about those articles and reports. As a review of the literature may be conducted for various reasons, users of the Schema are also asked to **record** the decisions that were made about the scope of the review. Thus the critical appraisal of articles or reports, and interpretation of the evidence that they present, are conducted within the context of an explicitly defined review setting.

WHAT THE SCHEMA DOES NOT DO

The Schema is **not** a guide to the processes involved in formulating an appropriate review question or a guide on how to search the literature. It is essential, however, that these steps are done thoroughly. You may find it useful to refer to publications from the Australian NHMRC, the UK Centre for Reviews and Dissemination, and the Cochrane Collaboration on how to prepare for systematic reviews.^{3,4,5}

We emphasise that the Schema is **not** a guide on how use the review once it is completed. The intention is that evidence reviews prepared using the Schema will be of high quality and make useful contributions to decisions about public health interventions. However, the Schema was developed to enhance the technique of critical appraisal, and it does not address the range of social and political factors that determine public health policy or feed into decisions about resource allocation. If you are involved in making decisions about public health policy and practice, you may wish to refer to the National Public Health Partnership documents on planning in public health practice or planning an intervention portfolio.^{6,7} Figure 2 shows what the Schema does and does not do.

Figure 2: What the Schema does and does not do



FORMAT OF THE SCHEMA

The Schema comprises a series of critical appraisal questions to be applied to literature that reports on the evaluation of public health interventions. The structure of the Schema is designed to arrange and group the questions in a logical sequence. We hope that this structure will help users to appraise the literature systematically and comprehensively, thereby enhancing the rigour and utility of their reviews. The aim is to assist reviewers to reach sound conclusions on what is known, and what is not known, about public health interventions.

The questions in *Version 4* of the Schema are arranged in five sections. In Section 1 users of the Schema identify the purpose of their review and their review needs. In Sections 2, 3 and 4, users appraise and summarise individual papers or evaluation reports. Section 5 leads users to prepare a consolidated statement about all the papers and reports in the body of available evidence. Conclusions (Section 5) should be linked to the aims of the review (Section 1).

THE SCHEMA IN SUMMARY

Section 1: Recording the purpose and scope of your review

Section 1 is done only once – at the start of your review. On completing this section, you will have a clear record on why you are doing the review and what the review is about. Refer to this statement on the purpose and scope of your review as you appraise individual papers (Sections 2 to 4), and when you are forming your conclusions about the evidence (Section 5).

A clear statement about the purpose and scope of your review will also assist subsequent users of your review to interpret your conclusions. The key points to be addressed in Section 1 are to identify the question you want to answer in the review and describe how your findings are to be used.

Please note that the Schema has been prepared to provide a comprehensive checklist of critical appraisal questions, some of which may not be relevant to your particular review. Thus when you have clarified the purpose and scope of your review, you will consider the questions posed in Sections 2 to 4 and determine the value of applying the Schema in full. Those questions that appear less relevant to your needs should be considered optional.

Section 2: Evaluating each article in the review

Section 2 is done separately for each paper or report. In Section 2 you will examine each paper closely to understand and assess:

- the intervention being evaluated;
- the context in which the intervention took place;
- the background to the evaluation of the intervention; and
- how the evaluation was carried out.

By completing this section for each paper, you will be able to decide whether to include the paper's findings in your review (i.e. whether to go on to apply Sections 3 to 5 to that paper) or whether to reject the paper and exclude it from further detailed consideration. Section 2 has five sub-sections – 2A to 2E.

2A PUBLICATION DETAILS

Sub-section 2A asks you to record the publication details of each paper that you examine.

2B SPECIFYING THE INTERVENTION

Sub-section 2B guides you to describe precisely the intervention(s) that have been evaluated in each paper, as well as the origin and development of the intervention(s). Some research reports only include a summary description of the intervention. If the research paper does not provide adequate information about the intervention, you may need to consult any additional documentation that may be available.

On completing sub-section 2B, you will have either a complete and detailed description of the intervention, or a good idea of what information is lacking on the intervention. This will help you assess the relevance of the research paper to your review.

2C IDENTIFYING THE INTERVENTION CONTEXT

Sub-section 2C focuses on the context (or setting) in which the intervention took place. It is important to determine whether contextual factors are critical or integral to the intervention. It is also important to examine whether adequate details were provided about the context to inform subsequent decisions about the applicability of the intervention to other settings.

The findings from 2C will help you to assess the transferability of the intervention(s), and to determine the relevance of the intervention to your review when you apply Section 4 to the article. If the article does not provide enough information on the intervention context, sub-section 2C will help you to identify that gap in the evidence. Lack of information about the intervention context may influence your decision on how to use the article's findings.

At the end of Section 2C you will consider if the intervention and the context described in the article are relevant to the scope of your review (as identified in Section 1). If the article is relevant (or partly relevant) you will continue to apply the Schema. If you decide that the article is not relevant (and thus not useful) you will move on to the next article (paper or report) to be appraised.

2D THE EVALUATION CONTEXT – BACKGROUND, PURPOSE AND QUESTIONS ASKED

Sub-section 2D guides you towards an understanding of the background of the evaluation that is reported in the article, and to identify the evaluation questions. In some articles the intervention and its evaluation are part of the same project. In other articles you may find that an evaluation was planned and conducted as a separate project rather than as part of the implementation of the intervention. Key issues are how the timing of the evaluation relates to the development and implementation of the intervention, and who conducted the evaluation and why.

Sub-section 2D helps you to assess the context of the evaluation and to consider whether that evaluation was appropriate and adequate for its purpose. This will assist in interpreting the reported evaluation results in Section 4. If the article provides insufficient information on the evaluation context, additional information could be sought, including contact with the author.

2E THE METHODS USED TO EVALUATE THE INTERVENTION

Sub-section 2E requires you to appraise the quality and rigour of the method(s) used to evaluate the intervention. You will be asked to apply one or more critical appraisal checklist(s) (included in the Appendices), depending on the research methods used in the study. The key question addressed is whether the design and conduct of the evaluation methods were adequate to produce credible evaluation results.

This assessment of the evaluation methods will help you decide whether to include the evaluation results from any particular article in your review findings (i.e. whether to continue with your appraisal of the article by applying Sections C and D).

At the end of Sub-section 2E you will decide if the findings of the evaluation are likely to be credible (based on the rigour of the evaluation methods used). If you conclude that the findings may be credible you will continue applying the Schema. If you conclude that the findings are not likely to be credible, you will identify any lessons to be learned from the article, and then move on to the next paper or report to be appraised.

Section 3: Describing the results from articles selected

Like Section 2, Section 3 is done separately for each paper or report. In Section 3 you will describe the results that were reported in the article. This includes both beneficial and adverse effects of the intervention, and the sustainability of those effects. Section 3 also deals with the relative effect of the intervention on disadvantaged groups, and examines the influence of the study context on the evaluation findings

You will use your findings from Section 3 in Section 4 to examine the applicability of the reported results to the scope of your review.

Section 4: Interpreting each paper

Like Sections 2 and 3, Section 4 is done separately for each paper or report. In Section 4 you will identify the relevance of the study and its findings to the scope of your review and your review questions. You will also draw out other potentially useful lessons from each article, such as requirements or barriers for successful implementation of the intervention strategies.

If the purpose of your review is to formulate recommendations for a particular policy or practice setting, in Section 4 you will also assess the transferability of the intervention, and decide whether the findings could be replicated in your own setting.

Section 5: Summarising the body of evidence

Section 5 is done only once – to complete your review. After you have reviewed each individual paper or report (using Sections 2, 3 and 4), Section 5 guides you to prepare a summary of all the papers or reports collectively. This summary is then used to formulate your conclusions about the body of available evidence. Section 5 is completed with reference to the purpose and scope of your review, and the review questions identified in Section 1. Section 5 has two Sub-sections – 5A and 5B.

5A GROUPING, RATING AND WEIGHING-UP YOUR PAPERS

In Sub-section 5A you will group the articles according to their research questions and intervention strategies. You will then:

- grade the overall quality of the studies within each group;
- assess the consistency of the findings among the stronger studies; and
- determine the applicability to your review of the strongest studies with the most consistent findings.

5B FORMULATING A SUMMARY STATEMENT

In Sub-section 5B you will form conclusions about the strengths and weaknesses of the evidence for your purposes (as specified in Section 1), and identify gaps in the evidence.

OPTIONS FOR USING THE SCHEMA

The Schema may be used by any group or individual conducting a literature review. However, the Schema requires user(s) to describe and judge different aspects of research reports; inevitably such judgements are strengthened by multi-disciplinary perspectives. We therefore recommend that, wherever possible, an appropriate multi-disciplinary group be convened to apply the Schema. This accords with the NHMRC recommendation for the use of multi-disciplinary groups to develop and evaluate practice guidelines.⁸

The multi-disciplinary group should include individuals with expertise in the types of evaluation research to be considered in the review and individuals with expertise in the relevant area of public health practice.

Preferably, the review group should also reflect the interests of those who will decide on the future implementation of the intervention(s) under consideration, and those potentially affected by such intervention(s). It is important to note that a *single* community perspective rarely exists. For example, Aboriginal and Torres Strait Islander communities in different regions may have different perspectives on whether a particular public health intervention is appropriate or feasible in their local setting. Perspectives may also vary depending on whether nominated representatives come from rural or urban groups, and whether they are community-based or professional groups.

The application of Section 1 (in which the scope of the review is defined) and Section 5 (in which a summary statement on the body of evidence is formulated) were particularly envisaged as group processes. The reading, review and summary of individual articles in Sections 2, 3 and 4 may be completed in a number of ways, and can involve one or more individuals.

Some options for the use of Sections 2–5 of the Schema are as follows:

Option One

- One person applies Sections 2, 3 and 4 of the Schema to review all the individual papers, and prepares a critical appraisal summary of each paper.
- Other members of the group read all the summaries of all the papers for review.
- The review group is convened to apply Section 5 of the Schema and prepares a collective group statement on the body of available evidence.

In this option, Sections 2, 3 and 4 of the Schema are used to *extract and summarise information* to be considered by the wider review group. The primary reviewer should therefore aim to provide a clear description of each article and to explain the basis of his/her judgements in Sections 2, 3 and 4.

Option Two

- All members of a review group read all of the original articles.
- Each person applies Sections 2, 3 and 4 to each article and produces his/her own summary statement on the body of evidence using Section 5.
- The group is convened and each member presents his or her conclusions on the evidence from Section 5.

- Members discuss areas of agreement and disagreement to formulate a consolidated group statement on the overall body of available evidence.

In this option, Section 5 is used by individuals to make judgements about the body of evidence. These individual judgements are then debated among group members to arrive at a group judgment on the evidence. Thus each reviewer uses Sections 2, 3 and 4 as a basis for his or her *reasoning and supporting facts* that lead to his or her conclusions in Section 5.

Option Three

Option Three is a form of compromise between the Options One and Two. Small groups or pairs of reviewers work through Sections 2, 3, 4 and 5 of the Schema. The body of literature is divided among the pairs or groups to share the workload. At the end, the entire review group convenes to compare notes and formulate a joint statement on the evidence.

Advantages and disadvantages of these options

The advantage of the Option One is its relative efficiency, with one person reading all the collated articles and summarising the relevant information for others. An important requirement of Option One is that the person allocated the task of reading the articles is highly skilled at critical appraisal of research publications. He or she must also be conscientious in extracting all the details required by others (who do not read the original articles) so that they can make an informed assessment of the evidence.

Option Two may be preferred when there are not many articles to review. It is possible that reviewers from different disciplinary backgrounds will answer some of the questions in Sections 2, 3 and 4 differently, particularly about studies that employed diverse methods. It may be important that everyone in the review group gains a direct or “first hand” opinion of the available articles before forming their conclusions in Section 5.

Option Three may be preferred if there is a large body of literature, but reviewers wish to check the inter-rater reliability of their reading and review process for individual articles, before they rely on these summaries to complete Section 5 of the Schema.

SOME ADDITIONAL POINTERS FOR USING THE SCHEMA

Expect some overlap between Sub-sections 2B–2D

The Schema deliberately makes a distinction between an intervention (e.g. a set of policy and practice strategies), the setting and context of that intervention (e.g. the social, organisational, political environment), the evaluation context (e.g. the background to and purpose of the evaluation, and the questions to be answered by the evaluation) and the evaluation methods (e.g. the study design and the rigour with which the study is conducted). We believe that this distinction will help you in categorising and appraising different aspects of a paper or report. We acknowledge, however, that some questions in Sections 2B–2D may appear to overlap, particularly

if the distinction between an intervention, its context, and the evaluation that was conducted are not clear in the article. Rarely does only one 'right' or 'wrong' way exist to allocate information from the paper to different Sections of the Schema. It is also anticipated that some users of the Schema will decide that not all of the questions are essential in their particular setting, and that they will modify their review accordingly.

Difficulties in reviewing literature

Evaluating evidence on interventions can be arduous and demanding. If you feel confused or frustrated, the following pointers may help you:

- Refer back to your review question in Section 1 to help you to define which particular aspects of a program you wish to focus on and learn about, and thus what you will include within your definition of an 'intervention'.
- Make the reasons for your decisions and judgements about the evidence explicit and clear
- Where a question seems to require information that you have already recorded, make cross-references between sub-sections of the Schema to avoid duplication of effort.

Reviewing systematic reviews

If you are reviewing a systematic review, you may find that some questions in Sub-sections 2B–2D may be difficult to apply because they were designed for individual evaluations, rather than for reviews of evaluations. Under these circumstances, you may wish to use the questions in Sub-sections 2B–2D as a checklist to consider whether the systematic review has adequately considered all the important aspects of the papers that it covers, and to identify the type of information that was overlooked or omitted in the systematic review.

In other respects, systematic reviews may be treated in much the same way original articles. You can identify the publication details in Sub-section 2A. A checklist for appraising the methods of a systematic review is included in Sub-section 2E. The findings of a systematic review can be summarised in Section 3 of the Schema, and the applicability of those findings can be discussed in Section 4. Your conclusions regarding a systematic review can be incorporated into your overall summary of the evidence in Section 5. You may also choose to use existing reviews as a comparison to your own review if they address similar literature.

Other uses of the Schema

In addition to conducting reviews of evidence, several alternative uses for the Schema have been identified by those who participated in, or commented on, its development. It could be used:

- as a reference for assessing proposed evaluations of an intervention,
- as a checklist for preparing evaluation papers and reports,
- to compare and assess past reviews of evidence, and
- to identify new review questions, or further research questions, based on gaps in the available evidence.

Although the Schema was not developed for these purposes, and has not been tested for them, criteria for the critical appraisal of evidence can potentially be used to influence future research and evaluation. We encourage others to conduct further research and development in these areas of potential application of the Schema. The aim of such activities should be to continue to improve the quality and utility of evidence (and summaries of evidence) on public health interventions, and to promote links between public health research and public health policy and practice.

A few words of caution

In the context of evidence-based policy and practice, the criteria that are used to judge the credibility and value of evaluation research, and the application of those criteria, should themselves be carefully examined. The systematic and explicit nature of the Schema has been widely commended and supported. However, a few concerns about the potential misuse of the Schema were raised in our consultations and discussions with the public health workforce. Their words of caution are acknowledged and noted here.

- The perspectives and values of reviewers influence the way in which many of the questions in the Schema are answered. Reviewers should be selected according to their experience and expertise, but their personal priorities and perspectives must be recognised and acknowledged.
- The Schema has value in promoting the rigorous assessment of research evidence. However, if application of the Schema reveals a lack of evidence on an apparently important intervention, or shows up the poor quality of the available evidence, this should not preclude the implementation of the intervention. For example, many public health initiatives intended to improve Aboriginal and Torres Strait Islander health have not been adequately evaluated. Action to implement these initiatives should not be delayed just because of a lack of good-quality evidence for their effectiveness – but an adequate evaluation of them should be strongly encouraged.
- By intention, the Schema draws attention to evaluation research and evidence-based decision making. We emphasise that the Schema is a guide to appraising the strengths and limitations of evaluation research, rather than a guide to making policy and practice decisions. It should be noted, however, that the emphasis on evidence can frame and drive debates about intervention options. Care should be taken to ensure that the available evidence is used to provide information on the range and scope of initiatives considered, rather than limiting them.

Section 1: The scope of your review

Before you proceed to appraising individual papers and reports, you should be able to state why you are doing the review and what the review is about, and to describe the setting in which you are doing the review.

Thus the review questions will need to be agreed and finalised before you can complete Section 1 of the Schema.

1A THE FOLLOWING IS A CHECKLIST OF ITEMS THAT YOU SHOULD RECORD ABOUT THE SCOPE OF YOUR REVIEW.

- What is the question that you want to answer in the review?
- How are you (and possibly others) going to use the findings of the review?
- Who asked for the review to be done?
- How has the review been funded?
- Who is actually carrying out the review?

CASE EXAMPLE*

Does available evidence support the introduction of meningococcal vaccine on a population-wide basis in Australia? This question is to be examined with regard to professional and community perceptions of the burden of meningococcal disease, vaccine efficacy, acceptability of the vaccine by the community, and cost of including the vaccine in existing schedules. The review findings will be summarised in a paper to be considered by an expert committee that will make recommendations on whether or not the vaccine should be incorporated in the Australian immunisation schedule. The review was funded by the Commonwealth Government, and is being undertaken by the department of infectious diseases in a major Australian university (named). The review team comprises an infectious diseases physician, an epidemiologist, a statistician, a health economist, an academic general practitioner, and two representatives from a health consumers' association. (All members of the team are named, and their titles and positions or affiliations are identified).

* All case examples in the Schema are hypothetical.

CASE EXAMPLE

What are the most effective ways of increasing the consumption of fruit and vegetables among people living in remote parts of Australia? The review to answer this question is being undertaken at the request of, and with funding from, the Regional Health Service covering a remote area. The Health Service is seeking to implement Statewide nutrition policy, and is in the process of planning local interventions. The review findings, and recommendations based on them, will be presented to the Health Service Executive. The review is being coordinated by the Population Health Unit of the Health Service, and the review team comprises a community nutritionist, a health promotion officer, a representative of one of the local shires, and a member of the local chamber of commerce with a knowledge of food distribution and retailing. (All members of the review team are named and their titles and positions or affiliations are identified).

CASE EXAMPLE

Are school-based programs effective in preventing youth suicide? The review to answer this question is intended to be included in the Cochrane Library of Systematic Reviews. The project is being done on a voluntary basis (without specific funding) by two psychiatrists, one of whom has received training from the Cochrane Collaboration, and an educational psychologist employed in the State school system. (All members of the review team are named, and their titles and positions or affiliations are identified).

Note: when you have clarified the purpose and scope of your review in Section 1, move on to the questions in Sections 2 to 4.

Any questions that do not appear to be relevant to your needs can be considered optional. You may also consider omitting questions in Sections 2 to 4 in order to reduce the time it will take to apply the Schema

However, if you do omit questions, you should weigh the benefit of saving time against the potential value of applying the Schema in full.

Section 2: The Papers in the Review

2A PUBLICATION DETAILS

Identify the publication details for each paper or report to be appraised using the Schema.

Include the following information.

- Title
- Authors
- Date
- Publication information, e.g. name of journal, volume and page numbers or publisher. If the report is unpublished, this fact should be noted, and the agencies that commissioned and produced the report should be noted (often only one agency, such as a government department, will have been involved.)
- Type of article

For example, is this a peer-reviewed journal article (primary evaluation or review), a report from a government department, a non-government agency publication or another type of publication? (specify)

- Have related papers or reports been published? If so, what are they?

For example, you may be aware of prior or later evaluations of the same program from the references given in the articles, or from your own knowledge of the literature. If you have not yet examined these articles, note the citations, and decide whether you will include the articles in your review.

2B SPECIFYING THE INTERVENTION

(QUESTIONS 2.1 TO 2.4)

2.1 Exactly what intervention was evaluated in the study?

Look for the following information to describe the intervention fully.

- What was/were the main intervention(s)? Note whether the intervention(s) involved multiple strategies or a singular strategy. If multiple strategies were evaluated, how were they combined?
- What problem was the intervention intended to address?
- Where and when did the intervention take place? [cross-refer to 2.5]
- Was it a one-off, time-limited intervention, or was it an ongoing or repeated intervention?
- For how long was the intervention implemented?
- Were different types and different levels of intervention studied?

For example, printed materials compared to a television campaign to promote sun protection; or one week compared to three weeks of a daily television campaign.

- Whom did the intervention target?
- Who implemented the strategies used (what skills and/or qualifications did they have)?

CASE EXAMPLE

The intervention was a campaign to promote healthy eating among all staff in a teaching hospital in the western suburbs of Sydney. The intervention consisted of two strategies: a change to the food available for sale in the staff canteen (more low-fat, high-fibre and vegetarian dishes) and halving the price of low-fat and high-fibre food. A trial of the intervention was conducted for a six-month period from January to June 1997. The strategies evaluated in 1997 in this study were developed and implemented by the hospital Catering Department (catering manager and cooks), in consultation with the Department of Nutrition and Dietetics.

2.2 What was origin of the intervention?

Examine how the intervention came to be implemented. In your description of the origin of the intervention, consider the following:

- Who was responsible for deciding that the intervention should occur?
- Was the intervention planned and developed for identifiable health policy and/or program objectives?
- Was the intervention carried out for the purposes of research?
- Did the intervention evolve without a formal research or health policy framework? If so, how did it come about?
- Was the target community involved in initiating and/or developing the intervention? To what degree?

CASE EXAMPLE

The intervention started out as a youth work-experience scheme. It was introduced in 1996 by a regional chamber of commerce, and it involved local business interests. Members of the local youth centre participated in planning and promoting the scheme. The intervention was not originally planned to fulfil health objectives, and was not implemented for research purposes. Rather, the scheme evolved from the community's desire to address local issues of youth unemployment. Local business owners were concerned about the disenfranchisement and disengagement of unemployed teenagers from the rest of the community, resulting in vandalism and graffiti in the business district. A few years later the State Health Department considered the potential of this scheme as an intervention to reduce youth suicide and commissioned an evaluation of its effectiveness for this purpose.

CASE EXAMPLE

The intervention was a program carried out in another country (specified) to alter infants' sleeping positions to reduce the incidence of cot death. The program was implemented by that country's Health Department following the publication of several observational studies suggesting a link between infants' sleeping position and cot death. It was a policy initiative that was planned and implemented as a result of published research evidence and media reporting of that evidence.

2.3 If the origin of the intervention involved a degree of formal planning, what was the rationale for the strategies selected?

When describing the rationale for the intervention consider what information is available about the following:

- What reasons were given in the article for selecting the particular intervention strategies that were used, rather than other strategies?

For example (with reference to the case example described in 2.2), why was parent education in post-natal wards and through early childhood health centres chosen as the intervention, rather than a mass media campaign?

- Did the authors conduct a review of the evidence (i.e. did they examine evaluation research conducted on similar interventions) or cite existing systematic reviews of evidence?
- Was a formal theory identified as a basis for the intervention strategies adopted (e.g. educational, behavioural, environmental or community development theories)?

2.4 What organisations or individuals sponsored the intervention (with funding or in-kind contributions)? Where relevant, give details of the type of sponsorship provided.

2C IDENTIFYING THE INTERVENTION CONTEXT (QUESTIONS 2.5 TO 2.7)

2.5 What aspects of the context in which the intervention took place were identified in the article?

Examine the information provided in the paper or report about the context in which the intervention was implemented. Include planned and unplanned changes to the context that supported the intervention, as well as those that may have hindered the intervention. Look for the following information.

- The time and place of the intervention [cross-refer to **2.1**].
- The local policy environment, including institutional policies and management support for the intervention.
- The broader political environment.
- Other social and cultural factors, concurrent social changes or social movements.

- Other organisational factors, structural or physical environment.
- Economic climate and the availability of resources.
- Training, skills and experience of those implementing the intervention.

CASE EXAMPLE

The hospital-based canteen intervention described in 2.1 above was a part of a five-year program (1995–2000) to develop and implement a hospital-wide food and nutrition policy for staff and patients. The food and nutrition policy was initiated by the hospital general manager and supported by other senior managers, clinical staff and staff unions. The policy was prepared by the Department of Nutrition and Dietetics in consultation with hospital staff and the local community between 1995 and 1996. Extra resources were allocated for implementing the policy in the hospital budget (although the amount is not specified in the paper). Catering staff received two half-days of training about the modifications implemented in their canteen. The intervention coincided with two articles in the local newspaper about healthy eating. Although the newspaper articles were not planned as part of the intervention (they were initiated by the journalist), the authors comment on their contribution to raising awareness among the hospital staff target group.

2.6 Was enough information provided in the article to enable you to describe the intervention and its context as requested above?

If not, identify the major deficiencies in the information provided.

2.7 How relevant to the scope of your review (as recorded in Section 1) are the intervention and the context described in this article? Give details.

- **If you conclude that the article is relevant (or partly relevant) to the scope of your review, go to Sub-section 2D.**
- **If you conclude that the article is not relevant to the scope of your review, record why it is not relevant, and then move on to the next paper or report to be appraised.**

2D THE EVALUATION CONTEXT – BACKGROUND, PURPOSE AND QUESTIONS ASKED

(QUESTIONS 2.8 TO 2.20)

In addition to describing the intervention and its context, it is important to consider the background and purpose of the evaluation of the intervention. An understanding of the context in which an evaluation was conducted will help you to interpret the evaluation findings.

2.8 Who requested or commissioned the evaluation and why?

The aims of an evaluation determine the research questions that are asked, the measures taken, and the qualitative observations that are made. In attempting to describe why the evaluation was conducted, look for the following information:

- Is enough information provided for you to determine the overall purpose and framing of the evaluation that was conducted?
- Was the evaluation conducted for the purpose of research, for a decision about funding or continued funding of an intervention, to support an identified course of action, or for other reasons? (specify)
- Were the questions framed from a researcher perspective, a policy maker's perspective, a community interest group perspective, or some other perspective? (specify)

2.9 Exactly what research questions were asked in the evaluation reported in the study?

2.10 What measures of effect or intervention outcomes were examined?

For example, evaluation of an immunisation program may focus on one or more of the following outcomes – access to immunisation services, distribution and appropriate storage of vaccines, the delivery of immunisation services, immunisation rates, or the effect of an intervention on occurrence of vaccine-preventable disease.

- Were the measures of effect, or intervention outcome measurements, clearly identified in the article, and were they adequately defined?
- How well did the observations or measures address the research questions asked in the study?

2.11 What was the anticipated sequence of events between the intervention strategies and the measures of effect or intended intervention outcomes?

This may also be referred to as the intervention's mode of action or causal pathway, or the evaluation's program logic.

CASE EXAMPLE

A canteen-based nutrition campaign used price changes in the canteen to promote the purchase of low-fat and high-fibre foods and other healthy food choices. Thus it was anticipated that canteen price changes would lead to canteen users purchasing more of the healthy foods and less of the unhealthy foods. This substitution would contribute to an overall increase in daily fibre intake and a reduction in daily fat consumption among the hospital staff who eat at the canteen (program logic). These effects would be identified in pre- and post-intervention 24-hour food records to determine daily fibre and fat consumption (measures of effect).

2.12 Were the measures of effect or intervention outcomes achievable and compatible with the sequence of events outlined in 2.11?

Consider whether the anticipated intervention mode of action (causal pathway) was realistic, and thus whether the measures of effect or intervention outcomes were theoretically achievable, that is, whether the evaluation program logic was sound.

For example, it may be realistic to measure the effect of an education program on participants' knowledge and skills, but unrealistic to expect an effect on health outcomes, particularly if there is no change in participants' behaviours, and if other determinants of health remain the same.

2.13 What was the timing of the evaluation in relation to the implementation of the intervention?

Evaluations conducted within research frameworks are often planned in advance of the implementation of the intervention. Conversely, interventions that emerge from prevailing circumstances are often evaluated long after they were initially conceived and implemented. The timing and sequence of events in an evaluation will affect the type and quality of data collected, e.g. whether evaluators are able to collect pre-intervention baseline data.

- Was the evaluation planned before or after the intervention was implemented?
- Did the evaluation begin before or after the intervention was implemented?

2.14 Was the intervention adequately implemented in the setting in which it was evaluated?

Knowing how well an intervention and its various strategies were implemented will help you to interpret the evaluation findings. This is particularly important if the evaluation findings are negative or inconclusive. For example, if an intervention was poorly implemented, an evaluation study would be unlikely to show that the intervention was effective.

Look for information to answer the following questions.

- Were monitoring and process measures collected about the implementation of the intervention? If so, specify.
- Did the intervention implementation differ from what was planned? If so, what was the rationale or explanation for the difference?
- Did the evaluators check whether the intervention had reached the correct target group, and reached a sufficient number of members of the target group, before process and outcome measurements were made?
- Was enough information provided in the report to enable you to determine how well, overall, the intervention was implemented?

2.15 Was the intervention ready for the type of evaluation that was conducted?

Consider the following information:

- What was the stage of development of the intervention at the time of its evaluation?

For example, some interventions go through several stages of development and evaluation. An intervention developed for the first time may undergo detailed process evaluation. Interventions that have been shown to be effective in some settings may undergo dissemination evaluation to determine whether they can be implemented more widely.

- Was a pilot study or ‘process’ evaluation done before an ‘impact’ or ‘outcome’ evaluation?
- Have evaluations of similar interventions been done in other settings that have shown the intervention to be effective? If so, were the references and findings cited in the article?

CASE EXAMPLE

Canteen-based pricing and product strategies to promote healthy eating have been tested in the USA and Europe and been shown to have an effect on staff food choices (references given). The intervention evaluated in this study was pilot tested in 1996 (reference given), and the local strategies appeared to be ready for the impact and outcome evaluation reported in the paper.

2.16 Were the measures of effect or intervention outcomes validated or pilot tested? If so, how?

2.17 Did the observations or measures include the important individual and group-level effects?

For example, a campaign to promote healthy eating could be assessed by its effect on the diet of defined individuals (individual-level effects) or by its effect on supermarket sales of certain food items (group-level effects).

- If important measures were missing, specify which ones.

2.18 Was there a capacity to identify unplanned benefits and unanticipated adverse effects?

It is possible to look for unplanned or unanticipated effects by including exploratory observations or measures in an evaluation.

- If so, was there some scope to identify unplanned or unanticipated effects among disadvantaged or minority groups, such as Aboriginal or Torres Strait Islander communities?

2.19 If the research was not primarily an economic evaluation, were economic factors considered?

Look for information to answer the following questions:

- How much did the intervention cost?
- What resources were needed to implement the intervention?
- Was value for money assessed? Were intervention costs and resources weighed against intervention benefits (or harms)?

2.20 Was there a significant potential for conflict of interest (in the way the intervention and/or its evaluation were funded and implemented) that might affect interpretation of the findings?

Consider the following questions:

- What was the nature of the relationship (if any) between those involved in planning, implementing and funding the intervention (their objectives and interests) and those conducting the evaluation of that intervention (their objectives and interests)?
- What is the potential effect of such a relationship on the study findings?
- Was enough information provided to make an assessment about potential conflicts of interest? If not, what information is missing?

2E THE METHODS USED TO EVALUATE THE INTERVENTION

(QUESTIONS 2.21 TO 2.27)

2.21 What types of research methods were used to evaluate the intervention?

For example, were epidemiological methods used, or qualitative methods, or economic evaluation methods, or a combination of methods?

2.22 What study designs were used in the evaluation?

2.23 How appropriate were the research methods and study designs in relation to the questions asked in the study?

Take into account the intervention context and the evaluation questions asked about the intervention to determine if the study design was appropriate, or if it could have been improved. Consider the following:

- What are the best methods to answer the research questions?
- What would the best feasible study design have been?
For evaluations that aim to quantify intervention effectiveness, take into account the hierarchy of study designs outlined in Appendix 1.
- Did the study design used differ from the best feasible study design? If so, how?
- Were reasons given for the choice of the study design that was used? If so, what were they?

2.24 Was the evaluation conducted from a single perspective or multiple perspectives? Give details.

Multiple perspectives can strengthen an evaluation. Multiple perspectives can be achieved by combining research methods, using a number of different observations or measures, taking effect and outcome measures over several time points, and comparing the perspectives of multiple observers.

2.25 Appraise the rigour of the research methods used in the study using the relevant checklist(s) in Appendix 2.

Appendix 2 includes several guides for critically appraising the rigour of research methods. You should select the relevant critical appraisal guide according to the method(s) used in the study.

- Systematic review on intervention effectiveness (use Supplementary Guide 1)
- Randomised Control Trial (Supplementary Guide 2)
- Observational study (Supplementary Guide 3)
- Economic evaluation (Supplementary Guide 4)
- Qualitative study (Supplementary Guide 5)

Evaluations of public health interventions often rely on a combination of research methods (e.g. quantitative and qualitative). Where two or more methods (or study designs) are reported, you will need to refer to more than one critical appraisal guide. This may highlight discrepancies in the quality of different components of the study, or discrepancies in the adequacy with which those components were reported.

Once you have appraised the study using the relevant checklist(s) go to question 2.26.

2.26 What are your conclusions about the adequacy of the design and conduct of the research methods used to evaluate the intervention?

Base your conclusions on your answers to questions 2.21 to 2.24 and your application of the checklists in question 2.25.

2.27 Are the reported findings of the evaluation likely to be credible?

Take into account your conclusions about the methods used in the study (2.26), as well as any other limitations to the evaluation, and potential conflicts of interest, identified in Sub-section 2D.

Consider the following in your answer:

- Are the measures of effect or intervention outcomes likely to be believable?
- Are observed effects attributable to the intervention, or do other possible explanations exist?

■ **If you conclude that the reported findings are (or may be) credible go to Section 3.**

■ **If you conclude that the findings are not (or unlikely to be) credible, go to Section 4 and answer question 4.2 only, and then move on to the next paper or report to be appraised.**

Section 3: Describing the results from the papers selected

3.1 What findings were reported in the study?

Look for the following information:

- What were the beneficial and adverse effects? Where relevant, give the estimates of effect and their confidence intervals.
- Were there any unanticipated effects? (i.e. did the intervention have an effect on the community or setting other than the stated objectives?)
- Were there any side effects or harms for the target group, for other groups in the community, or for those recruited to carry out the intervention?
- What was the cost and cost effectiveness of the intervention?

3.2 If the study specified measurable or quantifiable targets, did the intervention achieve these objectives?

For example, an objective of an intervention may be to achieve a 50% reduction in the prevalence of smoking, but the results report a 20% reduction, therefore the study did not achieve its stated objectives, although the findings were statistically significant.

3.3 Were reported intervention effects examined among sub-groups of the target population?

Sub-groups may be identified by (for example) gender, age, socio-economic status, ethnicity, or culture.

If sub-group effects were examined, answer the following questions:

- What sub-groups were identified and explored?
- What were the reasons given for the sub-groups examined?
- Were the relative effects on disadvantaged groups assessed?
- How did the effects of the intervention differ among the sub-groups explored?
- Were reasons for different intervention effects in the sub-groups explored? If so, what was the explanation given?
- Was the sub-group analysis a pre-planned component of the evaluation or was it a post-hoc analysis of the data?

3.4 Should any other important sub-group effects have been considered that were not considered?

If yes, identify important sub-group effects that should have been considered.

For example, where relevant, did the evaluation assess the relative effectiveness of the intervention on Aboriginal or Torres Strait Islander groups?

3.5 Was the influence of the intervention context on the effectiveness of the intervention investigated in the study?

For example, researchers may report whether an intervention was easier or more difficult to implement in some settings compared to others. They may also report whether it was more or less effective in different physical settings or social circumstances. Insufficient detail about the context is often an important limitation in the evidence when it comes to appraising the applicability of the study findings and the transferability of the intervention to other settings.

Include the following in your answer:

- Was enough detail provided about the influence of the context on the effectiveness of the intervention so that enabling factors could be replicated and adverse effects avoided in the future?
- Should any potentially important contextual influences have been examined that were not examined? If so, what are they?

3.6 How dependent on the context is the intervention described in the article?

Intervention strategies and intervention effects vary in the degree to which they are context dependent.

For example, the efficacy of a screening procedure may not depend much on the social context, while a campaign promoting its uptake by the general public is likely to be highly dependent on the context.

3.7 Were the intervention outcomes sustainable?

It is quite common for evaluations to measure effects immediately after an intervention, but not to consider the sustainability of those effects. This is often an important gap in the evidence. Consider the following:

- For how long were data collected about the effects of the intervention?
- How long were the reported intervention effects actually sustained?

3.8 Did the study examine and report on the value of the measured effects to parties interested in or affected by them?

A specific target group, the wider community, local practitioners, fund holders or policy makers are all likely to place different values on reported intervention effects. The evaluation of an intervention may include the step of presenting its findings to interested parties and asking for feedback on the utility and value of those findings.

For example, with reference to a particular intervention, practitioners may be interested in the clinical importance of the effects of the intervention, while managers may focus on obtaining value for the money spent on the intervention, and policy makers may be interested in the extent to which intervention effects support a political agenda.

- If there was a report on the value of the reported effects, was this speculative on the part of the authors, or based on empirical data? Give details.

Section 4: Interpreting each article

4.1 How well did the study answer your review question(s)? Give details.

4.2 Are there other lessons to be learned from this study?

For example, the article may not have answered your question about the effectiveness of a particular intervention, but you may identify useful lessons relating to the development and implementation of this type of intervention, or factors that would improve future evaluations of similar interventions.

If you are conducting the review for the purpose of making recommendations for a particular policy or practice setting answer questions 4.3 to 4.8, otherwise go to Section 5.

4.3 Are the essential components of the intervention and its implementation described with sufficient detail and precision to be reproducible?

4.4 Is the intervention context, as described in the article being examined, comparable to the intervention context that is being considered for future implementation of the intervention?

- If the intervention context is similar, what are the similarities?
- If the intervention context is different, what are the differences?
- Is your capacity to implement the intervention in the new setting comparable to the capacity described in the article?

For example, the capacity to implement the intervention may depend on resources, skills of local people, organisational factors, and/or the policy and political environment.

4.5 Are the characteristics of the target group studied in the article comparable to the target group for whom the intervention is being considered?

The effectiveness of an intervention can be highly dependent on characteristics of target communities, such as education level, social and economic status, and access to basic services such as housing, water, power supply, sanitation and health care.

- If the target groups are similar, what are the similarities?
- If the target groups are different, what are the differences?
- If there are significant differences, what are the potential implications for the applicability of the research? Why are these implications important?
- Overall, can the reported findings be applied to the new population for whom the intervention is being considered?

For example, evidence demonstrating the effectiveness of an antenatal program for urban women with Anglo-European origins may not be applicable to women in rural Aboriginal communities.

4.6 If an economic evaluation was conducted, did the paper or report include and address the details required in order to make an informed assessment about the applicability and transferability of the findings to other settings?

4.7 If enough information was provided, are the findings of the economic evaluation relevant and transferable to your setting?

4.8 Are the effects of the intervention likely to be considered important in your setting?

Consider the following:

- The perspectives of those who may be affected by the intervention (there may be quite diverse views among different community groups).
- The perspectives of those who will make decisions about its future funding and implementation.

You may need to consider the magnitude of the intervention effects as well as their statistical significance, the individual (clinical) and population-wide importance of the effects, and also their social, political and economic importance.

Section 5: Summarising the body of evidence

You are now ready to make a judgement on the entire body of evidence (i.e. all the papers and reports that you have reviewed). It is desirable for this process to be a team effort. Your judgement will be based on your consideration of individual papers and reports as described in Sections 2, 3 and 4. You should relate your judgement to the original objectives of your review, as identified in Section 1.

Section 5 has two parts – Sub-sections 5A and 5B. In Sub-section 5A you are asked to:

- group the articles according to their research questions and intervention strategies;
- grade the overall quality of the studies within each group;
- assess the consistency of the findings among the stronger studies within each group; and
- determine the applicability of the studies to your review.

Your formulations in 5A will help you in 5B to reach a conclusion about the strengths and weaknesses of the evidence for your purposes (as specified in Section 1), and to identify gaps in the evidence.

5A GROUPING, RATING AND WEIGHING UP THE PAPERS AND REPORTS

(INSTRUCTIONS 5.1 TO 5.4)

To reach an overall formulation, you should group together papers and reports in which similar research questions are asked and similar intervention strategies are used (5.1). You should then work through points 5.2 to 5.4 for one group of articles, before moving on to the next group.

5.1 Grouping articles by research questions and intervention strategies

Group articles with similar research questions and similar intervention strategies. For each group, enter the research question and a brief description of the intervention into the first column of the summary table in Appendix 3.

The selection of articles for each group will be determined by the questions that you have posed for the review. For example, you might form two groups of papers which ask ‘How effective are pricing modification strategies for reducing dietary fat consumption?’ One group would comprise articles investigating pricing strategies applied in workplace canteens, and the other group would comprise articles investigating pricing strategies applied in supermarkets.

5.2 Quality of studies

Assess the quality of each study in each group on a scale from 1 (poor) to 5 (excellent). This assessment should take into account the adequacy of information provided about the intervention and its evaluation and the design and conduct of the study (i.e. the methods). Mark each study on the scale provided on the summary form (Appendix 3, second column).

5.3 Consistency of findings

Assess and grade the consistency of the findings among the stronger studies in each group on a scale from 1 (findings among the stronger studies are inconsistent) to 5 (findings among the stronger studies are highly consistent with each other). Mark your assessment on the scale provided on the summary form (Appendix 3, third column).

5.4 Applicability of studies

Determine whether or not the stronger studies, which have consistent findings, can be applied to your review context, which you have described in Section 1. Categorise applicability as:

- applicable
- possibly applicable
- not applicable

5B FORMULATING A SUMMARY STATEMENT (QUESTIONS 5.5 TO 5.10)

When you have completed 5.1–5.4 for all of the articles which you reviewed, you will have developed a ‘map’ of the available evidence. Use this information to formulate a summary statement on the body of evidence by answering the questions below.

5.5 Did studies that examined similar intervention strategies, with similar research questions, produce consistent results?

5.6 Did studies with different research questions produce compatible results?

5.7 Overall, what does the body of evidence tell you about the intervention?

In your summary, include the following:

- Has the intervention been shown to work? To what degree?
- In what context has it been shown to work?
- For whom did it work, and for whom did it not work?
- Can you identify factors that are necessary for similar interventions to be effective in other settings?

5.8 Are there important gaps in the evidence? If so, what are they?

5.9 To what degree are the review findings useful for your purposes, as identified in Section 1?

Consider the following:

- How well do the review findings answer your review question(s)?
- If relevant, are the review findings applicable and useful for your policy or practice setting?

5.10 What are your recommendations based on this review?

Formulate recommendations according to the purpose of the review identified in Section 1. If appropriate, you may wish to consider making recommendations on:

- Further research and evaluation on the public health intervention(s) examined.

This may be strongly recommended if the review findings are inconclusive. For example, if the evaluation methods were generally poor, but the described intervention seems promising, you may recommend a better evaluation of such an intervention in another setting. If the evaluation methods were good but the intervention was inadequately planned and implemented, you may recommend better development of the intervention prior to further evaluations.

Alternatively, the effectiveness of the intervention may still need to be evaluated among disadvantaged groups such as Aboriginal and Torres Strait Islander communities, or there may be a need for research to examine the effect of the intervention on inequities in health status.

- Other reviews of available evidence

For example, you may identify related review topics on which there appears to be available literature that has not yet been reviewed.

- Current or future public health policy and practice

For example, your review findings may indicate that it would be desirable to modify current public health policy or practice, or your findings may strongly support the continued funding of existing interventions.

Appendix 1: Hierarchies of study design and designation of levels of evidence

(REFER TO QUESTION 2.23 IN THE SCHEMA)

Study design	Level of evidence*
Systematic review of all relevant randomised control trials (RCT)	I
Properly designed RCT	II
Well-designed pseudo-randomised controlled trial (e.g. alternate allocation)	III-1
Comparative studies (or systematic reviews of such studies) with concurrent controls and allocation not randomised, cohort studies, case-control studies, or interrupted time series with a control group	III-2
Comparative studies with a historical control, two or more single arm studies, or interrupted time series without a parallel control group	III-3
Case series, post-test or pre-test/post test, with no control group	IV

* Lower numbers indicate a higher level of evidence; higher numbers indicate a greater potential for bias.

This hierarchy of study designs (also commonly referred to as “levels of evidence”) is used by NHMRC and others^{9,10} to indicate the relative potential for bias in alternative epidemiological studies (there is an increasing potential for bias down the table).

It is important to be aware of the relative strengths of alternative epidemiological designs. However, when deciding whether a study is “good enough” to inform your conclusions about an intervention (Sub-section 2E of the Schema), it is also important to note that in some public health settings it may only be feasible, or politically and/or ethically acceptable, to conduct observational studies.

Appendix 2: Supplementary Guides 1 to 5

(REFER TO QUESTION 2.25 IN THE SCHEMA)

SUPPLEMENTS TO QUESTION 2.25

Supplementary Guide 1: Appraising reviews

Supplementary Guide 2: Appraising randomised controlled trials

Supplementary Guide 3: Appraising observational studies

Supplementary Guide 4: Appraising economic evaluations

Supplementary Guide 5: Appraising qualitative studies

Supplementary Guide 1: Appraising reviews

A SUPPLEMENT TO QUESTION 2.25

This is a checklist of questions to help you decide if the methods used in an existing review were rigorous.

When you have appraised the methods using this supplementary guide, return to question 2.26 of the main Schema.

This guide has been adapted from checklists available from:

The Critical Appraisal Skills Program (CASP), Public Health Resource Unit, Institute of Health Sciences, Oxford, UK, <http://www.phru.org.uk/~casp/resources/>

Undertaking Systematic Reviews of Research on Effectiveness, CRD's Guidance for those Carrying Out and Commissioning Reviews. Stage I, Appraising available reviews. <http://www.york.ac.uk/inst/crd/report4.htm>

Oxman AD et al. Users' Guides to the Medical Literature, VI How to use an overview. JAMA 1994; 272(17): 1367–1371.

G1.1 Did the review address a clearly focused research question?

A research question should be 'focused' in terms of the:

- population studied
- intervention given (exposure)
- outcomes considered

G1.2 Did the review include the right type of studies?

Studies should:

- address the review's research question
- have an appropriate study design

G1.3 Did the reviewers try to identify all relevant primary studies?

Did they:

- search all relevant bibliographic databases?
- follow-up from reference lists?
- use personal contact with experts?
- search for unpublished studies?
- search for non-English studies?

G1.4 What criteria were used to include or exclude primary studies and how were they applied?

G1.5 Did the reviewers assess the quality of the primary studies included in the review?

Did they:

- use explicit criteria?
- check the criteria were applied appropriately and consistently?

G1.6 How were the data from the primary studies synthesised?

- Were differences between studies investigated?
- Were the reasons for any variations in results discussed?
- Were the results of the primary studies combined (pooled)? If so, was it reasonable to combine the results given the similarities and differences between the studies reviewed?
- Do the reviewers' conclusions flow from the evidence examined?

PLEASE RETURN TO QUESTION 2.26 IN THE MAIN SCHEMA.

Supplementary Guide 2: Appraising randomised controlled trials

A SUPPLEMENT TO QUESTION 2.25

This is a checklist of questions to help you decide if a randomised controlled trial was rigorous.

When you have appraised the methods using this supplementary guide, return to question 2.26 of the main Schema.

This guide has been adapted from checklists available from:

The Critical Appraisal Skills Program (CASP), Public Health Resource Unit, Institute of Health Sciences, Oxford, UK, <http://www.phru.org.uk/~casp/resources/>

Guyatt GH, Sackett DL, Cook DJ. Users' Guides to the Medical Literature. II. How to use an article about therapy or prevention. JAMA 1993; 270: 2598–2601 and 271: 59–63

G2.1 Was the assignment of study participants to the intervention (exposure) randomised?

- Was the method of randomisation appropriate?

Note: study participants can be randomised individually or as a 'cluster' (group). The method of randomisation is appropriate if it allows each study participant (or cluster) to have the same chance of receiving each intervention option. Investigators determining participants' eligibility and allocating interventions should not be able to predict which intervention is next in line. Alternate methods of allocation are not random.

G2.2 Were all the participants who entered the trial accounted for at its conclusion?

- How complete was the follow-up of study participants?
- Were participants analysed in the groups to which they were randomised?

Note: participants who were included in the study but did not complete the observation period or who were not included in the analysis should be described. The number and the reasons for withdrawal in each group must be stated.

G2.3 Were all the groups similar at the start of the trial?

Consider factors, other than the intervention, that might affect the outcome, such as age, gender, social class or ethnicity.

G2.4 Aside from the intervention being tested, were the groups exposed to similar influences before and during the trial?

Consider environmental factors that were not part of the intervention but which may effect the outcome, e.g. reports in local media or concurrent political developments.

G2.5 Aside from the intervention being tested, were the groups treated equally?

- Were the groups reviewed at the same time intervals?
- Did they have the same follow-up procedures?

PLEASE RETURN TO QUESTION 2.26 IN THE MAIN SCHEMA.

Supplementary Guide 3: Appraising observational studies

A SUPPLEMENT TO QUESTION 2.25

This is a checklist of questions to help you decide if the methods used in observational studies (cohort, case control or case series) were rigorous.

When you have appraised the methods using this supplementary guide, return to question 2.26 of the main Schema.

This guide is an extract from a report produced by The University of York NHS 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) March 2001. Stage II, Conducting the Review, Phase 5 Study Quality Assessment. Khalid S Khan, Gerben ter Riet, Jennie Popay, John Nixon & Jos Kleijnen.

The full text of this report can be viewed at:
<http://www.york.ac.uk/inst/crd/report4.htm>

Section 2.5.5.2 of this report includes notes on appraising observational studies.

COHORT STUDIES

- G3.1 Is there sufficient description of the groups and the distribution of prognostic factors (which may affect the outcome)?
- G3.2 Were the groups assembled at a similar point in their disease (or risk factor) progression?
- G3.3 Was the exposure to the intervention reliably ascertained?
- G3.4 How comparable were the groups in all important confounding factors?
- G3.5 Was there adequate adjustment for the effects of these confounding variables?
- G3.6 Was a dose–response relationship between intervention and outcome demonstrated?
- G3.7 Was outcome assessment blind to exposure status?
- G3.8 Was follow-up long enough for the outcomes to occur?
- G3.9 What proportion of the cohort was followed up?
- G3.10 Were drop-out rates and reasons for drop-out similar across intervention and unexposed groups?

CASE CONTROL STUDIES

- G3.11 Is the case definition explicit and adequate?
- G3.12 Has the disease-state (outcomes) of the cases been reliably assessed and validated?
- G3.13 Were the controls representative of the same population source as the cases?

- G3.14 How comparable are the cases and controls with respect to potential confounding factors?
- G3.15 Were interventions and other exposures assessed in the same way for cases and controls?
- G3.16 How was the response rate defined?
- G3.17 Were the non-response rates and reasons for non-response the same in both groups?
- G3.18 Is it possible that over-matching has occurred in that cases and controls were matched on factors related to the exposure (intervention)?
- G3.19 Was an appropriate statistical analysis used (matched or unmatched)?

CASE SERIES

- G3.20 Is the study based on a representative sample selected from a relevant population?
- G3.21 Are the criteria for inclusion explicit?
- G3.22 Did all individuals enter the survey at a similar point in their disease (or risk factor) progression?
- G3.23 Was follow-up long enough for important events to occur?
- G3.24 Were outcomes assessed using objective criteria or was blinding used?
- G3.25 If comparisons of sub-groups are being made, was there sufficient description of the sub-groups and the distribution of prognostic factors?

PLEASE RETURN TO QUESTION 2.26 IN THE MAIN SCHEMA.

Supplementary Guide 4: Appraising economic evaluations

A SUPPLEMENT TO QUESTION 2.25

This is a checklist of questions to help you decide if the methods used in an economic evaluation were rigorous.

When you have appraised the methods using this supplementary guide, return to question 2.26 of the main Schema.

This guide is based on checklists available from:

The Critical Appraisal Skills Program (CASP), Public Health Resource Unit, Institute of Health Sciences, Oxford, UK, <http://www.phru.org.uk/~casp/resources/>

Drummond MF, Stoddard GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programs. Oxford, Oxford University Press, 1987.

G4.1 Was a well-defined question posed?

- Is it clear what the authors were trying to achieve?

G4.2 Was a comprehensive description of the competing alternatives given?

- Can you tell who did what to whom, where and how often?

G4.3 Does the paper provide evidence that the program would be effective?

- Does the program do more good than harm?
- How good is the evidence of effectiveness?

G4.4 For each alternative option, were all the important and relevant consequences (in terms of resource use and health outcome):

a) identified?

What perspectives were taken?

b) measured accurately in appropriate units prior to evaluation?

c) valued credibly?

Have opportunity costs been considered?

G4.5 Were resource use and health outcome consequences adjusted for different times at which they occurred (discounting)?

G4.6 Was an incremental analysis of the consequences and costs of alternatives performed?

G4.7 Was an adequate sensitivity analysis performed?

- Were all the main areas of uncertainty considered?

PLEASE RETURN TO QUESTION 2.26 IN THE MAIN SCHEMA.

Supplementary Guide 5: Appraising qualitative studies

A SUPPLEMENT TO QUESTION 2.25

This is a checklist of questions to help you decide if the methods used in a qualitative study were rigorous.

When you have appraised the methods using this supplementary guide, return to question 2.26 of the main Schema.

This guide has been adapted from a checklist available from:

The Critical Appraisal Skills Program (CASP), Public Health Resource Unit, Institute of Health Sciences, Oxford, UK, <http://www.phru.org.uk/~casp/resources/>

G5.1 Was there a clear statement of the aims of the research?

Can you tell:

- what the authors were trying to find out?
- why it was important?
- how it was relevant?

G5.2 Was a qualitative method appropriate to address the aims of the research?

- Did the research seek to understand, illuminate or explain the subjective experience or views of those being researched in a defined context or setting?

G5.3 Was the sampling strategy appropriate to address the aims of the research?

- Is it clear
 - a) from where the sample was selected and why?
 - b) who was included in the sample and why?
 - c) how they were selected and why?
 - d) whether some people approached chose not to participate and why?
- Was the sample size justified?

G5.4 Were the data collected appropriately to address the research issue?

- Is it clear:
 - a) where the data were collected and why that setting was chosen?
 - b) how the data were collected and why (focus group, interview, observations etc.)?
 - c) how the data were recorded and why?
 - d) if the methods of data collection were modified during the process and why?

G5.5 Was the data analysis sufficiently explicit?

- Is it clear:
 - a) how the analysis was done?
 - b) how the categories/themes/concepts were derived from the data (i.e. is there adequate description of the researchers process and analytical decisions)
- Are you confident that all relevant data were taken into account?

G5.6 Was the data analysis sufficiently rigorous?

- Have steps been taken to assess the credibility of the findings?
- Was there adequate discussion of the evidence both for and against the researchers' arguments.
- Did researchers attempt to feed their results back to the study participants?
- Where appropriate, did they use and compare different sources of data about the same issue (triangulation)?
- Was the analysis repeated by more than one researcher to ensure reliability?

G5.6 Has the relationship between the researchers and participants been adequately described and considered?

- Did the researchers critically examine their own role, potential influences and biases?
- How was the research explained to the participants?

G5.7 Is there a clear statement of the findings?

- Are the findings explicit and easy to understand?

G5.8 Did the researchers indicate links between their presented data, their findings, and their conclusions?

PLEASE RETURN TO QUESTION 2.26 IN THE MAIN SCHEMA.

Appendix 3: A table for summarising all the papers reviewed

(REFER TO SUB-SECTION 5A)

Note: this summary table is an example only. To summarise your own review you may need to enlarge the table, or expand it with additional rows. Alternatively you may wish to modify its layout to suit your own needs.

<p>Group papers by research question and intervention strategies examined</p>	<p>Rate quality of individual studies 1 _____ 5 Scale: 1 (poor) to 5 (excellent)</p>	<p>Decide on consistency among good quality studies 1 _____ 5 Scale: 1 to 5 1 (findings are inconsistent) 5 (findings are highly consistent)</p>	<p>Decide on applicability of the consistent findings Responses: • Applicable • Possibly applicable • Not applicable</p>
<p>Group 1 What are the effects of pricing strategies applied in workplace canteens on dietary fat intake of staff?</p>	<p>11 papers in group 1 Rating 1 = papers 1, 7, and 11 Rating 2 = papers 2, 3, and 10 Rating 3 = paper 8 Rating 4 = papers 6, 9 and 5 Rating 5 = paper 4</p>	<p>Consistency of findings among papers 6,9 and 4 is rated '4' Consistency between papers 4 and 5, 6, 9 is rated '3'</p>	<p>The findings of papers 6, 9 and 4 are 'Applicable' to our review context</p>
<p>Group 2 What are the effects of pricing strategies applied in supermarkets on dietary fat intake of local customers?</p>			

References

- 1 Rychetnik L, Frommer M. A Proposed Schema for Evaluating Evidence on Public Health Interventions. National Public Health Partnership. May 2000. <http://www.nphp.gov.au/ppi/evidence/isspaper/>
- 2 Rychetnik L, Frommer M, Hawe P, Shiell A. Criteria for evaluating evidence on public health interventions. *Journal of Epidemiology and Community Health* 2002; 56: 119–127.
- 3 National Health and Medical Research Council (NHMRC). “How to Review the Evidence: Systematic Identification and Review of the Scientific Literature.” AusInfo, Canberra, 2000.
- 4 NHS Centre for Reviews and Dissemination. “Undertaking Systematic Reviews on Effectiveness, CRD’s Guidance for Carrying Out or Commissioning Reviews.” University of York, 2001. Online: <http://www.york.ac.uk/inst/crd/report4.htm>
- 5 Mulrow CD, Oxman AD. Cochrane Collaboration Handbook for Reviewers, Cochrane Library. Update Software; Issue 4, 1999. Online: <http://www.cochrane.org/cochrane/hbook.htm>
- 6 National Public Health Partnership (NPHP). A Planning Framework for Public Health Practice. NPHP, 2000. Online: <http://www.nphp.gov.au/planfrwk/index.htm>
- 7 Sheill, A. Deciding and Specifying an Intervention Portfolio. National Public Health Partnership, Melbourne, 2000.
- 8 National Health and Medical Research Council (NHMRC). A Guide to the Development, Implementation and Evaluation of Clinical Practice Guidelines. AusInfo, Canberra, 1999.
- 9 National Health and Medical Research Council (NHMRC) How to Use the Evidence: Assessment and Application of Scientific Evidence, p8. AusInfo, Canberra, 2000.
- 10 NHS Research and Development Centre for EBM. Levels of Evidence, 2001. Online: <http://163.1.96.10/docs/levels.html>