Report on developing the European review protocol and completed feasibility test reports by the partners, Getting Evidence into Practice project

09-08-2005

Räty Sanna*, Parry-Langdon Nina**, Drahonovska Hanna***, Laubli Loud, Marlène****, Bos Vivian*****. Aro Arja R.*

In collaboration with Getting Evidence into Practice Strand I Project team members and partners

* National Public Health Institute (KTL), Finland
** Health Promotion Division Research and Evaluation Branch, Wales, UK
*** National Institute of Public Health (NIPH), Czech Republic
**** Swiss Federal Office of Public Health, Switzerland
***** Netherlands Institute of Health Promotion (NIGZ), the Netherlands

The Getting Evidence into Practice Project, (Evidence Consortium, GEP, Grant agreement no 2003123 (790841)) is funded by the European Commission (DG Sanco/G). Neither the European Commission nor any person on its behalf is responsible for any use that may be made of the information contained therein.
Colophon

Project Getting Evidence into Practice Project  
(Evidence Consortium, GEP, European Commission Grant agreement no 2003123 (790841).

Report on developing the European review protocol and completed feasibility test reports by the partners, Getting Evidence into Practice project, 09-08-2005

Räty, Sanna*, Parry-Langdon, Nina**, Drahonovska, Hanna***, Laubli Loud, Marlène****, Bos, Vivian***** Arja R.*

In collaboration with Getting Evidence into Practice Strand I Project team members and partners

* National Public Health Institute (KTL), Finland  
** Health Promotion Division Research and Evaluation Branch, Wales  
*** National Institute of Public Health (NIPH), Czech Republic  
**** Swiss Federal Office of Public Health, Switzerland  
***** Netherlands Institute of Health Promotion (NIGZ), the Netherlands

September 2005

The central co-ordination of this project is carried out by  
Netherlands Institute of Health Promotion and Disease Prevention (NIGZ)  
P.O. Box 500  
3440 AM Woerden  
The Netherlands  
Visiting address: de Bleek 14, 3447 GV Woerden, the Netherlands  
Tel: +31 348 439853  
Fax: +31 348 437666  
Email: gettingevidence@nigz.nl  
website: www.nigz.nl/gettingevidence

Project staff:  
Dr. Gerard Molleman, project manager  
Jan Bouwens, project coordinator  
Hans Saan, senior advisor  
Marianne Smit, project assistant

© 2005 This publication has been produced with the financial support of the European Commission (DG Sanco/G) as part of the project Getting Evidence into Practice Project, (Evidence Consortium, GEP, Grant agreement no 2003123 (790841). Neither the European Commission nor any person on its behalf is responsible for any use that may be made of the information contained therein.
Content

1. Introduction ........................................................................................................................................... - 4 -

2. Development process............................................................................................................................ - 6 -
   2.1 Inventories of the existing review protocols and quality criteria.................................................... - 6 -
   2.2 Delphi process................................................................................................................................... - 6 -
   2.3 Feasibility test ................................................................................................................................. - 7 -

3. Further challenges .................................................................................................................................. - 7 -

Appendix 1: Feedback report on first Delphi round of the strand .......................................................... 9
Appendix 2. Feedback report on the second Delphi round of the strand I.............................................. 23
Appendix 3. Feedback report on the consensus meeting.......................................................................... 33
Appendix 4. Feedback report on the feasibility test............................................................................... 37

References
Abbreviations

EBM = Evidence Based Medicine
GEP = Getting Evidence into Practice project
HAD = Health Development Agency
HP = Health Promotion
HP/PH = Health Promotion and Public Health
IUHPE = International Union of Health Promotion and Education
NICE = National Institute of Clinical Excellence
QA-tool = Quality assessment tool
RCT = Randomized Control Trials
1. Introduction

Evidence-based health promotion requires conducting extensive cross-disciplinary literature search, selecting the most effective of the relevant programmes and applying rules of evidence and appraisal of study quality to determine the validity of the findings. It also requires attention to recognized health promotion concepts, sociocultural factors and organizational factors. (1) In addition, much of the practice in health promotion consists of looking outside the health promotion field (policy, social sciences etc.) for what works in health promotion. This implies a very broad understanding of how evidence is gathered. (2)

There are many systematic review initiatives internationally, several of which have focused specifically on health promotion and public health topics and interventions (3). However, most of these initiatives have used review protocols designed for medical and clinical studies (and applied to health promotion). Speller et al (4) have cautioned that considering health promotion with the tools used in Evidence Based Medicine (EBM) carries the risk that health promotion may be designated ‘not effective’ because it is assessed with inappropriate tools. In addition, the selection of studies is done on the basis of the quality of research only, not on the quality of the health promotion interventions. (4) In addition, most of the existing review protocols concentrate on finding evidence from research (randomized control trials (RCT) and from published documents but little emphasis has been given to finding evidence from practice and from expert opinion.

There has appeared a clear need to develop a review protocol suitable for health promotion and public health topics. Strand I, Review protocol is one of the strands within Getting Evidence into Practices project (GEP). It aims to gather existing review protocols and the quality criteria for reviewing process in the field of health promotion and public health, compare these, and find the key issues to be addressed in conducting reviews in HP/PH. Based on the analysis of the existing review protocols and quality criteria a consensus based review protocol is being developed.

The aim of this report is to describe the development process of the European review protocol for health promotion version 1.2 and the collaboration with the partners.
2. Development process

The development of the European Review protocol for health promotion version 1.2 was based on the extensive background work and active collaboration between the 18 GEP Strand I partner and project team member institutes.

The development process was divided into three main phases:
1. Collecting and analysing the existing review protocols and quality criteria
2. Carrying out Delphi rounds
3. Carrying out a feasibility test by the selected Strand I participant organisations.

2.1 Inventories of the existing review protocols and quality criteria for HP

The first phase started with an extensive literature search and survey by the Strand I participant organisations. More detailed information on this phase can be found in the following separate documents:
1. Inventory report on the existing review protocols (5)
2. Inventory report on the quality criteria in the existing review protocols of health promotion (6).

2.2 Delphi process

Based on the comprehensive analysis of the existing review protocols (5, 6) and notion of the key challenges when reviewing information for HP/PH topics the first draft proposal of the structure and content of the European review protocol for HP was developed. This version was sent to the Strand I project team members and partners to be commented and to enhance the broadening of the evidence base beyond research information in the protocol.

Altogether two electronic Delphi rounds (7) were carried out in the time period of 1\textsuperscript{st} of April - 20\textsuperscript{th} May 2005. The aim of the Delphi rounds was to reach a consensus on the content, structure and critical issues of the review protocol. More detailed information on the Delphi rounds can be found in the appendices 1-2.

After the two electronic Delphi rounds a consensus meeting was held at the Stockholm pre-conference on 31\textsuperscript{st} of May with the help of 14 experts present representing different areas. The main aim of the consensus meeting was to achieve a consensus on the broad definition of evidence. Feedback report on the consensus meeting can be found in Appendix 3.
2.3 Feasibility test

Based on the inventory of the existing review protocols and two Delphi rounds (Appendices 4A-B) the first version of the European Review protocol for health promotion, version 1.1 was developed. The aim was to test the usefulness of the review protocol and to further develop it with the help of outside experts. Altogether eight outside review experts from four countries were interviewed and their feedback was collected into a separate document (Appendix 4). Based on the analysis of the feasibility test results and including also feedback from the consensus meeting the latest version of the European review protocol for health promotion, version 1.2 was developed.

3 Further challenges

The development of the European Review protocol for health promotion has been a challenging and demanding task. The GEP project has touched a new area and took the first steps to bring the broad definition of evidence into the reviewing process. As the HP/PH research not yet fully takes into account the complex nature of health promotion, the research information may not be enough for the reviewing process. The development work of this new review protocol does not solely aim to support the idea of including expert opinions and grey literature into the reviewing processes, but it aims also to enhance the quality of health promotion research. As far as the research world cannot answer all needs of health promotion e.g. take into account the context, theory and complex nature of HP/PH interventions there is a need to collect information also outside research to make the review results more useful for real actors in the field.

The European review protocol for health promotion has tried to answer these challenges. However, there remain still some further challenges to make the protocol more suitable and user-friendly for the users. During the development work of the protocol consensus was achieved on the main structure and content of the protocol. Also the idea of including the broad definition of evidence into the reviewing process was strongly supported by the consensus meeting members (Appendix 4C). Also feedback from the outside experts was mainly positive with one exception. Changing the well-rooted clinical thinking may cause resistance and sceptism, and thus adaptation of new thinking requires both time and effort in the terms of training system.

The protocol has been found to be especially useful for HP/PH topics (Appendix 4D). Further development of the review protocol is still needed in the second phase of the GEP. During the first phase of the GEP no-consensus was achieved on which of the existing quality criteria for the reviewing process to recommend. In addition, the quality criteria as well as grading system for the non-research information and expert opinion are still lacking. There is a need to continue with these issues in the future. Development of the GEP grading system was strongly supported by the project
partners as well as outside experts (Appendix 4C-D). One approach would be to use the HDA's/ NICE's idea (as presented in the GEP protocol at the moment) as a starting point. Also the first steps have been taken in forming a smaller expert group to support this development work.

It is crucial to guarantee the continuity of the GEP project and the development work of the review protocol. Further challenges are to develop more user-friendly formats (e.g. shorter version, electronic version with an index) and to support the dissemination and training the use of the protocol. Especially in the eastern part of the Europe, where review protocols are rarely used, training on how to review and where to find evidence is very welcome. When ready the European review protocol could enable the comparison of the results and collaboration between countries when developing evidence-based HP/PH interventions. Shared norms and standards would support the evidence base in all arenas of health promotion: practice, policy and research.
Appendix 1. FEEDBACK REPORT ON FIRST DELPHI ROUND OF THE STRAND I

FEEDBACK REPORT ON FIRST DELPHI ROUND OF THE STRAND I

Räty S. and Aro A.R.
3th of May 2005

Getting Evidence into Practice, GEP project. EC project no.20031123
(Evidence Consortium 790841)
1. Background

The first Delphi form was developed on the basis of the Delphi exercise held in Woerden in the Strand I workshop and on the analysis of the existing review protocols. The Delphi form was sent to all project team members (N=8) and partner organisations (N=10) of the strand I. Altogether 4 organisations answered before the deadline (15th of April) and 5 organisations asked more time. One reminder was sent before the first deadline, and after the deadline the non-respondents were contacted and informed on the importance of the participation. Altogether 10 organisations participated in the Strand I 1st Delphi round. The last answers were received on 27th of April 2005.

2. Analysis of the results

The first Delphi form included the first draft proposal for the European review protocol and the participants were asked to comment the draft paragraphs and familiarize with the given www-links when relevant. The analysis of the results is descriptive.

Delphi 1 round included 22 questions on the context of the proposed chapters of the review protocol. Table 1 in Appendix 1 summarizes the main questions and demonstrates on which questions the consensus was achieved and which of them require the 2nd Delphi round. The Delphi questions were open questions where the respondents were asked give comments and impression, additions or changes.

The consensus was defined: all respondents (n=10) agreed on the content of the chapters, or they did not comment at all, or they did not have anything to add (e.g. they answered “no” to the question “do you want to add/change something”.

No consensus was defined: one or more respondents explicitly opposed the proposed content of the chapter.

All comments and answers have been analyzed. The technical suggestions (language e.g. corrections) have been taken into account (however, a professional language check will be done at the later stages). In addition, the suggestion to include, exclude, change or improve the structure, headings, definitions and contents a have been summarized (Table 1 in Annex 1).

3 Results

3.1 General comments on the structure of the review protocol

Generally the structure of the review protocol was seen good and logical. The following slight changes were suggested:

1. Change the name of the chapter 4 Inclusion criteria ➔ ‘Selection criteria’ or ‘inclusion and exclusion criteria’ (by three respondents, n=3)
2. Re-number the sub-chapter of the Planning phase (n=2)
3. First section: Introduction not instruction
4. Own chapters for data extraction and data synthesis (n=1)
5. Own main chapter for getting evidence into practice (n=2)
New order for chapters 3 to 7 (see also Appendix 2, the structure of the protocol):

<table>
<thead>
<tr>
<th>Delphi 1</th>
<th>Suggestion 1</th>
<th>Suggestion 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Channels of finding information</td>
<td>5. Inclusion and exclusion criteria / Selection criteria</td>
<td>5. Inclusion and exclusion criteria / Selection criteria</td>
</tr>
<tr>
<td>8. Data extraction and synthesis</td>
<td>8. Data extraction</td>
<td>8. Data extraction</td>
</tr>
</tbody>
</table>

In addition three respondents suggested to add some information on the working process and some background information on the GEP project. It was also suggested to include aim, short introduction to the review protocols in health promotion, differences with these protocols, development process, and instruction on how to use this protocol at the introduction part of the protocol.

In addition, according to the respondents the protocol should include:
1. A chapter or description on how to transfer reviews to recommendations (grading the evidence) (n=2)
2. Definition of the terms ± Glossary of main terms (n=4)
3. More practical examples (n=3)
4. Step by step instructions (n=2)
5. References with clear distinction what is based on the consensus of the GEP consortium and what on the existing review protocols

3.2 Comments to the Delphi questions

3.2.1 Comments on question 1.

Comments to the introduction chapter, would you like to add/change/remove something?

Generally the introduction chapter was seen to be good and informative (n=5). Most of the respondents (n=5) emphasized the importance and highlighted the added value of this protocol. In addition, it was strongly recommended to give some background information on the working process and why the development of the European review protocol was essential.

Suggested changes:
1. Specify more clearly **who are the potential users** of the review protocol ± “It should be aimed also others than researchers or what is the definition of the researchers”
2. Specify the **overall purpose of the protocol;** what for and why, is it for doing a systematic review or perhaps narrative review, or what?
3. Highlight **the new elements of this protocol** (what is new)
4. Add some information on the **working process** and some background information on the GEP project: include aim, short introduction to the review protocols in health promotion, differences between these protocols, development process, and instruction on how to use this protocol.

5. Specify the **main steps of the protocol** and the steps to be taken inside each main steps e.g. a follow chart of the main steps.

### 3.2.2 Comments on question 2.

<table>
<thead>
<tr>
<th>Comments on the resources chapter, would you like to add / change/ remove something?</th>
</tr>
</thead>
</table>

Only slight modifications were suggested to the resource chapter. The respondents (n=4) suggested to complete the timetable list so that it covers all the main phases of the reviewing process. In addition, two of the respondents suggested changing the order of the chapters so that we would start the planning phase from the reviewers, advisory group, and other resources.

**Suggested changes:**
1. Complete the **list of timetable**:
   - Specify training for whom and for what? Can the training be instruction meeting/exploratory meeting?
   - The timetable should include also: reading phase, translations, analyze and classification, dissemination, implementation and conclusion phase.

2. **Change the order of the headings**: reviewers, advisory group, other resources (see chapter 3.1)

3. Define more clearly: who does the search: librarian, statistician etc. Define more clearly the persons and skill required for the review process.

### 3.2.3 Comments to question 3

<table>
<thead>
<tr>
<th>Comments on the advisory group chapter, would you like to add/ change/remove something?</th>
</tr>
</thead>
</table>

Most of the respondents (n=5) emphasized the importance of specifying more clearly the number of the persons to be included to the advisory group and also explain a bit further the roles (job descriptions of the participants of the advisory group). Four respondents did not have any comments or opinion on the chapter.

**Suggested changes:**
1. Include **definition for the advisory group**

2. Clarify the forming of advisory group: with a leadership/chair person + notification of the administrative support

3. Notice the **HP/PH expertise**

4. **Specify the number of people in the advisory group**: what is a small group and specify the number of the group: small group performing the review 2-4 persons, including supervision and expert group of 5 to 8 persons?

5. Specify the **role of the persons** of the advisory group

6. Specify also how to identify experts/ how to locate them
3.2.4 Comments to question 4

Comments on the necessary number of the reviewers. Would you like to add/change/remove something on this chapter?

It was agreed that it's good to have a coordinator who takes the responsibility of the reviewing process. However, most of the respondents (n=5) wanted that the number of the reviewers is more explicitly said in the protocol. In addition, the specification of the reviewer skills needed was requested by four respondents. It was stressed that the skills and expertise should more clearly reflect HP/PH topics (now the list of experts relies far too much still on clinical experts). Three respondents suggested more expert areas/research designs to be included in the list of potential reviewers.

Suggested changes:
1. Change the order of the chapters (see the suggestion in the introductory chapter)
2. Specify how many reviewers are recommended: more clearly the number of recommended reviewers e.g. minimum 2, a group of 2-4 person.
3. Specify all the persons who are needed to make review: reviewers, assistants, administrative people, librarian, statistician, and skills and competences for all of them
4. Complete the list of expertise with skills needed (review methodology, skills in systematic literature search, other skills according to the topic to be reviewed which differ from review to review).
   † Specify what skills are needed: the list of expertise
   ‡ Make the list more relevant to the HP/PH (not as clinical as it is at the moment):
   - Determinants of health
   - Not medical statistics † biomedical etc. more HP oriented
   - Health and social measurement
   - Medical and social statistics
   - Social epidemiology or social and educational sciences
   - Qualitative and quantitative research
   - (Clinical) and social epidemiology
   - Political scientists/experts

3.2.5 Comments to question 5

Comments on forming the review question chapter, would you like to add/change/remove something?

Two of the respondents did not want to add anything to the chapter of forming the review question. However, most of the respondents wanted some clarification for the chapter.

Suggested changes:
1. Specify more clearly who makes the decision on the review question? The project manager or the advisory group?
2. Some considerations of how to select the questions could be useful: Ideas for items of possible reviews could come from many places: Practitioners, politicians, researchers, health care planners.
3. Examples of good reviewing questions could be useful to illustrate the text.
4. What to take into account in formulating a review question”:
   • “IDENTIFYING THE REVIEW QUESTIONS” should be linked to the reflection of why we want this review of evidence and what do we intend to do with the results.
• Contexts, contextual factors and management / processes of the interventions / policy development processes etc.

• CONDITIONS - i.e. some reference to the different contexts of the interventions reviewed / or the contexts in which this evidence was developed. The question of What works is always limited if we don't get some understanding of What works for whom and under what conditions

5. I do not agree on the text that is included in the paragraph 'how contextual and…'. The question does not set out the rationale and does not explain why the questions are important. Instead: the research question defines the subject area and the aim of the review. The rationale and relevance should come from knowledge on state of the art of the field. The question does also not have to contain a conceptual discussion etc. It should be based on the knowledge on relevancy, theory etc. etc.

6. Make the text more relevant to the broadening the evidence base: more information on the other documents than research and modify the text reflecting this (e.g. remove the clinical terms and add examples on the HP/PH)!

3.2.6 Comments to the question 6

Comments on classes of information sources chapter, would you like to add/change/remove something?

Four respondents had nothing to add to the information sources chapter. One respondent suggested adding evaluation studies to this category and four suggested the following changes:

Suggested changes:

1. Include or exclude lay people
   • Where a health promotion intervention is a secondary intervention (e.g. CHD) ‘knowledgeable patients’ are known to be very informed and informing. This is particularly relevant for health promotion and disability issues (including chronic illness).
   • Could allow exploring the acceptability of interventions, the perception of interventions and the place of intervention compared with other interventions in different topics.
   • Lay people as sources of information should not to be included in the review, but of course they can be included in the process of defining the question and in the advisory board and so on.

2. Include or exclude experts opinions
   • Could be a source of information about effectiveness or efficiency
   • What about the subjectivity? The method?
   • I feel that expert opinion is as rigorous as RCT! Both sources provide different and useful information, and that has to be explained and developed!
   • Explain how to involve experts? How many opinions are needed? Specify also how to identify experts/ how to locate them

3. Include or exclude non-research
   • Explain more: What is the value to add of such information as expert opinion, non-research etc.?
   • Include information on evaluation studies

4. Discuss which study designs are appropriate for which questions included in this review protocol

5. What about review up-dating? Should we add some explanations about how to up-date or adapt existing reviews

6. Include some examples
3.2.7 Comments on question 7

Comments on the search strategy and language question would you like to add/change/remove something?

Only couple of respondents (n=3) commented the search strategy. Suggestion was given to include also some general instructions for developing a search strategy for non-research literature and not only for research literature. Almost all the respondents (n=7) regarded the language question problematic. There were suspicious comments against including different languages to the reviews because of the translation cost and possible poor quality of the non-English research results. Two of the respondents did not have anything to add to the chapters.

Suggested changes:
1. The language question was seen problematic:
   - Not realistic to include many languages!
   - Translation is expensive, impossible where budgets are tight and where the project is fairly small.
   - Include research reviewers in the team who have the necessary linguistic skills.
   - The up-dating process should be developed in this protocol.
2. Some general instructions for developing a search strategy also for unpublished literature
3. Include definitions

3.2.8 Comments to question 8

List here databases that you would like to be included (name of the database and www-address). Would you like to include also national database? If yes list the national databases and www-addresses. Any other comments?

Five respondents suggested including also national databases. Two respondents could not list them in a given timeframe and one informed only that national databases should be included in the protocol. It was also reminded (n=1) that the databases and their addresses change frequently, thus, such a list needs to be updated continuously. Listed national databases were:
- Central European Journal of Public Health - is in English, edited in Prague
- Political science databases
- Banque de données en santé publique (Public health databases): http://www.bdsp.tm.fr/Base/QbeA.asp
- Banque de données en Sciences, technologies et médecine PASCAL (Sciences, technology and medicine database): http://www.inist.fr/PRODUITS/pascal.php
3.2.9 Comments on question 9

List here journals that you would like to be included in the list (name of the journal). Would you like to include also national journals? If yes, list the national journals.

Five respondents suggested including more journals in the list. The hand searching was seen as a way to identify not only trials but also other published studies. The journals suggested were:

- Critical Public Health [http://www.tandf.co.uk/journals/authors/ccphauth.asp](http://www.tandf.co.uk/journals/authors/ccphauth.asp)
- American Journal of Community Psychology,
- Health and Social Care in the Community,
- Canadian Review of Social Policy.
- Provide also example of subject matter journals - e.g. AIDS, something on sexual health etc. etc
- Add national journals, for exemple Revue française d'épidémiologie et de santé publique, Santé de l'Homme (revue de l'INPES)
- Evaluation journals

One participant did not answer the question and one found it not feasible to suggest hand searching at all: “It is not realistic to ask reviewers to hand search a wide range of journals - and for how many years back?” It was suggested that hand search can be an option if it is not possible to locate the evidence from the electronic literature databases, but not something to do always.

3.2.10 Comments on question 10

List here research register(s) (HP/PH) that you would like to be included in the list (name of the registers and www-addresses). Would you like to include also national registers? If yes, list the national research registers and/or link(s) to these registers (give also www-addresses).

Only two respondents listed national research register:

- Danish Data Archive [http://www.dda.dk](http://www.dda.dk)

One respondent suggested that national registers could be included, but could also be included in a national version of the protocol. One respondent regarded the trials registers not so relevant in the protocol and it was doubted that there are no registers for HP/PH research making the statement
'research registers on the topic area are good to be searched' less valid. Five respondents did not have any opinion of the question.

3.2.11 Comments on question 11

List here the sources of grey literature that you would like to be included in the list (name of the sources and possible www-addresses). Would you like to include also national sources in the list? If yes, list the national sources and give also the www-addresses

One respondent did not want to include any national sources here (can be included in national version of the protocol) and one emphasized that it is extremely difficult and irrelevant to search grey literature for each subject! Four respondents did not have any opinion of the question.

Three respondents mentioned some national databases including grey literature:
- French databases include grey literature: BDSP, Pascal, Francis …
- Danish Data Archive http://www.dda.dk lists also grey literature

3.2.12 Comments on question 12

List here the search engines that you would like to be included in the list (name of the search engines and the www-addresses )?

Eight respondents did not have any opinion of the question. One respondent pointed out that it is not very useful to search through www-pages: “waste of time to use these internet sources - go directly to the science, otherwise you'll get a lot of rubbish to get rid of”

3.2.13 Comments on question 13

If expert opinion is to be included as a source of information should we explain more how to find/collect this information?

Three respondents answered only ‘no’ or they did not have opinion of the question. One respondent was strongly against including expert opinion in the systematic reviews: “absolutely not, just adds special people’s values; a systematic review should not include such horrible things!” Five respondents supported including experts as sources of information. However, it was stressed that more information is needed from where and how to find this information (provide some guidance). It was suggested that:
- (e.g. University home pages list areas of expertise - it is also useful to search literature by [a] topic for a listing of authors for potential experts)
- ‘Health Promotion Source’ (IUHPE) could be sources.
- In addition, it was stressed that this channel (experts) can also be used to collect evidence from practice. More information needs to be given how information from practice can be obtained and used for the review.
3.2.14 Comments on question 14

Do we need also some other selection criteria? If yes, what?
Do we need selection criteria for other information than research, e.g. for policy documents? On what basis do we include policy and other documents in the review?

Only one respondent did not want to comment the questions. Two respondents would like to include discussion on equality and empowerment as well as theory base in the inclusion criteria but they pointed out that it would however exclude maybe too many papers. Two respondents suggested additional inclusion criteria: sometimes it is relevant to restrict the inclusion for time period; according to the review question (e.g. all studies up to 5 years ago etc.). In addition, two respondents expressed their opinion on selection criteria for policy documents. It was stated that selection criteria for policy documents could refer to the topic area (e.g. if the evidence is about smoking interventions) and/or the strategic area (e.g. young peoples' health), and/or a particular population (e.g. ethnic minority). These criteria might cover other documents (short reports, web-based information, etc.). Inclusion criteria could be clear information of the origin of the paper, time, purpose and target group. It was stressed that policy papers will not always be relevant to be included all depends on the question.

Changes suggested:
1. Call this section 'Selection criteria'
2. Definitions for the inclusion criteria and explanation on where are we now in the process of reviewing.
3. An example of inclusion criteria
4. Include time period to the inclusion criteria
5. Including also other sources (than only study dg) more visibly to the components of inclusion criteria?
6. Combine selection criteria and quality criteria chapters

3.2.15 Comments on question 15

Is this a good way to present the quality criteria of quantitative studies or should we recommend e.g. only some of the existing quality criteria? Please check the links provided. Would you recommend some of these? And if yes, why?

It was seen (n=4) that the list of links provided a comprehensive list of quality checklists (two of them had no opinion and two regarded this form of presenting the information good (sufficient information can be obtained thought the links). However, it was stressed that for an experienced reviewer this is not a problem but for a health promotion specialist who has minimal support the information in the links can be daunting. Thus, it might be useful to provide a core set (must do list) of criteria to determine the 'minimum quality threshold'. In addition, three respondents stressed that it could be useful to give a brief description of the content of each link or to be more precise (for example: give an example of relevant questions) or suggest only some to be used. Two respondents wondered if it is possible to use somehow the strand II QA-tool for this. One respondent pointed out that it's important to be aware that quality criteria includes different value judgments based on hidden theory of science preference and that they all have biases that effects the endpoint value of a systematic reviews. As a reader of a systematic review it's important to check these value judgments and how they were used. One respondent stressed that it should also mention of quality standards for the Meta evaluation of evaluation studies e.g. the American
Evaluation Society’s standards or even the Swiss Evaluation Society’s standards (there may even be European Union evaluation standards).

Changes suggested:
1. Create own quality assurance list and give also list of the existing ones
2. Brief description of the content of each link
3. Possible links to the strand II
4. Include quality standards for the meta evaluation of evaluation studies
5. Specify how strict these criteria should be applied
6. Specify when information is seen as evidence and how this is done? How many articles and of what quality can be seen as providing evidence?
7. Include steps to be taken

3.2.16 Comments on question 16

Is this a good way to present the quality criteria of qualitative studies or should we recommend e.g. only some of the existing quality criteria? Please check the links provided. Would you recommend some of these? And if yes, why?

Four respondents found this way of presenting the quality criteria of qualitative research good: ‘It is good to have several sources; I agree with suggestion above, well it’s a good example list’. One suggested including also quality standards for the Meta evaluation of evaluation studies (e.g. the American Evaluation Society’s standards or even the Swiss Evaluation Society’s standards / European standard). Two respondents suggested to include description of the links, one recommend to make a synthesis of these tools/checklists (GEP own quality list) and one of crating an own quality checklist besides the recommended links. One respondent found the Quality in qualitative evaluation: a framework for assessing research evidence” http://www.policyhub.gov.uk/docs/a_quality_framework.pdf the most interesting and more relevant for this purpose providing good and comprehensive explanation about qualitative studies design and quality criteria.

Changes suggested:
1. Decide which one of the provided links to recommend or make a synthesis of them
2. Brief description of the content of each link
3. Include quality standards for the meta evaluation of evaluation studies

3.17 Comments on question 17

How to assess the quality of other information than research?

Five respondents agreed on the proposed chapter. Three respondents found the HDA suggestion on general quality criteria is sufficient for assessing the quality of other information than research. Two of the respondents pointed out the importance of evaluation studies/reports: One should also mention of quality standards for the (Meta) evaluation of evaluation studies. One respondent had a feeling that other information has to use the same quality criteria as research and published studies. Two of the respondents did not have a clear opinion of the question.
Changes suggested:
1. Develop own quality criteria
2. Include steps to be taken.
3. Formulate inclusion and exclusion criteria on quality and apply these to the studies
4. Example could be used to illustrate the process

3.18 Comments on question 18

Is this a good way to present the data extraction chapter?
Please check the links provided. Would you recommend some of these? And if yes, why? Do we need some further information on data extraction?

Three respondents regarded that the provided text is sufficient. Four had no opinion of the chapter and of which two did not check the links. A suggestion was given to organize an own Delphi round for data extraction forms. One of the respondents highlighted the Community guide’s data extraction form as most suitable for this purpose giving good explanation with examples. One stressed that ideally, GEP should summarize / provide its suggested data abstraction form. Another way would be to give good explanations why those listed were selected, and whether they can be recommended on the basis covering community based/settings/social systems interventions. The others regarded that the proposed content was either too general (n=1) or too specific (n=5). It was suggested to include some examples of formats for data extraction forms in the text.

Changes suggested:
1. Example for both methods (paper or electronic data collection) and list of advantages and disadvantages.
2. Summarize / provide own data abstraction form and/or why those listed were selected and whether they can be recommended and why
3. Specify who should do the abstraction and the assessment of information quality? The reviewer? The "commissioner"? Independent judges? A worked example might be useful to illustrate the process.
4. Own main chapter (not together with synthesis)

3.19 Comments on question 19

Is this a good way to present data synthesis chapter or should we give more practical instructions/examples? Do we need some further information on data synthesis?

Four respondents stressed that practical examples on, what should be done and how, could make the chapter more precise. It was suggested to provide references to examples of data synthesis and give practical examples. Also explaining the purpose of making the data synthesis was stressed (WHAT should be done and HOW this should be done).

One respondent did not have any opinion of the question and one referred not to have enough expertise to be able to answer the question. Three respondents regarded that the proposed chapter gave needed information and more detailed instructions were not needed.
Changes suggested:
1. Specify the steps to be taken
2. Mention and explain why to do it (useful to explain briefly why synthesizing studies with variable methods is far from simple)
3. Describe explicitly what should be done by using examples
4. Own chapter

3.20 Comments on question 20

Is this a good way to present (e.g. rely on the existing information) combining different sources of information? Or would you suggest some other format for this chapter? If yes, what kind? Do we need some further information / clarification?

Four respondents regarded the proposed way a useful way of presenting the information (there is no point is duplicating work that is already well done). However, a bit more explanation on why this phase is far from simple could be added to the chapter. Three respondents did not have any opinion of the question and one referred not to have enough expertise to be able to answer the question. One respondent wanted to have more information and would suggest references in the text for more information and one regarded the text too complicated and specific: ‘This part is written for researchers and an explanation and examples could be very useful.

Changes suggested:
1. Changes to the content of the chapter:
   a. Triangulation should be discussed at the beginning of the chapter
   b. Move away or delete the Bayesian hierarchical modeling from statistical combination
2. Complete the information:
   a. Include references on statistical combination: people will need more information if they want to do this and I do not think it is necessary to include any further information on this topic here.
   b. Explain also in a bit more detail consensus methods and the pros and cons of this in comparison to triangulation.

3.21 Comments on question 21

Is this a good way to present the interpretation of the results chapter? Do we need some further information?

Three respondents regarded that this is a good way to present the interpretation of the results chapter. One respondent pointed out that developing the ‘knowledge base’ in reviewing is essential - it might be useful to add a paragraph that suggests possible mechanisms to ensure that the learning obtained through the process of review is recorded and disseminated (e.g. hold a feedback session for other reviewers/potential reviewers) Note: disseminating the learning is not the same as reporting and disseminating the review report.

Three respondents regarded the chapter too complicated. It was suggested to concentrate to the most relevant question such as:

- What works
- for whom
- under which circumstances
- and how well

One respondent suggested using the layout of the discussion section of scientific papers: first short summary of the results, than discussion of methodological pitfalls and limitations, than comparison
with other research on same topic, than explanation of results in light of the research questions, than implications for practice, policy (applicability) than conclusion. The whole chapter should be written from the perspective of the research question.

Changes suggested:
1. Modification of the text; what works, to whom, contexts
2. Stress the importance of learning
3. More specific instruction on the topics/issues to be handled in the conclusion chapter

3.22 Comments on question 22

<table>
<thead>
<tr>
<th>Is this a good way to present reporting and disseminating and how to get evidence into practice? Do we need some further information?</th>
</tr>
</thead>
</table>

Five respondents regarded that this chapter is a good way to guide on writing the review and on how to disseminate. One respondent proposed small changes to the given structure of the review report. Four respondents thought that getting evidence into practice is very difficult and might require a protocol of its own - or at least its own chapter in this review protocol. It was stressed that the review protocol can only provide advice rather than instructions on this topic. One respondent pointed out that getting evidence into practice is a big job and properly done as time consuming as the review itself. It was suggested that maybe it should be done by others than those that did the review, people with other skills in communication and perhaps jointly and interactive with practitioners. In addition, it was stressed that a review is not recommendations. Some relevant questions were posed: “How can a review be useful for public health managers and public health actors? We don't know how the public health managers could use and transfer results of the review into their practice, into interventions”. Some examples could be a good way to clarify this.

Suggested changes:
1. Own chapter for getting evidence into practice
2. Complete the list of proposed headings of the ‘review report’:
   1. Cover Sheet
   2. Executive Summary
   3. Background
   4. Objectives / review question
   5. Methods
   6. Recommendations and Conclusions (including answer to review question)
   7. References
   8. Appendices
   9. Tables
   10. Acknowledgements
3. Include some practical examples
Appendix 2. FEEDBACK REPORT ON THE SECOND DELPHI ROUND OF THE STRAND I

FEEDBACK REPORT ON THE SECOND DELPHI ROUND OF THE STRAND I

Räty S. and Aro A.R
23th of May 2005

Getting Evidence into Practice, GEP project. EC project no.20031123
(Evidence Consortium 790841)
1. Background

The Delphi form was sent to the participants of the first Delphi round (n=10) and to one additional expert (who was not able to send the answers in time to the first Delphi round). Altogether five organizations answered before the deadline (13th of May) and 3 organizations asked for more time. One reminder was sent before the first deadline, and after the deadline the non-respondents were contacted and informed on the importance of the participation. Altogether 10 organizations participated in the Strand I 2nd Delphi round. The last answers were received on 20th of May 2005.

2. Analysis of the results

The second Delphi form concentrated on the issues where consensus was not received as well as to the issues which required further development. The analysis of the results was descriptive.

The 2nd Delphi round included seven questions. In addition comments were requested for the two additional topics: the proposed structure of the review protocol and to the content of the introduction chapter. The Delphi questions were open questions where the respondents were asked to give their comments and impressions, additions or changes.

The consensus was defined: all respondents (n=10) agreed on the topics/issues, or they did not comment at all, or they did not have anything to add.

No consensus was defined: one or more respondents explicitly opposed the proposed topic/issue.

Based on the majority: at this stage in the issues on no consensus the contents accepted by the majority are kept in.

3. Results

3.1 Comments on the changes made based on the 1st Delphi round

All participants (n=10) agreed on the proposed changes made based on the first Delphi round.

‡ The proposed changes based on the first Delphi round were accepted and the corrections have been made to the review protocol. (Consensus achieved on the content of the chapters and changes proposed)

3.2. Opinions on including an example of the review question in the review protocol

In the first Delphi round it was request to include an example of the review question. In this second round we suggested to include an example on the forming of the review question from the handbook of Systematic Review of Health Promotion Interventions (Jakson, 2005 http://www.vichealth.vic.gov.au/cochrane/training/Unit%20Five%20to%20print.pdf)

Eight of the ten respondents either agreed on including this as an example or suggested to use it as a starting point. One respondent did not understand the question and one did not want to include the proposed example because it lacked the reference to “under what conditions” as well as reference to contextual issues. One of the respondents pointed out that it is a bit dangerous to select only one example: ‘I would prefer to see a longer list and one that provides examples from across Europe.’ The respondent pointed out also that the GEP project should also provide guidance.
on how to set review questions as this is one of the key stages for getting an informed view about what sources of evidence would be appropriate.

† Based on the majority the proposed example will be included in the review protocol and it will be used as a starting point (e.g. modified) as an example of the review question and how to develop a review question (no consensus achieved).

### 3.3 Consensus on the definition of the evidence

Nine of the ten respondents agreed on the definition of the evidence to cover also non-research information and expert opinions besides the research information (RCT etc.) Even if supporting the main idea of the evidence five respondents wanted still to include some clarification on what circumstances the ‘lower level of evidence’ would be appropriate. It was suggested to make it clearer that the evidence from different sources can be used for different things. It was stressed that for example in order to assess the effectiveness of an intervention (what works and how well) the research literature would give the most appropriate answers. If information from research literature could not be found searching for grey literature could be recommended (knowing that the strength of the evidence presumably would be weaker). It was stressed that when it comes to knowledge of for whom the intervention works for and under which circumstances the research literature will often not be able to help. Here the grey literature and expert opinions can be helpful.

In addition, it was stressed the importance to describe how the assessment of the quality of the expert opinion will be done and develop a list of inclusion and exclusion criteria (selection criteria e.g. include if the ‘source’ of expertise is clear i.e. the experts are recognized as such, and have been cited; Secondary sources, internet: include if the ‘source’ of evidence is properly referenced e.g. to known sites of expertise; if the secondary source is cited elsewhere). One respondent would like to get (even if supporting the broad definition) more precise definition on how far the broadening can go in the GEP.

Only one of the respondents did not fully support the idea of including experts and non-research as a source: ‘I disagree with the idea that expert opinion/ or their views should be included in systematic reviews.’ However the respondent pointed out that the inclusion of practitioners’ views and experience is important in the overall health promotion work and not the least are they key players getting evidence to work in practice but they should not be used as source of information’.

† Based on the majority the European review protocol will rely on the GEP-project definition of evidence covering both research information and information derived from expert opinions and practice. (No consensus achieved)

† Clarification of the text will be done according to the feedback

† Attention will be paid on how to analyze the quality of the expert opinion and on how to grade the evidence (3rd Delphi Round in Stockholm)

### 3.4 How would you develop a search strategy for non-research literature?

Nine of the ten respondents supported the idea to search also non-research literature. Only one respondent did not want to recommend to search grey literature because of the experience from searching grey literature in the field of work health promotion and promoting physical activity have showed that it is a lot of work with very little valuable results.

All the other respondents gave some ideas on where to search, but generally did not specify how. One respondent pointed out that there are no instructions available of how to develop a search.
strategy for non-research literature, probably because it is difficult to develop a standard search strategy, as the search strategy is totally dependent on the research question. Therefore the respondent suggested to rely on the assistance of an information specialist in identifying non-research literature: ‘An information specialist will often be able to identify much non-research literature form his/her own country – but will probably have serious problems, if non-research literature outside her own country has to be identified as well. One way to approach this problem could be to collaborate making the review between more countries with experiences in a certain intervention – and bringing the non-research literature from the collaborating countries together – and not including systematic searches of non-research literature from other countries.’

In addition, it was suggested that the search strategy of non-research should

1) Rely on key words building on: setting, population and intervention specified in the research question would be the main approach
2) Searched for/agree sources of non-research literature with librarian (or other relevant knowledgeable source)
3) Include (e.g. define) own selection criteria [of non-research literature] relevant to the topic area.
4) Include and apply quality criteria systematically
5) Include discussion of findings with other reviewer/s.

In addition, it was emphasized that if the non-research information has been published and searched from electronic databases then similar rules apply as for research literature. However, if it is unpublished then there are a variety of methods including internet searches and setting networks of experts and putting out calls for literature against specific criteria.

Also a couple of good sources on how to search grey literature were proposed:

2) http://www.paramedic.com/research/researchfeature_view.asp?cid=120

Two of the respondents did not like the term non-research literature and, in addition, they had found controversy in the terms (non-research and unpublished studies). It was suggested to change the term non-research to the non-peer reviewed research.

‡ Based on the majority searching for non-research information will be included in the review protocol (No consensus achieved)
‡ Modification and adding of the search strategy for non-research will be done according to the given feedback.
‡ The coherence of terms will be checked and modified

3.5 About the creation of the quality assessment tool/checklist for the European review protocol

Almost all respondents (n=8) proposed ‘the must to do list’ when assessing the quality of the information. The proposed items have been collected and the recommended must to do list is under construction. In addition, some ideas on where to start (e.g. contact information) were given.

‡ Use the must to do list as a starting point in the 3rd Delphi round at the Stockholm pre-conference
‡ Familiarize with the given documents and create the draft proposal for the grading system for further develop at the Stockholm pre-conference
3.6 On what basis to start the development process of the quality checklist for the GEP project

Most of the participants (n=9) suggested that the development of the possible GEP quality checklist should start on the basis of the existing checklists (for research literature). One respondent pointed out that we have to carefully think what is the added value of producing something over and above what already exists. In addition, two respondents suggested concentrating on the policy documents and expert opinion level information where there are no quality checklists available so far. Some suggestion for the further information on this topic was also given.

One of the respondents suggested relying on the strand II protocol.

‡ GEP project should propose some of the already existing quality assessment tools
‡ At the Stockholm pre-conference we have to agree on:
   - Which quality assessment tool to recommend (of the already existing ones)
   - If we start the development of our own checklist: what is the added value?
   OR
   - Should we concentrate on creating the quality assessment only in the area where the information is missing? (No consensus achieved)
‡ In addition to the quality assessment of the review material as such the quality of the interventions included in the review material needs to be judged by using relevant criteria (the possibilities to develop a must to do list of the GEP, Strand 2 QA-tool, for the needs of review will be further discussed).

3.7 About including practical examples in the review protocol

Most of the (n=6) respondents were unable to give any practical examples. Three respondents recommended using some of the already existing review as examples. Two respondents stressed that examples are good but we cannot make this document too long. Including examples can be done electronically. However, it was stressed that whilst we might not want to go down the route of a standard European form for each of these issues – we do need to highlight a core framework and minimum standards for how to do things. One respondent did not answer the question.

‡ Highlighting some of the given examples as a minimum standard (no consensus achieved)

3.8 Comments on the proposed structure

The structure of the protocol was accepted by all respondents.

‡ Consensus on the structure of the protocol was achieved

3.8 Comments on the introduction chapter

Generally the introduction chapter was seen good (n= 6). However, it was suggested to make it more specific why we think this protocol adds value to any of the other nationally developed protocols. In addition, some changes in the order of the chapter were suggested. One respondent stressed that the usability of the document should be kept in mind and thus to keep the introduction as short and concrete as possible.

‡ Added value will be said more clearly
‡ Some modification of the chapter will be done
<table>
<thead>
<tr>
<th>Delphi question</th>
<th>Consensus</th>
<th>Suggested changes, accepted</th>
<th>Topics/issues requiring still consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>General comments on the structure</td>
<td></td>
<td>New suggestions for the headings/ordering</td>
<td></td>
</tr>
<tr>
<td>Q1. Comments to the introduction chapter</td>
<td></td>
<td>Clarify:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• The potential users of the review protocol</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• The overall purpose of the protocol</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• The new elements of this protocol</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Include background information (on review protocols, GEP project, development process)</td>
<td></td>
</tr>
<tr>
<td>Q2. Comments on the resources chapter</td>
<td>The chapter content OK (only some slight modification)</td>
<td>Change the order of the headings</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complete the list of timetable, persons and skill required for the review process</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>✪ will be done according to feedback</td>
<td></td>
</tr>
<tr>
<td>Q3. Comments on the advisory group chapter</td>
<td>Forming of advisory group and including experts, practitioners with different background to the group</td>
<td>Clarify the number of people in the advisory group and the role of the persons of the advisory group: ‘Number of the people in the advisory group depends on the subject area or expertise needed...’</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Notice also HP/PH experts</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>✪ will be included according to feedback</td>
<td></td>
</tr>
<tr>
<td>Q4. Comments on the number of the reviewers</td>
<td>One of the reviewer is the coordinator of the work</td>
<td>Specify how many reviewers are recommended: ‘Number of the reviewers depends on the subject area or expertise needed. Basic rule is that not alone e.g. minimum of two reviewers.’</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Include examples of persons who are needed to make review and expertise with skills needed (HP/PH perspective)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>✪ will be included according to feedback</td>
<td></td>
</tr>
<tr>
<td>Q5. Comments on forming the review question chapter</td>
<td>Main components of question: 1. population 2. interventions 3. outcomes relevant to the objectives of the review 4. designs of primary studies</td>
<td>Include examples of the ideas where the possible items of the reviews could come from</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Including also other sources (than only study dg) more visibly to the components of questions</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>✪ will be done according to feedback</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Examples of good reviewing questions</td>
<td></td>
</tr>
</tbody>
</table>
Table 1. Summary on the 1st Delphi round

| Q6. Comments on classes of information sources chapter | Include discussion on how the review is updated (one sentence to the introduction part) † will be included
| NOTE: The GEP project builds on a broad consensus regarding the definition of 'evidence'. Evidence is not restricted to the results of "hard" scientific research: Besides written information, attention should be given to expert knowledge and secondary sources including the Internet | Include discussion on which study designs are appropriate for which questions † will be included
| | Some general instructions for developing a search strategy also for unpublished literature

| Q7. Comments on the search strategy and language | Include clarification to the language question: ‘this protocol recommends including all the relevant languages (and notice language skills of the reviewers). Languages to be included depend on the review question and context.’
| Search strategy ok | National versions of the protocol will be developed in the second phase of the GEP if relevant
| | Some general instructions for developing a search strategy also for unpublished literature

| Q8. Databases that you would like to be included | Include statement: ‘Database list included is an example of the existing ones. Note: the databases and their addresses change frequently, thus, such a list needs to be updated continuously!’
| Database list accepted | National versions of the protocol will be developed in the second phase of the GEP if relevant
| | Some general instructions for developing a search strategy also for unpublished literature

| Q9. List here journals that you would like to be included | Add new journals suggested
| List of suggested journals ok with some adding | NOTE! The list of journal is an example and it’s not exhaustive list
| | National versions of the protocol will be developed in the second phase of the GEP if relevant
| | Some general instructions for developing a search strategy also for unpublished literature

| Q10. List here research register(s) (HP/PH) | Include a clarifying statement to research registers: ‘Search whenever relevant (it’s not something to do always, it’s an option).
| Generally ok | Include EPPI-Centers Trials Register of Promoting Health Interventions as an example
| | National versions of the protocol will be developed in the second phase of the GEP if relevant
| | Some general instructions for developing a search strategy also for unpublished literature

| Q11. List here the sources of grey literature | ‘Search whenever relevant (it’s not something to do always, it’s an option).
| Generally OK | National versions of the protocol will be developed in the second phase of the GEP if relevant
| | Some general instructions for developing a search strategy also for unpublished literature

| | National versions of the protocol will be developed in the second phase of the GEP if relevant
| | Some general instructions for developing a search strategy also for unpublished literature
Table 1. Summary on the 1\textsuperscript{st} Delphi round

| Q12. List here the search engines (internet) | Include a clarifying statement to the search engines: ‘Internet is most useful in finding information on relevant organizations and experts…’ |
| Q13. If expert opinion is to be included | NOTE: The GEP project builds on a broad consensus regarding the definition of ‘evidence’. Specify more clearly who are the experts and how to locate them |
| Q14. Selection criteria | Agreed on the proposed items of the selection criteria: • population • interventions/HP approaches • outcomes • study designs NOTE! time period can be part of the search strategy, but with out specified time period (e.g. 5 year; the time period depends on the question) Combine the selection criteria and quality criteria chapters: The selection of the information (e.g. include or exclude articles, documents) is the first step to select the relevant information for the review. Quality assessment will be done only to the articles relevant for the review question (e.g. for those included). Thus we prefer not to combine them. |
| Q15. Quality criteria of quantitative studies | Create our own quality criteria for all study designs? Create the grades of the evidence Create must do list |
| Q16. Quality criteria of qualitative studies | Create our own quality criteria for all study designs? Create the grades of the evidence Create must do list |
| Q17. How to assess the quality of other information than research | Create our own quality criteria for all study designs? Create the grades of the evidence Create must do list |
| Q18. Data extraction chapter | Generally OK Examples of the good data extraction forms: for each review question the reviewers may have to design own abstraction form, thus the creation of GEP abstraction form is not perhaps necessary Specify more clearly who should do the abstraction and the assessment of information |
| Q19 Data synthesis chapter | Generally OK Include some practical examples |
### Table 1. Summary on the 1st Delphi round

| Q20. Combining different sources of information? | Generally ok | Include references on statistical combination  
Explain also in a bit more detail consensus methods and the pro's and con's of this in comparison to triangulation  
Move away or delete the Bayesian hierarchical modeling from statistical combination and Triangulation should be discussed at the beginning of the chapter |
|---|---|---|
| Q21. Present the interpretation of the results | Generally ok | Modify the chapter according to given feedback:  
• what works  
• for whom  
• under which circumstances  
• how well |
| Q22. Reporting and disseminating and how to get evidence into practice | Generally ok | Own chapter for getting evidence into practice |
### Appendix 2. Suggested changes to the structure of the review protocol

<table>
<thead>
<tr>
<th>Delphi 1 structure</th>
<th>Suggestion</th>
<th>Proposed structure 1</th>
<th>Proposed structure 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>I INSTRUCTION</td>
<td>I INTRODUCTION</td>
<td>I INTRODUCTION</td>
<td>I INTRODUCTION</td>
</tr>
<tr>
<td>II BACKGROUND</td>
<td>New</td>
<td>II PREPARATORY PHASE</td>
<td>II PREPARATORY PHASE</td>
</tr>
<tr>
<td>III PREPARATORY PHASE</td>
<td>1 Planning phase</td>
<td>2 Forming review question</td>
<td>2 Forming review question</td>
</tr>
<tr>
<td>Delete</td>
<td>New</td>
<td>II PREPARATORY PHASE</td>
<td>II PREPARATORY PHASE</td>
</tr>
<tr>
<td>1 Planning phase</td>
<td>1 Planning phase</td>
<td>1 Planning phase</td>
<td>1 Planning phase</td>
</tr>
<tr>
<td>1.1 Resources (money /time)</td>
<td>1.1 Reviewers</td>
<td>1.1 Reviewers</td>
<td>1.1 Reviewers</td>
</tr>
<tr>
<td>1.2 Need/use of an advisory group</td>
<td>1.2 Advisory group</td>
<td>1.2 Advisory group</td>
<td>1.2 Advisory group</td>
</tr>
<tr>
<td>1.3 Number of reviewers</td>
<td>1.3 Resources</td>
<td>1.3 Resources</td>
<td>1.3 Resources</td>
</tr>
<tr>
<td>2 Forming review question</td>
<td>2 Forming review question</td>
<td>2 Forming review question</td>
<td>2 Forming review question</td>
</tr>
<tr>
<td>IV CONDUCTING THE REVIEW</td>
<td>3 Classes of sources of information</td>
<td>3 Classes of sources of information</td>
<td>3 Classes of sources of information</td>
</tr>
<tr>
<td>3 Classes of sources of information</td>
<td>3 Classes of sources of information</td>
<td>3 Classes of sources of information</td>
<td>3 Classes of sources of information</td>
</tr>
<tr>
<td>3.1 State of the art</td>
<td>DELETE</td>
<td>DELETE</td>
<td>DELETE</td>
</tr>
<tr>
<td>3.2 Research information</td>
<td>3.1 Research information</td>
<td>3.1 Research information</td>
<td>3.1 Research information</td>
</tr>
<tr>
<td>3.3 Un-published research</td>
<td>3.2 Non-research information</td>
<td>3.2 Non-research information</td>
<td>3.2 Non-research information</td>
</tr>
<tr>
<td>4 Search protocol/ Strategy</td>
<td>Change to chapter 5</td>
<td>4 Channels of finding information</td>
<td>4 Channels of finding information</td>
</tr>
<tr>
<td>4.1 List of search terms</td>
<td>4.1 Electronic literature databases</td>
<td>4.1 Electronic literature databases</td>
<td>4.1 Electronic literature databases</td>
</tr>
<tr>
<td>4.2 Languages included</td>
<td>4.2 Hand searching</td>
<td>4.2 Hand searching</td>
<td>4.2 Hand searching</td>
</tr>
<tr>
<td>5 Channels of finding information</td>
<td>Change before classes of sources</td>
<td>4.3 Research registers</td>
<td>4.3 Research registers</td>
</tr>
<tr>
<td>5.1 Electronic literature databases</td>
<td>4.4 Grey literature</td>
<td>4.4 Grey literature</td>
<td>4.4 Grey literature</td>
</tr>
<tr>
<td>5.2 Hand searching</td>
<td>4.5 Free Internet search</td>
<td>4.5 Free Internet search</td>
<td>4.5 Free Internet search</td>
</tr>
<tr>
<td>5.3 Research registers</td>
<td>4.6 Interviews of experts</td>
<td>4.6 Interviews of experts</td>
<td>4.6 Interviews of experts</td>
</tr>
<tr>
<td>5.4 Grey literature</td>
<td>4.7 Other personal contacts</td>
<td>4.7 Other personal contacts</td>
<td>4.7 Other personal contacts</td>
</tr>
<tr>
<td>5.5 Free Internet search</td>
<td>5 Selection criteria</td>
<td>5 Selection criteria</td>
<td>5 Selection criteria</td>
</tr>
<tr>
<td>5.6 Interviews of experts</td>
<td>6 Search protocol/ Strategy</td>
<td>6 Search protocol/ Strategy</td>
<td>6 Search protocol/ Strategy</td>
</tr>
<tr>
<td>5.7 Other personal contacts</td>
<td>6.1 List of search terms</td>
<td>6.1 List of search terms</td>
<td>6.1 List of search terms</td>
</tr>
<tr>
<td>6 Inclusion criteria</td>
<td>Change to selection criteria</td>
<td>6.2 Languages included</td>
<td>6.2 Languages included</td>
</tr>
<tr>
<td>7 Quality Criteria</td>
<td>Combine with selection criteria</td>
<td>7 Quality Criteria</td>
<td>7 Quality Criteria</td>
</tr>
<tr>
<td>8 Data extraction and synthesis</td>
<td>Divide to own main chapters</td>
<td>8 Data extraction</td>
<td>8 Data extraction</td>
</tr>
<tr>
<td>8.1 Data extraction</td>
<td>Delete (form a main chapter)</td>
<td>9 Data synthesis</td>
<td>9 Data synthesis</td>
</tr>
<tr>
<td>8.2 Data synthesis</td>
<td>Delete (form a main chapter)</td>
<td>9 Data synthesis</td>
<td>9 Data synthesis</td>
</tr>
<tr>
<td>8.3 Combining different sources of information</td>
<td>Inside the data synthesis chapter</td>
<td>9 Data synthesis</td>
<td>9 Data synthesis</td>
</tr>
<tr>
<td>9 Interpretation of results</td>
<td>10 Interpretation of results</td>
<td>10 Interpretation of results</td>
<td>10 Interpretation of results</td>
</tr>
<tr>
<td>V REPORTING AND DISSEMINATION</td>
<td>IV REPORTING AND DISSEMINATION</td>
<td>IV REPORTING AND DISSEMINATION</td>
<td>IV REPORTING AND DISSEMINATION</td>
</tr>
<tr>
<td>10 Reporting and dissemination</td>
<td>11 Reporting and dissemination</td>
<td>11 Reporting and dissemination</td>
<td>11 Reporting and dissemination</td>
</tr>
<tr>
<td>10.1 Writing the report</td>
<td>10.1 Writing the report</td>
<td>10.1 Writing the report</td>
<td>10.1 Writing the report</td>
</tr>
<tr>
<td>10.2 Disseminating the report</td>
<td>10.2 Disseminating the report</td>
<td>10.2 Disseminating the report</td>
<td>10.2 Disseminating the report</td>
</tr>
<tr>
<td>10.3 Getting evidence into practice</td>
<td>Own main chapter</td>
<td>12 Getting evidence into practice</td>
<td>12 Getting evidence into practice</td>
</tr>
</tbody>
</table>
Appendix 3. FEEDBACK REPORT ON THE CONSENSUS MEETING

Feedback report Consensus meeting - European Review Protocol for HP Workshop Pre-conference 31\textsuperscript{st} of May 2005

Räty S. & Aro A.R.

20\textsuperscript{th} of June 2005
1 BACKGROUND

Getting Evidence into Practice third project meeting was held in Stockholm on 31st of May – 1st of June 2005 as a pre-conference of the Best Practice for Better Health 6th IUHPE European Conference on the Effectiveness and Quality of Health Promotion. At the pre-conference Strand I organized a workshop concentrating on the critical issues of the European Review Protocol for health promotion (HP) version 1.1. The workshop acted as a face to face consensus meeting e.g. the third Delphi round for strand I. The aim of the meeting was to discuss on the broad definition of evidence and to reach a consensus on quality assessment criteria and grading system of the evidence as well as to get some feedback and ideas on the finalization for the review protocol.

The European Review protocol for HP version 1.1 (based on the inventory of existing review protocols and two Delphi rounds by strand I project team members and partners) was sent to all GEP projects’ pre-conference participants as background reading for the workshop. Altogether twelve project team members and partners took part in the workshop. Six of the participants were researchers, three policy makers (or researchers mainly involved in policy subjects) and the rest were either researchers or policymakers with more practical perspective. Three of the participants had taken part in the development of some review protocol or actively worked with systematic reviews.

2 RESULTS

2.1 Broad definition of the evidence

The GEP project defines the evidence to go beyond the randomized control trials (RCT) and takes into account also other information than research such as grey literature and knowledge from practice. All participants of the workshop agreed on the definition. It was seen that the European review Protocol should answer to the current situation of the health promotion and public health (HP/PH) research. As the research does not (always) take into account the complex and multi-level nature of the HP/PH interventions we cannot only rely on the information coming from research. Thus we have to use whenever relevant also other information sources than research. It was suggested that the European review protocol should be seen as an answer to the current ‘handicapped situation of the research world’ but also in the long run as a facilitator to improve the quality and appropriateness of HP/PH research information.

However, it was stressed that the research information should be seen as a primary source of information and if this information is not enough comprehensive e.g. in covering also the contextual factors then also other information sources should be searched. When including other information sources it should be made crystal clear in the end-report (review) what information and which interpretations of the results are based on the research information and which on other sources.

The importance of making a clear distinction between implicit and explicit information and knowledge of the effective HP information was discussed. This would have consequences for the use of the terms research and non-research.
Figure 1. Implicit and explicit information and knowledge in HP: Implicit knowledge is practical and subjective know-how of effective HP interventions. Grey literature does bring some of this know-how above surface. However, only when information is published and thus judged by the scientific community it becomes solid and transparent evidence e.g. explicit information.

2.2 Quality criteria for HP/PH reviewing process

It was agreed that at this stage the European Review protocol will list the existing quality criteria. No consensus was achieved on which of these criteria to recommend as the most appropriate. However, it was suggested that the quality assessment of the non-research information should include more practical instructions, but at this stage of the project and with in the given time limits the production of these remains a task of the second phase. NOTE: the quality criteria for project-based interventions are the work of strand II

2.3. Grading the evidence

In general it was agreed that all forms of information should be graded in same way. The GEP project will not suggest developing separate grading systems for research and non-research information. Concrete development of this remains a further challenge.
3 CONCLUSIONS

- Small changes to the review protocol:
  - More clearly: Go for the research first, if not found, then go for the non-research information.
  - Introduction: the description of the current situation (‘Handicapped’ situation of HP/PH research)
  - One aim of the review protocol is in long run to increase the quality and appropriateness of the HP/PH research e.g. use of theory and taking the contextual factors into account.

- Further development in the second phase of the GEP
  - Recommend some of the existing quality criteria
  - Develop an own grading system for the evidence using the HDA’s/ NICE’s idea (as presented in the protocol at the moment) as a starting point.
Appendix 4. FEEDBACK REPORT ON THE FEASIBILITY TEST

FEEDBACK REPORT ON THE FEASIBILITY TEST OF STRAND I, GEP

Räty S. Parry-Langdon N., Läubli Loud M., Drahonovska H., Bos V. Aro A.R.

15th of July 2005

Getting Evidence into Practice, GEP project. EC project no.20031123
(Evidence Consortium 790841)
1. Background

Based on the inventory of the existing review protocols and two Delphi rounds (Appendices 4A-B) the first version of the European Review protocol for health promotion, version 1.1 was developed. The aim was to test the usefulness of the review protocol and to further develop it with the help of a feasibility test.

The idea of the feasibility test was discussed and established during the last steering group meeting 1st of February in Woerden, the Netherlands. The aim was to carry out the feasibility test in six Strand I participant countries. All Strand I project team members and partners were given the possibility to participate. Four project team members were willing to carry out the test. The feasibility test started on 23rd of May, and the deadline for the feasibility test was 15th of June 2005.

Four selected project team members interviewed two review experts in their own country. A review expert was defined as a person with experience in reviewing literature and preferably also other sources of information (grey literature, documents) of health promotion and public health topics. The project team members were instructed to contact the experts and agree on the interview time (about 1,5 h/ interview). Experts were asked to critically read the European review protocol keeping in mind the overall aim of the GEP project to broaden the evidence base beyond the research-based evidence to health promotion/public health (HP/PH) experts and practice.

The interview was based on the evaluation of the latest version of the European review protocol version 1.1 with the help of the following critical questions:

1. General comments on the protocol: structure and outlook
2. What is your overall impression of the aims and usefulness of the protocol?
3. How different is the European review protocol from the protocol/techniques you have been used to apply in the reviewing process?
4. On the basis of the previous reviews you have been conducted (or read/evaluated), how might your findings (the findings of review) differ if you had used the European review protocol for health promotion / public health (HP/PH)?
5. What can/could the European review protocol add to your work in doing a review?
6. What are the strengths and weaknesses/ challenges of the protocol?
7. At the moment there are no agreed quality criteria for non-research information (e.g. governmental/ non-governmental documents, evaluation reports, expert opinion etc.) in HP. How would you start the creation of quality criteria for non-research information?
8. At the moment there are no generally accepted grading systems for non-research information in HP/PH. How would you create the grading system non-research information?

After the interviews the interviewers were asked to make a short report with conclusions (Appendices 4A-D).
2 Results

Altogether eight experts (outside GEP) were interviewed by four project team members. Two project team members carried out the test before the deadline (15th of June) and one organization asked for more time. The last results were received on 8th of July 2005. Full version of the feedback reports and the results of the interviews can be found in appendices 4A-D.

Most of the outside experts (6/8) found the Strand I work very essential and important. It was seen that the work so far done would lead the way of evidence-based thing to the directions that support the health promotion and public health interventions better that the traditional clinical review protocols. Only one of the interviewed experts did not find the protocol useful or agree with the GEP idea of broadening the evidence base beyond research information. All others found the broad idea useful for the health promotion and public health perspective. It was highlighted that the protocol is a good tool to collect especially policy-related information. In addition, the protocol was seen as a good possibility to enhance the HP/PH education, especially in the Eastern European countries.

The respondents suggested minor changes concerning mainly the spelling and clarity of the text. For example, it was suggested to broaden the target group from researchers and reviewers to the practitioners. It was seen that also practitioners should do the reviews and they are able to do them with somewhat more precise instructions.

In addition to minor changes also some further development ideas were given. It was suggested to develop a smaller version of the protocol and make it more user-friendly by including e.g. an index.

The review protocol was found not so different from the traditional formats of the review protocols, and thus using might not give that different results. However it was seen that the contextual factors and theory base would become more clearly covered with the help of this protocol. It was also seen that the European review protocol would be used besides already existing review protocol because some of its content are not enough rigorously covered yet. The main challenges in the future would be to recommend some of the quality criteria as a European Standard and develop the grading system to support the broad definition of evidence. Also the inclusion of practical examples would be needed in the long run. That would give unified and rigorous instructions on how to do the critical phases of the reviewing process and enable also the comparison of the results and collaboration between countries.

3 Conclusions

Generally the European review protocol for health promotion was seen very welcome in the field of health promotion. Minor changes according to the feedback received will be carried out for the end product: European review protocol for health promotion, version 1.2. However, some major changes and also issues will remain further challenges and they will be carried out in the second phase of the project and with the help of smaller expert group. Fore example the quality criteria as well as grading system for non-research information and expert opinion are still lacking. The first steps to encounter these challenges have already been made: the forming of a smaller expert group to support the further development work has been started.
1. General comments on the protocol: structure and outlook

- The expression ‘non-research’ evidence should not be used as a generic term to cover all grey literature and unpublished research – these sources very often include high quality research (e.g. those that have found negative findings) and good evaluation designs (e.g. government reports for policy development that are not in the public domain)
- It is good that there is an emphasis on expert consensus
- The protocol may be used by all practitioners (expert reviewers to those with some experience – although this latter group will need guidance and direction e.g. when to seek help from expert reviewers

2. What is your overall impression of the aims and usefulness of the protocol?

- The protocol is very big – there needs to be a shorter version (e.g. a core set of ‘must do’ processes)
- If it is to be Euro-wide the protocol must be tightened up i.e. it must be more prescriptive to allow reliable comparisons between countries
- Everyone must conduct the review in the same way so as to achieve consensus on the quality of the evidence e.g. use the same data extraction sheet
- If it is not a step-wise process (i.e. accept certain level of evidence at the outset) – what is reasonable to include as it may be just to big and time-consuming to conduct (unless it is a particularly unusual topic area which requires wide scoping)
- These are issues to do with quality control which must have core recommendations - possible trial the process in a small number of countries on one particular topic and recommend a method (of data extraction)

3. How different is the European review protocol from the protocol/techniques you have been used to apply in the reviewing process?

- Again – use of the term ‘non-research’ literature (see response to question 1 above)
- Need to consider if the evidence is salient and feasible (relevant and appropriate) – not just classified as grey/non-research
4. On the basis of the previous reviews you have been conducted (or read/evaluated), how might your findings (the findings of review) differ if you had used the European review protocol for health promotion / public health (HP/PH)?

• There is not enough time to look at everything at the same time – what should be the absolute minimum requirements to conduct acceptable reviews?
• It would be a useful exercise to run shorter ‘rapid reviews’ on a small number of topics and compare findings with the full length review results.

5. What can/could the European review protocol add to your work in doing a review?

• There is a need for tighter, higher quality standards – especially for Europe wide work
• This protocol – as long as it can achieve consensus – will be extremely useful.

What are the strengths and weaknesses/ challenges of the protocol?

• There is a lack of specific guidance and direction – therefore not comparable findings
• Great strength in locating public health expertise as European collaboration
• There need to be worked examples (or links to) of how to synthesise data – this is very difficult and may be done very badly

7. At the moment there are no agreed quality criteria for non-research information (e.g. governmental/ non-governmental documents, evaluation reports, expert opinion etc.) in HP. How would you start the creation of quality criteria for non-research information?

• Must treat every piece of information has potentially the same contribution to make – it is wrong to start with the question on what is low quality evidence (see response to question 1)
• There are no worked examples (or links to) on how to judge quality of qualitative information
• Not all grey literature are qualitative studies (e.g. evaluation studies by government departments)

8. At the moment there are no generally accepted grading systems for non-research information in HP/PH. How would you create the grading system non-research information?

• This protocol does not grade the quality of the evidence
• There is a need to provide guidance/direction (critical appraisal) on how to interpret grades of evidence for recommendation for use by expert reviewers.
II Interview with Lesley Sander
Information Specialist, SURE (Support Unit for Research Evidence), Cardiff University.
14th June 2005

1. General comments on the protocol: structure and outlook
   • As a finished product the protocol as it stands (in landscape) is unwieldy and would be easier to handle in portrait orientation
   • For small projects, practitioners may try to dip in for certain points. Practitioners with little or no knowledge of the review process should start at the beginning – the flow chart should make this clear
   • May be used by all practitioners (expert or less expert) – but needs clear guidance

2. What is your overall impression of the aims and usefulness of the protocol?
   • As it stands it is an additional/complementary useful tool
   • Some of the components are not rigorous enough e.g. not enough discussion [on the importance] of pre-defined parameters such as inclusion/exclusion criteria (CRD make a strong case for this)

3. How different is the European review protocol from the protocol/techniques you have been used to apply in the reviewing process?
   • It is good to comply with other review protocols such as CRD and Cochrane (establishing and maintaining standards)
   • Although including Expert Opinion may be seen as a retrograde step [by some expert reviewers] it is essential that non-research information/grey literature is included in reviews
   • Needs to be more guidance and direction on where to find grey literature (e.g. SIGLE has been omitted form sources of grey literature)
   • Grey literature should not be routinely equated with lower level evidence

4. On the basis of the previous reviews you have been conducted (or read/evaluated), how might your findings (the findings of review) differ if you had used the European review protocol for health promotion / public health (HP/PH)?
   • Findings would not differ - although this ultimately depends upon the inclusion and exclusion criteria established for a [any] review

5. What can/could the European review protocol add to your work in doing a review?
   • It is good to encourage multiple methods for evaluating outcomes
   • Bear in mind that for the review process there are still the recognised difficulties such as self-reported data etc.
What are the strengths and weaknesses/challenges of the protocol?

- Need a note to emphasise that the data extraction sheet/s must be piloted
- Need a note to emphasise that expert advice should be sought to develop the search strategy
- Synthesis of results is potentially difficult (p37)
- Practitioners with less expertise should be advised on when (at what point in the process) they should consult an Expert Reviewer for guidance
- All practitioners should be advised on when they should consult wider expertise e.g. economist, statistician, etc.

7. At the moment there are no agreed quality criteria for non-research information (e.g. governmental/non-governmental documents, evaluation reports, expert opinion etc.) in HP. How would you start the creation of quality criteria for non-research information?

- This is a very difficult question to answer
- This must be established before a grading system for non-research information
- For comparison purposes, it may be useful to submit a review using the GEP protocol for review by Cochrane

8. At the moment there are no generally accepted grading systems for non-research information in HP/PH. How would you create the grading system non-research information?

- Also a very difficult question to answer
Appendix 4B. FEASIBILITY TEST OF STRAND I, GEP (Marlène Läubli Loud)

Interviews with Dr. Bernard Burnand of the Institute of Social and Preventive Medicine, University of Lausanne and Prof. Jürg Rehm, Director of the Institute for Addiction Research, Zurich
20th June 2005

Background

- Dr. Bernard Burnand of the Institute of Social and Preventive Medicine, University of Lausanne: a doctor / epidemiologist whose main working experience has been in epidemiological research. He is also a member of the Swiss and International Health Technology Assessment Group.
- Prof. Jürg Rehm, Director of the Institute for Addiction Research, Zurich: professor in public health sciences, psychologist by background, with many years experience in the field of drug dependency research both nationally and internationally. On several occasions, he has served as an advisor about “what works” to health departments in various countries which were at the time, planning prevention measures in relation, for example, to alcohol abuse. Experienced in the use of review protocols, particularly for more medical interventions e.g. heroin treatment, methadone etc.

For the first person, we were able to discuss the protocol “face-to-face”, but for the 2nd person, being in the US at the time, I called him to discuss the protocol by ‘phone. This didn’t prove a problem at all; the line was good, and we were not disturbed.

Below I have set out the main points of our discussion around the questions provided by the GEP Strand 1 Project Team

General Comments – pros and cons of the protocol

- Both considered that the Project was ambitious but innovative!
- The idea of developing a standardised review protocol for European context was most welcomed.
- The proposal was thought to be particularly relevant to policy issues e.g. for establishing evidence in the field of health promotion and disease prevention.
- The proposed review protocol could help identify the cultural differences which could explain differences in effectiveness reported between different countries. For example, one mentioned that in a current study in which he’s involved, there are differences between countries in the reported effectiveness of a particular type of intervention. “We haven’t taken enough account of the cultural / contextual differences in our reviews – this protocol made me more aware of this”.


• Integrating qualitative data, and grey literature into the analysis was thought to be challenging!
• Lots of interesting and detailed ideas, but too dense! Needs index to help find relevant information when needed. “I’ve been too restricted to medical intervention field – this gives me more ideas, new information – can always learn something”.

**The negative sides:**

• the protocol promised much but in the end, didn’t differ that much from current protocols. Seemed to follow classic approach instead of offering better ways of integrating information on cultural/contextual/procedures into analysis on evidence.
• Too detailed and should serve as a reference only – when actually conducting such a review, one would have to be far more pragmatic. For example, one respondent said that if you try to adhere to all these criteria, you could end up with too few studies (minimal number) on which to base your evidence. He tends to define the inclusion/exclusion criteria beforehand – by looking at a ‘good’ review and its methodology section to draw up his criteria. Will then look at abstracts to identify possible studies for “sensitivity analysis” but then adapts his original inclusion criteria wherever necessary if he can see that it isn’t bringing in enough studies. Uses therefore a two step process to finalise inclusion/exclusion criteria.
• One respondent also mentioned the danger of including theory as inclusion criteria. There is the risk of over relying on a theoretical model for the evidence! “Theories can become over theorised and “spoil” the evidence because they themselves always change too”.
• In general, protocol offers broad scope, but lacks validated instruments for establishing the evidence.
• Must define the questions and then use appropriate scientific approach.

**Challenge of including grey literature and expert opinions into evidence**

• Whilst both agreed that this idea was laudable, it poses problems with regard to the rigour of the data and ultimate “evidence” produced.
• Both respondents were sceptical about taking into account “policy documents” apart from one aspect: to provide some indication of culture / context. For example, in Japan policy documents warn against “drinking when depressed – only drink when happy” whereas in the USA, such documents warn against going over the internationally accepted limits.
• **Quality criteria for such data:** One respondent suggested must have some comparative scale. For example, would need to look at data from different countries to compare and see if there is a temporal trend across all – time and regional differences are most important comparative criteria. Maybe even have cross-cultural applicability as criteria for study inclusion. BUT
• **Inclusion of contextual differences within evidence**: Not a problem if you turn the argument around the other way. Rather than trying to establish evidence that is appropriate for all contexts, you have to do contextual analysis first and then look at your evidence to see which “tried and tested” interventions might be appropriate given what you know about the context. Quoted Campbell “establish principle of what works, then experiment in the society” Always problematic to use evidence and apply it beyond 1 country – always be contextual differences to take into account. To get contextual information between countries also poses problems. “I know the major players, context, experts etc here in Switzerland but wouldn’t know enough about Australian context to apply evidence there”

• **Grading? Both** said that if you want to have feasible evidence, you can’t avoid grading the data – in the traditional way. Expert opinion and grey literature should always be graded at the lower end of the scale. “When we looked at patterns of drinking problems throughout the world, we found published information for only 30 countries – practically all of which were from developed countries. So we conducted an expert survey and also asked them about what basis they used to make their judgment. Published literature, highest grade – expert opinion – lowest grade (if at all)!!

• The other respondent however, suggested looking at the RAND APPROPRIATENESS METHOD e.g. for methadone treatment. Identifies different scenarios, followed by systematic review the results of which are then validated with expert groups. Have to beware of the potential and level of bias.

**Conclusions**

The protocol is welcomed as an important tool for providing standardised guidelines and criteria – throughout Europe - for collecting and judging evidence. Main challenge is what to do with the proposed broadened definition of what evidence is, and therefore how we should collect and judge it!! This protocol goes further than current thinking and practice, but the major issues are still not resolved.

**Proposal**

• I would suggest that expert opinion be used to validate and judge the appropriateness of the collected evidence for its relevance to practice in a specific cultural and socio-political and economic context. Similarly at the other point in the cycle, expert opinion can be used to orient new research into the literature by helping formulate the appropriate research questions – based on their knowledge of practice.

• As for grey literature, based on both the results of my interviews from more classical systematic reviewers and the report from Cardiff, I would certainly stress the need to include quality evaluation reports and certain policy papers – however, this comes back to the problem of inclusion / exclusion criteria and the context / needs of the particular “evidence” study.
• As for grading evidence, this remains a controversial issue. The problem of assuring “rigour” (and therefore at least being sure of quality criteria) yet giving weight also to non-peer reviewed, unpublished opinion. How to do this?
• I strongly recommend that Phase 2 doesn’t go ahead with trying to pilot this protocol until these issues have been better resolved. It cannot hope to get European backing as it is without some more thought on the above issues.
• I recommend a small working party be established, to include Anthony Morgan, Erica Wimbush amongst others, to “brain storm” and come up with some quality criteria for the more controversial aspects of this protocol.
Appendix 4C. FEASIBILITY TEST OF STRAND I, GEP (Hanna Drahonovska)

Interviews with Hana Svinova, MD, NIPH, Head of Department for Alcohol, Tobacco and Illicit Drugs Abuse Prevention (Expert 1) and Lumir Komarek. Assoc. Professor, MD, PhD, NIPH, Head of Centre for Health Promotion, Diseases Prevention and Customers Protection (Expert 2)

Summary
The feasibility study had been carried out in weeks 23 and 24 in 2005 according to instruction sent by the coordinator of the Strand 1. Interview with the experts were done on 8th June 2005 with Dr. Svinova and 17th June with Professor Komarek. Time spent was about two hours for each interview.
Both experts were enthusiastic and willing to participate even they were surprised by length of the Review Protocol when they familiarised with. Experts have found the Protocol useful in general and both have quite clear imagination how to use it in the field of health promotion.

Results: answers on interview questions
Ad 1 General comments on the protocol structure and outlook
Expert 1: The protocol is structured as a systematic guidelines and contents all important items. A good thing makes sense for using in health promotion practise however limited to people with appropriate resources accessibility as written in chapter 1.3.
Expert 2: The structure is good as well as outlook; some part of text should be more structured in smaller paragraphs or parts. I was very happy with chapter on dissemination because it is health promotion marketing and researchers and practitioners often underestimate it. The protocol seems me a good tool for raising awareness on this issue.
GEP Project Member: Very good things are brief summaries at the end of each chapter

Ad 2: What is your overall impression of the aims and usefulness of the Protocol?
Expert 1: The protocol is an useful tool to make more effective and to some extend easier to search for evidence. It creates a system and thus findings from different reviewers could be comparable. It is a good thing that makes sense for using in health promotion practise however it is limited to people with appropriate resources accessibility (as written in chapter 1.3 of the protocol).
Expert 2: I can see the protocol as an excellent tool for education in health promotion. It could be used as a comprehensive learning material. I also see the concrete possibility to take the protocol as a part of our courses for health promotion professionals which have been already running.
GEP team member: The protocol, if widely use could avoid to circulation of “the gold true” or “evidence” which is based on telling of some officially established authority.
Ad 3) How different is the European Review Protocol from the protocol/techniques you have been used to apply in the reviewing process?

**Expert 1**: No comparable tool exists in our country – professionals in health promotion (mostly medical doctors or people with higher health education). Searching for evidence is based on the same principles as the protocol is but is dependent on individual opinion, education, knowledge and experience of every professional. In spite of it I do not think that findings of both process (the Protocol and Individual reviewing) could be dramatically different.

**Expert 2**: We use common scientific methods for evaluation of evidence from research and scientific literature (mostly clinical epidemiology and experimental studies). Disadvantage of these methods is that they are not applicable completely to health promotion sources of information. The European protocol is more tailored for health promotion. That is why findings could be different: our finding could lack of practised knowledge.

**GEP member**: The differences exist especially in way how to disseminate findings and in transferring evidence into practice because of lack of system process in our technique using for evidence search. From this point, the Protocol could be as the first step to create European projects in health promotion.

Ad 5) What can/could the European protocol add to your work in doing review?

**Expert 1**: If the Protocol could be widely applicable (at least in our Institute) I hope that a group of reviewers could be made up and I will serve as member of advisory group.

**Expert 2**: I see my part in using the protocol in process of education. I would expect is as a step to connect stakeholders, policy makers, practitioners, researcher and other professionals in health promotion and to learn them to speak the language to understand each other.

Ad 6) What are strengths and weaknesses/challenges of the protocol?

**Expert 1**: Strengths: High quality of background information
- Innovative system of reviewing
- Complexity of reviewing
- Good and logic Structure

Challenges: High demands on time, money, and people
- Necessity of training for potential users of the protocol

**Expert 2**: Strengths: Efficiency of the tool
- Systemic tool
- Specific for health promotion

Weaknesses: Demanding on resources
- Too complicated to be used for specially untrained persons
- Not used in general practice – only for special teams

Ad 7) How would you start the creation of quality criteria for non-research information?

**Expert 1**: - Create a group of experts (or to use the advisory group).
- Make an overview of all possible non-research information what are accessible and related to a review question.
- Apply the same criteria on non-research information as for the research one.
- Apply Delphi or other method to reach consensus on what quality criteria are possible to be applied on every judged non research information.
- Make a list of these “core” criteria and apply again on the judged information.

**Expert 2:** I take reports from health promotion projects as an example of non-research literature:
- Analysis of projects from National Health Promotion Programme classified as effective according Evaluation system using by experts of a funding body (the Ministry of Health)
- Search for common identical or similar indicators (items) of these effective projects.
- Experts group evaluation and checklist of these indicators.

**GEP member:** Somebody could doubt a little if the both processes described by experts are worth of doing it: they are time and other resources very consuming. Could not be easier and more effective to use only research information? I think that this process can be only beginning and after a list of criteria (indicators) of quality will be made, this process could be finished.

**Ad 8): How would you create the grading system non-research information?**

**Expert 1:**
- Scoring according to quality of a source of information (credibility)
- Scoring according to “core” list of criteria (question 7)
5 grades of evidence: e.g. very strong, strong, probable, less probable, no (or other words but not only yes, yes/no, no.

**Expert 2:** We have created the checklist of indicators (question 7) I suggested following:
- Make a scale of evaluation of individual indicators (items) from the list according their importance for effective intervention
- Create a scoring system evaluating information (projects report) from scale of the individual indicators
- Set down grading system of extracted evidence. (I have no preferences on number of grades or their expression)

**GEP member:** I try to complete both experts suggestion with few comments:
- Credibility of source could be evaluated regarding: previous good experience, if information source is credible professional body, international organisation as e.g. WHO, information was evaluated under national (regional) official (authority) protocol or method etc.
- “Core criteria” or minimum, threshold criteria could be: Source of information: credibility; Objective: clear described
Results: clear formulated with minimum but clear explanation how they were reached
Intervention: was evaluated to some degree at least.

**Conclusion**
Both experts and me as a GEP team member agreed that the Protocol is excellent piece of work, useful for using in the field of health promotion. If used properly it could contribute to development
evidence based health promotion on regional, national, European and worldwide level. Careful translation and modification is necessary for implementation of its great potential.

In Prague, 17.06.2005  Hanna Drahonovska
Appendix 4D. FEASIBILITY TEST OF STRAND I, GEP (Vivian Bos)

I Interview with Martijnje Bakker

1. General comments on the protocol: structure and outlook
   - Is ok. But before we look at the structure it is important to assess its relevance and scope. Why do we need a new protocol?

2. What is your overall impression of the aims and usefulness of the protocol?
   - Martijnje states that the problem is that prevention workers do not tend to use information that comes from reviews. This may partly have to do with the nature and content of reviews, but this may also have to do with other issues. Prevention workers at municipal health services do not regard it as their task to look into evidence on interventions. Martijnje considers it a task for the NIGZ (and similar institutions) to look at the evidence basis of interventions and communicate the results to the municipal health services. Martijnje doubts whether changing the nature of reviews would stimulate prevention workers to use information from reviews. She wants information on what works and what does not work and she is not helped with a review protocol on other forms of evidence. She thus questions whether a review protocol is the best solution
   - She acknowledges the problem that reviews do often not sufficiently describe interventions. The problem here lies however with the journals and not necessarily so with review protocols. It should become easier to publish process evaluations.

3. How different is the European review protocol from the protocol/techniques you have been used to apply in the reviewing process?
   - She mentions that for other forms of evidence it becomes increasingly important to perform a review with more than one person because the judgement of the information becomes increasingly difficult
   - We discuss the kind of research questions that could be answered by this review protocol. We think that reviews often deal with questions as ‘does it work or does it not work?’. As results of studies become more difficult to compare because the study designs are weaker, it will become increasingly difficult to answer such a question. So with lower forms of evidence it is probably not possible to answer the question whether a certain type of intervention works. Martijnje suggests that there should be a link between the type of question that is being asked and the level of evidence that is used to answer the question. Questions as ‘what has been developed at a certain area’ or similar more descriptive questions are more suitable to answer with ‘lower’ forms of evidence. After some brainstorming, Martijnje mentions as other examples of suitable questions: “which aspect of the intervention was performed as it was intended?”, “what were problematic and successful aspects of the intervention?”. She is not sure, however, that a
review protocol is the best way to proceed with such a type of question. Would a Delphi method not be better?

- Martijntje states that the review protocol should aim for publications in ‘other’ type of journals. For example more profession-related (beroepsgroep-geoorienteerde) journals. Prevention workers are more likely to read such journals (solving the first mentioned problem) and such journals are more likely to publish reviews that deal with more descriptive information.

4. On the basis of the previous reviews you have been conducted (or read/evaluated), how might your findings (the findings of review) differ if you had used the European review protocol for health promotion / public health (HP/PH)?

5. What can/could the European review protocol add to your work in doing a review?

- Martijntje sees as a strong point of the review protocol that it is aimed at other forms of evidence. She also thinks that more use should be made of such evidence. And she believes in combining the bits and pieces that are available.

**What are the strengths and weaknesses/ challenges of the protocol?**

- She regrets however that in the discussion on “how can we increase the visibility of such types of information and do more with it” no prevention worker has been included. She regards a review protocol as an academic product and doubts whether a review protocol is a good solution for the problem.
- She thinks that the review protocol could add if it focuses on more process orientated outcome measures (and not at effectiveness).

7. At the moment there are no agreed quality criteria for non-research information (e.g. governmental/ non-governmental documents, evaluation reports, expert opinion etc.) in HP. How would you start the creation of quality criteria for non-research information?

- For quality criteria, criteria that are used in more qualitative research could be tapped.

8. At the moment there are no generally accepted grading systems for non-research information in HP/PH. How would you create the grading system non-research information?
1. General comments on the protocol: structure and outlook

See question 2.

2. What is your overall impression of the aims and usefulness of the protocol?

Dr. Cuijpers is very skeptical about the usefulness of the protocol. He acknowledges that there are ‘other’ forms of information (i.e. besides those stemming from RCT’s) and that this information can/should be used more than is currently being done. He says that what you should aim for is to systematize this information in order to answer specific types of research questions. You cannot use this information to answer questions as ‘is this intervention effective or not’. For questions as this you’ll always need reviews that make use of information that is generated by RCT’s.

He has the opinion that you should not mix different kinds of information. These different types of information serve a different aim (answer a different question) and should be used according to this aim.

His has strong objections to the suggestion that this protocol raises. That is: that other types of information can help to generate answers that can otherwise not be found. Other types of information help to answer other types of questions. For example questions as:
- is this intervention able to meet the needs of people?
- what is the theoretical background of interventions as these?
- how many people are reached with the intervention and what were the factors that enabled/disabled this?

Reviewing more qualitative types of studies can help to generate information about these kind of questions. One should take into account that the validity of a review that looks at more qualitative research is rather limited. The problem is that the studies that are used in the review are not very comparable. It would be better to ensure that studies are implemented in a similar manner in different environments for example, and than to combine the information that comes from this. You should not mix findings that stem from such different sources.

3. How different is the European review protocol from the protocol/techniques you have been used to apply in the reviewing process?

Very different. He thinks that traditional review protocols are suitable to systematize non-research information. There is no need for a new protocol.
4. On the basis of the previous reviews you have been conducted (or read/evaluated), how might your findings (the findings of review) differ if you had used the European review protocol for health promotion / public health (HP/PH)?

He does not agree with the solution of combining different types of information.

5. What can/could the European review protocol add to your work in doing a review?

He does not want to use this protocol.

6. What are the strengths and weaknesses/challenges of the protocol?

The weakness is that different types of knowledge are combined and that this can generate answers to research question that would otherwise not be found. For different types of questions different types of information should be used. The solution is in the content of the research question. The choices that are made (inclusion criteria, type of research that is used etc) should be dictated by the research question. One should think about the question which kind of research can generate which kind of information?

7. At the moment there are no agreed quality criteria for non-research information (e.g. governmental/ non-governmental documents, evaluation reports, expert opinion etc.) in HP. How would you start the creation of quality criteria for non-research information?

There are all kind of quality criteria available. For example criteria for trails, for evaluation research, for screening, for public health research etc. These criteria should be used. There is no need to develop new criteria.

8. At the moment there are no generally accepted grading systems for non-research information in HP/PH. How would you create the grading system non-research information?

You should not combine research with different ‘grades’.
References


This report was produced by a contractor for Health & Consumer Protection Directorate General and represents the views of the contractor or author. These views have not been adopted or in any way approved by the Commission and do not necessarily represent the view of the Commission or the Directorate General for Health and Consumer Protection. The European Commission does not guarantee the accuracy of the data included in this study, nor does it accept responsibility for any use made thereof.