Belgian Red Cross

CHARTER:

DEVELOPMENT OF EVIDENCE-BASED PRACTICE GUIDELINES

AND

DEVELOPMENT OF A SYSTEMATIC REVIEW

By the Centre of Evidence-Based Practice (CEBaP) of the Belgian Red Cross-Flanders

De Buck E, Pauwels NS, Dieltjens T, Vandekerckhove P. *Use of evidence-based practice in an aid organization: a proposal to deal with the variety in terminology and methodology*. Int J Evid Based Healthc., 2014, 12(1):39-49.

Belgian Red Cross

INHOUD

A. INTRODUCTION 4

BACKGROUND 4
DEFINITIONS 4
SEMANTICS 7
PROJECT INITIATION 8

B. DEVELOPMENT OF AN EVIDENCE-BASED RECOMMENDATION AND PRACTICE GUIDELINE 9

CHOICE OF THE SUBJECT 9
DEVELOPMENT OF AN EVIDENCE-BASED PRACTICE GUIDELINE ACCORDING TO AGREE II 9
1) SCOPE AND PURPOSE 9
2) STAKEHOLDER INVOLVEMENT 10
3) RIGOUR OF DEVELOPMENT 10
3.1 SYSTEMATIC LITERATURE SEARCH 10
3.2 FORMULATING RECOMMENDATIONS 14
3.3 REVIEW OF THE PRACTICE GUIDELINE 14
3.4 UPDATE OF THE PRACTICE GUIDELINE 14
4) CLARITY OF PRESENTATION 15
5) APPLICABILITY 15
6) EDITORIAL INDEPENDENCE 15
PUBLICATION OF A PRACTICE GUIDELINE 15

C. DEVELOPMENT OF A SYSTEMATIC REVIEW 16

CHOICE OF THE SUBJECT 16
DEVELOPMENT OF A SYSTEMATIC REVIEW ACCORDING TO THE COCHRANE HANDBOOK 16
1) DEFINING THE REVIEW QUESTION AND DEVELOPING CRITERIA FOR INCLUDING STUDIES 17
2) SEARCHING FOR STUDIES 17
3) SELECTING STUDIES AND COLLECTING DATA 17
4) ASSESSING RISK OF BIAS IN INCLUDED STUDIES 17
5) ANALYSING DATA AND UNDERTAKING META-ANALYSIS 17
6) ADDRESSING REPORTING BIASES 18
7) PRESENTING RESULTS AND SUMMARY OF FINDINGS TABLES 18
8) INTERPRETING RESULTS AND DRAWING FINAL CONCLUSIONS FOR PRACTICE 18
PUBLICATION OF A SYSTEMATIC REVIEW 18

D. KNOWLEDGE MANAGEMENT 19

EVIDENCE SUMMARY DATABASE 19
REFERENCES 20
APPENDIX 1: EVIDENCE SUMMARY AND GUIDELINE SUMMARY TEMPLATE 21
APPENDIX 2: PRISMA CHECKLIST 23

APPENDIX 3: TABLE OF METHODOLOGICAL PRINCIPLES OF GUIDELINE DEVELOPMENT VERSUS SYSTEMATIC REVIEW DEVELOPMENT (BRCF) 25
APPENDIX 4: PROJECT INITIATION AND PROJECT FLOW 26
APPENDIX 5: FLOWCHART STUDY DESIGNS 27
APPENDIX 6: STANDARD WORDING OF EVIDENCE CONCLUSIONS 28
APPENDIX 7: SUMMARY OF THE GRADE APPROACH FOR THE ASSIGNMENT OF LEVEL OF EVIDENCE 33

A. INTRODUCTION

BACKGROUND

Belgian Red Cross-Flanders (BRCF) comes up for the rights of vulnerable people (social assistance, relief, training, international aid, blood transfusion,…). The mission of the Centre of Evidence-Based Practice (CEBaP) consists of formulating evidence-based recommendations to support the activities and interventions of BRCF. The recommendations are collected in evidence-based practice guidelines (developed according to AGREE II) and in systematic reviews (developed according to the Cochrane Handbook), thus guaranteeing knowledge dissemination by publication in peer-reviewed journals. The methodology that is being used by CEBaP is described in detail in this charter, and has been published in *International Journal of Evidence Based Healthcare* [1].

DEFINITIONS

AGREE II

The AGREE (*Appraisal of Guidelines for Research & Evaluation*) instrument was developed to promote the quality of practice guidelines [2]. It is a tool that assesses the methodological rigour and transparency with which a guideline is developed. The original AGREE instrument has been refined, resulting in the new AGREE II. The purpose of AGREE II is to provide a framework to (1) assess the quality of guidelines, (2) provide a methodological strategy for the development of guidelines and (3) describe what information ought to be reported in guidelines and in which format.

Bias

In the case of bias, the results or the interpretation of a study differ from reality by a systematic error. Bias can be the result of an error in any of the steps of a study, such as preparing the study, collecting data, analysing and interpreting the results and reporting.

Body of evidence

The quality of all available studies (and not of individual studies) is being assessed and summarised.

Cochrane

An international organisation that gives support in making informed decisions about health care by publishing systematic reviews and meta-analyses about the effect of health care interventions.

Evidence

Scientific data to support the answer to a specific question.

Evidence summary

Transparent and structured summaries of scientific literature related to specific practice questions. In an '*evidence summary*' the following categories are described: topic; subtopic; intervention; question (PICO); search strategy; search date; inclusion and exclusion criteria; characteristics of included

studies; synthesis of findings; quality assessment; conclusion; references. The *Evidence summary* template can be found in Appendix 1.

Evidence-based methodology

The process by which decisions are being taken based on the best available scientific evidence, practical experience, preferences of the target group and the available resources.

Expert panel

1 chair and additional panel members that as a minimum have expertise related to the content of the project to give evidence-based support for the practice guideline or recommendation.

Forest plot

A graphical representation of the results of different studies included in a meta-analysis. The point estimates and confidence intervals of every study are shown in a horizontal line below each other and beneath them the pooled result is shown as a diamond, the estimate of the global effect.

Funnel plot

A graphical way to assess publication bias while performing a meta-analysis. Therefore, for every study the effect is plotted against the sample size. The distribution of the points in this graph should take the form of a funnel, in which the dispersion increases as the sample size decreases. In case of asymmetry we can assume that studies are missing (for example, these are not published or not included based on the search strategy).

Good practice point

Where no good-quality evidence is available but consensus among experts with practical experience exists, consensus-based recommendations are given. Such recommendations are called 'Good Practice Points' (GPP).

GRADE

A methodology that can be used to assess the level of evidence of studies and the grade of recommendation based on the corresponding evidence.

Grey literature

Publications that are developed by the government, academics, business and industry, which are not published in easily accessible journals and may not appear in databases or via web searches. Examples are: government reports, conference proceedings, abstracts of the research presented at conferences, technical reports, dissertations or other types of documentation.

Meta-analysis

A statistical method in a systematic review in which the results of a number of comparable studies are pooled and recalculated. By doing this, it is possible to better estimate the true "effect size" and to be more confident about the effect of an intervention or treatment.

NICE

The 'National Institute for Clinical Excellence', an independent British organization that develops evidence-based clinical guidelines to improve people's health and prevent and treat ill health.

PICO

A PICO (population [P], intervention [I], comparison [C] and outcome [O]) question is a specific question which precisely defines the population, intervention, comparison and outcome under consideration.

PRISMA

PRISMA or 'Preferred Reporting Items for Systematic reviews and Meta-Analyses' is a checklist, composed of 27 items, that aims to increase transparency and clarity when publishing a systematic review (Appendix 2).

Practice guideline

A document with recommendations, advice and instructions to support daily practice, based on the results of scientific research, discussion and decision making, aimed at good practice.

Publication

Each document that is being published in a (scientific) journal, magazine, newspaper or handbook, each document that is submitted to a scientific congress/symposium, and each document that is present on an information carrier (e.g. CD or DVD).

Recommendation

Advice or statement in which a certain technique, intervention, process or activity is being recommended.

SIGN

The 'Scottish Intercollegiate Guidelines Network' which develops evidence-based clinical practice guidelines for the National Health Service in Scotland.

Steering committee

The CEBaP Steering committee is composed of the several directors and managers of the Belgian Red Cross-Flanders: the manager of CEBaP, the Director of Humanitarian Services Flanders, the scientific co-ordinator of the Humanitarian Services, the Medical Director of the Blood Service and the CEO.

Summary of findings table

Contains the major results in a transparent and simple way. In the table the most important outcomes, the effect sizes, and the number of participants are reported.

Systematic review

Literature summaries that aim to answer a specific question on the effectiveness of interventions by performing a systematic search in available literature. The term 'systematic' indicates that specific attention is given to formulating the methods of data collection and handling, in order to provide a transparent methodology to the reader who can then make a judgement about the quality of the

literature search. This profound method minimises the risk of bias and results in the “best available scientific evidence”.

SEMANTICS

1. **Systematic review**

A systematic review gives an overview of the best available scientific evidence collected by a literature search on a very specific topic or question en can be used to inform policy makers.

When using the term ‘systematic review’ the semantics are very important: a systematic review literally means ‘performing a literature review in a systematic way’. However, many variations and gradations exist in performing a systematic literature review. Peer-reviewed publications that call themselves ‘systematic review’, thus may differ significantly in methodological quality.

The Cochrane Collaboration uses the strictest methodological criteria for the development of systematic reviews. Any systematic review developed by BRCF will be performed using the methodological principles of Cochrane. These are described in detail in the charter below.

2. **Guideline**

In a guideline recommendations for practice are being made, to assist practitioners/volunteers on the field. These recommendations are the result of balancing the quality of the evidence, benefits and harms, costs, and the preferences of the target group.

There is also a range of methodological variants among guidelines. A guideline is not an ‘evidence-based guideline’ by definition. Currently the AGREE II checklist is being used as ‘the golden standard’ by guideline developers [2]. A guideline that is developed according to the domain “rigour of development” in this checklist can be called an ‘evidence-based guideline’.

For the development of practice guidelines BRCF uses the AGREE II checklist. This checklist recommends performing the literature search in a systematic way. However, we do not call this a ‘systematic review’ according to the strict Cochrane definition, because we make a compromise between the number of topics on the one hand and a reasonable time span for the development of the practice guideline on the other (cf. SIGN methodology). This results in a methodology that is systematic but less rigorous than is the case for a Cochrane systematic review: a specific search strategy instead of a sensitive search strategy, 1 reviewer instead of 2 reviewers and not handsearching instead of handsearching for evidence. However, for guideline development, additional expert opinion is added and practical recommendations are being formulated. The methodological principles for guideline development used by BRCF are described in detail in the charter below.

An overview of the different methodological aspects used for either guideline development or the development of a systematic review is given in a table in Appendix 3.

PROJECT INITIATION

Following a specific question from a certain Red Cross Service, the CEBaP performs a scoping review before a practice guideline or systematic review project is initiated.

The aim of this scoping review is to get a first idea about the content, quantity and quality of the available evidence. A quick search methodology is being used when performing a scoping review, using a specific search strategy and limited to one or two databases (The Cochrane Library, MEDLINE or Embase). In order to define the research question as good as possible, the input of the requesting Red Cross Service is included.

After finalising the scoping review, the steering group of the CEBaP decides if:

1) A practice guideline project will be initiated.
2) A systematic review will be developed.
3) No new project will be started up.

In order to make this choice, the following criteria are being taken into account:

- urgency
- potential impact
    - impact on practice, society
    - opportunity for a publication
    - intellectual property
    - quality of the body of evidence
- economic and financial impact on BRCF
- relevance for BRCF (does it fit into our core business, in our strategy?)

A systematic review will only be developed if:

- it can be used for policy change
- the answer to the question is not urgent
- there is a major chance that it will result in a peer-reviewed publication
- the quality of the body of evidence is moderate to high (preferably)

A diagram of the project initiation and project flow can be found in Appendix 4.

B. DEVELOPMENT OF AN EVIDENCE-BASED RECOMMENDATION AND PRACTICE GUIDELINE

CHOICE OF THE SUBJECT

In prioritizing a subject for a recommendation or practice guideline we take into account the following criteria:

- Urgency: high
- Potential impact of the practice guideline
    - Impact on practice, society
    - Opportunity for a publication
    - Intellectual property
- Economic and financial impact for BRCF
- Relevance for BRCF (does it fit into our core business, in our strategy?)

DEVELOPMENT OF AN EVIDENCE-BASED PRACTICE GUIDELINE ACCORDING TO AGREE II

The development of an evidence-based practice guideline by BRCF is based on AGREE II [2]. This instrument offers a framework for the development of qualitative guidelines in which the potential biases of guideline development have been adequately addressed. On the other hand, the recommendations are both internally and externally valid, and are feasible for practice. The assessment includes judgments about the methods used for developing the guidelines, the components of the final recommendations, and the factors that are linked to their uptake.

Based on AGREE II 6 domains are described in detail in all guidelines developed by BRCF:

1) SCOPE AND PURPOSE

A. The overall objective of the guideline. This deals with the potential health impact (e.g. prevention) for the target population and the expected outcome.

B. The PICO question(s) covered by the guideline. This contains the setting or context, the population, the intervention, the comparison and the outcome(s). For an extensive practice guideline, it is possible that in the PICO question describes not one but all effective interventions with relevance to a certain target group. No systematic literature search is started when the PICO concerns:

- A 'good practice point' or common sense
- The responsibility of professionals (such as a medical doctor or pharmacist)
- Interventions with only a long-term effect
- The practical organisation of activities
- Medico-legal aspects

- Anatomy or physiology

C. The target population (patients, public, etc.) to whom the guideline is meant to apply. This contains the age range, sex, clinical condition (if relevant), the severity/stage of the disease (if relevant), comorbidity (if relevant) and excluded populations (if relevant).

2) STAKEHOLDER INVOLVEMENT

A. The composition of the guideline development group including individuals from all relevant professional groups. The members of the guideline development group are all those involved at some stage of the development process. This item consists of the name, discipline/content expertise, institution or organisation, geographical location and a description of the member's role in the guideline development group. The guideline development group consists of: the members of the Steering Committee of the CEBaP, the staff members of the CEBaP that are responsible for collecting the evidence, the Red Cross service for whom the guideline is being developed and who is responsible for formulating the draft recommendations, and the expert panel that makes a trade-off between the quality of the evidence, benefits and harm and validates the final recommendations. The expert panel consists of a chairman, with expertise in evidence-based methodology and on the content of the project, and additional panel members, who at a minimum have expertise in the content of the project.

B. The views and preferences of the target population. The guideline development group receives information about the views and preferences of the target population from (1) the Red Cross service involved, which has expertise in the content or collects the necessary information (e.g. by composing a reading group or by interviewing the target population) (2) a literature search concerning the values, preferences and experiences of the target population and/or (3) a feedback round or pilot test. In addition the target population is represented in the guideline development group.

C. The target users of the guideline. This topic consists of a description of the target group and of how the practice guideline may be used by its target audience.

3) RIGOUR OF DEVELOPMENT

The 'rigour of development' domain consists of the literature search, the formulation of draft recommendations, the review of the guideline and the updating of the guideline.

3.1 SYSTEMATIC LITERATURE SEARCH

A. Systematic methods to search for evidence. We use a stepwise search strategy for collecting literature, which is described in detail. The search strategy consists of the sources of literature, the search terms, the use of methodological filters (if relevant) and the period from which articles are retrieved.

In AGREE II no detailed description is available of the methodology for the literature search. Therefore we based our methodology on that used by SIGN (Scottish Intercollegiate Guidelines Network) and NICE (National Institute for Health and Clinical Excellence) ('SIGN 50: A guideline developer's handbook', 'Section 6: Systematic literature review'; NICE: 'The guidelines manual 2009').

Sources:

- Guidelines: NGC (National Guideline Clearinghouse), GIN (Guidelines International Network) and MEDLINE (via PubMed interface).
- Systematic reviews: the Cochrane Database of Systematic Reviews (via Wiley interface), Database of Abstracts and Reviews of Effects (DARE - via Wiley interface), MEDLINE (via PubMed interface) and BestBETs (containing pragmatic systematic reviews).
- Experimental studies: Cochrane Central Register of Controlled Trials (via Wiley interface), MEDLINE (via PubMed interface), Embase (via Embase.com interface). Optionally: the choice of databases can be expanded if this is relevant for the search question (e.g. SPORTDiscus or PEDro for a practice guideline on the prevention of sports injuries).
- Observational studies: MEDLINE (via PubMed interface) and Embase (via Embase.com interface).

Search terms: The search terms (can) differ for every source (e.g. database of guidelines vs. database of individual studies), but will be described in detail for every source, in order to make the search reproducible (e.g. for the updating of the guideline). For the choice of the search terms, we pay attention to possible synonyms and we consult the MeSH thesaurus to identify possible related terms.

Methodological filters:

If possible we try to avoid the use of methodological filters.

- Guidelines: A non-validated filter for guidelines:
    - MEDLINE: "Guideline "[Publication Type] OR "Practice Guidelines as Topic"[Mesh] OR "Practice Guideline "[Publication Type] OR practice guideline*[TW]
    - Embase: 'practice guideline'/exp OR 'practice guideline' OR 'practice guidelines'
- Systematic reviews: based on a SIGN (Scottish Intercollegiate Guidelines Network - http://www.sign.ac.uk/) filter for systematic reviews
- Experimental studies: based on an EPOC (Effective Practice and Organisation of Care - http://epoc.cochrane.org/) filter for experimental studies

- Observational studies: a SIGN (Scottish Intercollegiate Guidelines Network - http://www.sign.ac.uk/) filter for observational studies; a validated filter described by Deville et al. [3] for diagnostic studies; a validated filter described by Wilczynski et al. for prognostic studies [4; 5].

Search period: We search for guidelines, systematic reviews, experimental studies and observational studies from the date of inception of the database until the date of the current search.

The **search strategy**: this takes into account that a practice guideline of BRCF consists of many different topics (> 40 topics) and therefore makes a compromise between the number of topics on the one hand and a reasonable time span for the development of the practice guideline on the other (cf SIGN methodology).

- The search strategy is developed by 1 reviewer and evaluated by a second reviewer.
- In a first step we search for guidelines as a source of systematic reviews and individual studies. In this step we also search for systematic reviews as a source of individual studies. In a next step we search for controlled experimental studies. In a third step we search for controlled observational studies. We only go to the next step of the search strategy if no evidence is found (cf. SIGN methodology) or if the evidence cannot be included based on the inclusion and exclusion criteria (see B).
- During the search for guidelines, systematic reviews, experimental studies and observational studies, additional references can be selected by checking the 20 related citations in PubMed and/or handsearching (e.g. in the reference list of an included reference). The additional references (guideline, systematic review, experimental study, observational study) are assessed with the inclusion and exclusion criteria.
- 1 reviewer for each topic selects and evaluates the evidence and describes the literature search in an 'evidence summary'. The evidence summary is made in a standard template (see Appendix 1). As an internal control the selection of evidence of a random selection of questions is performed periodically (cf. NICE methodology).

B. Inclusion and exclusion criteria for selecting evidence. These include criteria for the language, criteria for the content (population, intervention, outcomes and the context), and methodological criteria. No articles are selected if the intervention concerns:

- A 'good practice point' or common sense
- The responsibility of professionals (such as a medical doctor or pharmacist)
- Interventions with only a long-term effect
- The practical organisation of activities
- Medico-legal aspects
- Anatomy or physiology

If we search for risk factors, no articles are selected if the risk factor:

- does not precede the outcome
- is common sense
- is not modifiable (e.g. age, sex) (can be project-dependent)
- is no proximal risk factor (e.g. healthy diet and cardiovascular diseases, smoking and lung cancer) (can be project-dependent)
- is not valid for healthy people (can be project-dependent)

English, Dutch, French and German literature is selected.

No animal studies, *ex vivo* or *in vitro* studies are selected, but depending on the subject we can make an exception to this. Arguments for making this choice are described in the practice guideline.

Methodological criteria depend on the type of study design:

- A guideline: inclusion if the guideline gets an acceptable score in the 'rigour of development' domain of AGREE II.
- A systematic review: inclusion if the search strategy and selection criteria are clearly described and if at least the Cochrane Library, MEDLINE and Embase are searched.
- An experimental study: inclusion in case of one of the following study types: (quasi/non-) randomised controlled trial, controlled before and after study or controlled interrupted time series, and the data are available. These study types can be identified using the flowchart in Appendix 5.
- An observational study: inclusion in case of one of the following study types: cohort and case-control study, controlled before and after study or controlled interrupted time series, and the data are available. These study types can be identified using the flowchart in Appendix 5. Depending on the project and context it can be decided to include cross sectional, diagnostic and/or prognostic studies. For all study types the control group should be clearly described in the methods section.

C. <u>The strengths and limitations of the body of evidence.</u> The strengths and limitations of the body of evidence are being assessed using the GRADE approach [6].

The search strategy, inclusion and exclusion criteria, study characteristics, study findings and levels of evidence are described in an 'evidence summary'. Standard wording for the evidence conclusions in this evidence summary are described in Appendix 6.

3.2 FORMULATING RECOMMENDATIONS

A. The methods for formulating recommendations are clearly described. This topic consists of a clear description of the method that is used for formulating the recommendations, the outcomes of the recommendation development process and a description of how the process influenced the recommendations. A multidisciplinary expert panel is involved in formulating the final recommendations by informal consensus. If a consensus cannot be reached, the decision depends on the opinion of the majority by voting. The expert panel is responsible for reading through the whole guideline and for the assignment of the grades of recommendation (i.e. weak or strong).

B. The health benefits, side effects and risks. This topic consists of a description of the trade-off between benefits and harm, side effects or risks in formulating the final recommendations. The multidisciplinary expert panel makes this judgement during the assignment of the grades of recommendation. Therefore, the method of GRADE is being used.

C. Explicit link between the recommendations and the supporting evidence. This consists of a description of how the guideline development group links and uses the evidence to make informed recommendations. During the formulation of the final recommendations and the assignment of the grades of recommendation, the expert panel makes use of a table in which the corresponding evidence is presented for every draft recommendation.

3.3 REVIEW OF THE PRACTICE GUIDELINE

External review by experts. The external experts or peer reviewers are not involved in the guideline development group and can be target population representatives. Reviewers include experts in the clinical area as well as some methodological experts. For every practice guideline the group of reviewers consists of at least 1 expert on the content and 1 methodological expert who is preferably also an expert on the content.

The purpose and intent, methods for the external review, outcomes/information gathered from the external review and a description of how the information was used to inform the guideline, are described in the practice guideline. Furthermore, the name, discipline/content expertise and institution or organization are given.

3.4 UPDATE OF THE PRACTICE GUIDELINE

A procedure for updating the guideline. The practice guideline will be updated every 5 years, unless stated otherwise. To achieve this the literature search will be performed again from the end of the previous literature search until the start of the update.

4) CLARITY OF PRESENTATION

A. The recommendations are specific and unambiguous. This implies a precise description of which option is appropriate in which situation and in what population group. If evidence is inconsistent, this is described in the evidence summaries. In that case, the expert panel decides which option is the most appropriate.

B. The different options for management. The different possible interventions are presented in the guideline.

C. Key recommendations are easily identifiable.

5) APPLICABILITY

Depending on the type of project, context and target group, we can decide to complement the practice guideline with an implementation guide. This implementation guide can contain the following information: the facilitators and barriers to the application of the guideline, advice on how to put the recommendations into practice, the potential resource implications of applying the recommendations and monitoring and/or auditing criteria.

6) EDITORIAL INDEPENDENCE

A. Name of the funding body (BRCF) and statement that the content of the guideline is not influenced by the one that finances the project.

B. Statement of competing interests of guideline development group members.

PUBLICATION OF A PRACTICE GUIDELINE

In the publication of the practice guideline the topics mentioned above are preferably described in detail. In every case the methodology is described or we refer to a document containing the detailed methodology.

C. DEVELOPMENT OF A SYSTEMATIC REVIEW

CHOICE OF THE SUBJECT

During the development of evidence-based recommendations or practice guidelines it becomes clear where an up-to-date overview on the effectiveness of interventions is lacking. In these cases the Steering Committee of the CEBaP can decide to develop a systematic review.

In choosing a subject for a systematic review we take into account the following criteria:

- Urgency: low
- Potential impact of the practice guideline
    - Impact on practice, society
    - Opportunity for a publication in a peer-reviewed journal with a high impact factor (ISI Web of Knowledge – Journal Citation Report)
    - Intellectual property
    - Quality of the body of evidence
- Economic and Financial impact on BRCF
- Relevance for BRCF (does it fit into our core business, in our strategy?)

If more topics are suitable subjects of an evidence-based recommendation or practice guideline, we use the same criteria for prioritization of the question.

DEVELOPMENT OF A SYSTEMATIC REVIEW ACCORDING TO THE COCHRANE HANDBOOK

The systematic review provides an overview of the best available evidence collected by a literature search on a very specific topic or question. A systematic review can be used to inform policy makers. Making a trade-off between the estimated benefits, harm and the estimated costs, and thus making specific recommendations for an action, goes beyond the scope of a systematic review and is typically the task of guideline development group.

The systematic review will be included in the practice guidelines of BRCF when developing or updating a guideline, or can be used as a source of information for policymakers.

For the development of a systematic review we follow the methodology described in the Cochrane handbook (version 5.1.0) (http://www.cochrane-handbook.org/). Below, only the differences with the search strategy as described for practice guideline development will be highlighted.

1) DEFINING THE REVIEW QUESTION AND DEVELOPING CRITERIA FOR INCLUDING STUDIES

A focused research question is described in detail as a PICO question, in which the population, intervention, comparison and outcome are clearly indicated.

In addition, the types of study to be included as the source of evidence are clearly specified. In making this choice we consider a priori what study designs are likely to provide reliable data with which to address the objectives of the review.

2) SEARCHING FOR STUDIES

When making a Cochrane Review, we closely cooperate with the *Trials Search Co-ordinator* of the corresponding *Cochrane Review Group* in the search for studies.

At least the following databases are searched: the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE and Embase. Efforts are made to identify unpublished studies, conference abstracts, grey literature and ongoing trials.

We use a very sensitive search strategy. In the case of using methodological filters, the sensitive filters of Cochrane are used.

3) SELECTING STUDIES AND COLLECTING DATA

The study selection and data extraction are performed by at least 2 independent reviewers. A clear procedure for action is described in case of disagreement between the 2 reviewers.

Preferably the authors of studies are contacted when information in the study is missing.

4) ASSESSING RISK OF BIAS IN INCLUDED STUDIES

The following topics, which may be the source of risk of bias are assessed for every study: sequence generation (selection bias), allocation sequence concealment (selection bias), blinding of participants/personnel (performance bias), blinding of outcome assessment (detection bias), incomplete outcome data (attrition bias), selective outcome reporting (reporting bias) and other potential sources of bias.

To assess the quality and risk of bias in studies we use the 'Cochrane Collaboration's tool for assessing risk of bias', or GRADE for the assessment of "limitations in study design" in individual studies.

5) ANALYSING DATA AND UNDERTAKING META-ANALYSIS

If possible a meta-analysis will be conducted to statistically combine the data of several included studies. Heterogeneity between studies is determined statistically. In case of heterogeneity we search for an adequate way to address it (e.g. subgroup analysis).

If a meta-analysis is undertaken, we will also perform a sensitivity analysis to find out if the findings of the systematic review are sufficiently robust.

6) ADDRESSING REPORTING BIASES

If sufficient studies are included, a funnel plot can be used to search for *reporting bias*.

7) PRESENTING RESULTS AND SUMMARY OF FINDINGS TABLES

The characteristics of the studies are given in a table ('*Characteristics of included studies'*).
The results, i.e. the effect measures and confidence intervals, of the individual studies are shown in a '*summary of findings'* table. If possible, meta-analyses are generated and presented in a *forest plot*.
The quality of the individual studies is summarized in a '*quality of evidence'* table.

For every outcome of each intervention, a level of evidence is assigned to the body of evidence according to the GRADE approach and the evidence conclusion is formulated using a standard wording (Appendix 6).

8) INTERPRETING RESULTS AND DRAWING FINAL CONCLUSIONS FOR PRACTICE

Drawing final conclusions about the practical usefulness of an intervention entails making trade-offs between benefits, harm and costs. Finding this balance, and thus making specific recommendations, goes beyond a systematic review and is typically part of the development of a practice guideline.

In a Cochrane review a suggestion can be made on potential implications for practice by highlighting different possible actions. In addition, the implications for research can be discussed.

At this point of the review, it may be beneficial to call on the expertise of relevant external consultants to provide an independent appraisal of the quality and relevance of particular aspects of the review.

PUBLICATION OF A SYSTEMATIC REVIEW

For transparent reporting of the development of a systematic review, we use the PRISMA statements 2009 (http://www.prisma-statement.org/statement.htm). This is a 27-item checklist that aims to guarantee the quality of systematic reviews by clear and transparent reporting in a publication. The checklist can be found in Appendix 2.

D. KNOWLEDGE MANAGEMENT

EVIDENCE SUMMARY DATABASE

Each systematic literature search is documented in an evidence summary (Appendix 1) in the evidence summary database on the CEBaP project site. The evidence summary database is developed in Sharepoint and it ensures rapid document retrieval and offers the ability for future database enhancement.

REFERENCES

[1] De Buck E, Pauwels NS, Dieltjens T and Philippe Vandekerckhove. *Use of evidence-based practice in an aid organisation: a proposal to deal with the variety 1 in terminology and methodology.* Int J Evid Based Healthc 2014, 12:39-49

[2] Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. *AGREE II: advancing guideline development, reporting and evaluation in health care.* CMAJ 2010, 182(18):E839-E842.

[3] Deville WL, Buntinx F, Bouter LM, Montori VM, de Vet HC, van der Windt DA, et al. *Conducting systematic reviews of diagnostic studies: didactic guidelines.* BMC Med Res Methodol 2002, 2:9.

[4] Wilczynski NL, Haynes RB. *Optimal search strategies for detecting clinically sound prognostic studies in Embase: an analytic survey.* J Am Med Inform Assoc 2005, 12(4):481-5.

[5] Wilczynski NL, Haynes RB. *Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: an analytic survey.* BMC Med 2004, 2:23.

[6] Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. *Grading quality of 8 evidence and strength of recommendations.* BMJ 2004, 328(7454):1490.

APPENDIX 1: EVIDENCE SUMMARY AND GUIDELINE SUMMARY TEMPLATE

EVIDENCE SUMMARY

Topic	
Subtopic	
Intervention	
Question (PICO)	
Search Strategy	
Search date	
In/Exclusion criteria	

Characteristics of included studies

Author, year Country	Study design	Population	Comparison/Risk factor	Remarks

Synthesis of findings

Outcome	Comparison/Risk factor	Effect Size	#studies, # participants	Reference

Quality of evidence

Version 1: Quality of Experimental studies

Author, Year	Lack of allocation concealment	Lack of blinding	Incomplete accounting of outcome events	Selective outcome reporting	Other limitations

Version 2: Quality of Observational studies

Author, Year	Inappropriate eligibility criteria	Inappropriate methods for exposure and outcome variables	Not controlled for confounding	Incomplete or inadequate follow-up	Other limitations

Version 3: Quality of Diagnostic studies

Author, Year	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Could the patient flow have introduced bias?	Other limitations

Level of the body of evidence

	Initial grading, e.g. High	Downgrading due to
Limitations of study design	0	See table 'Quality of evidence'
Imprecision	0	
Inconsistency	0	
Indirectness	0	
Publication bias	0	
QUALITY (GRADE)	**Final grading, e.g. Low**	

Conclusion(s)	
Reference(s)	
Evidence used for	
Project	
Reviewer(s)	
URL	
Omleidings-URL2	

APPENDIX 2: PRISMA CHECKLIST

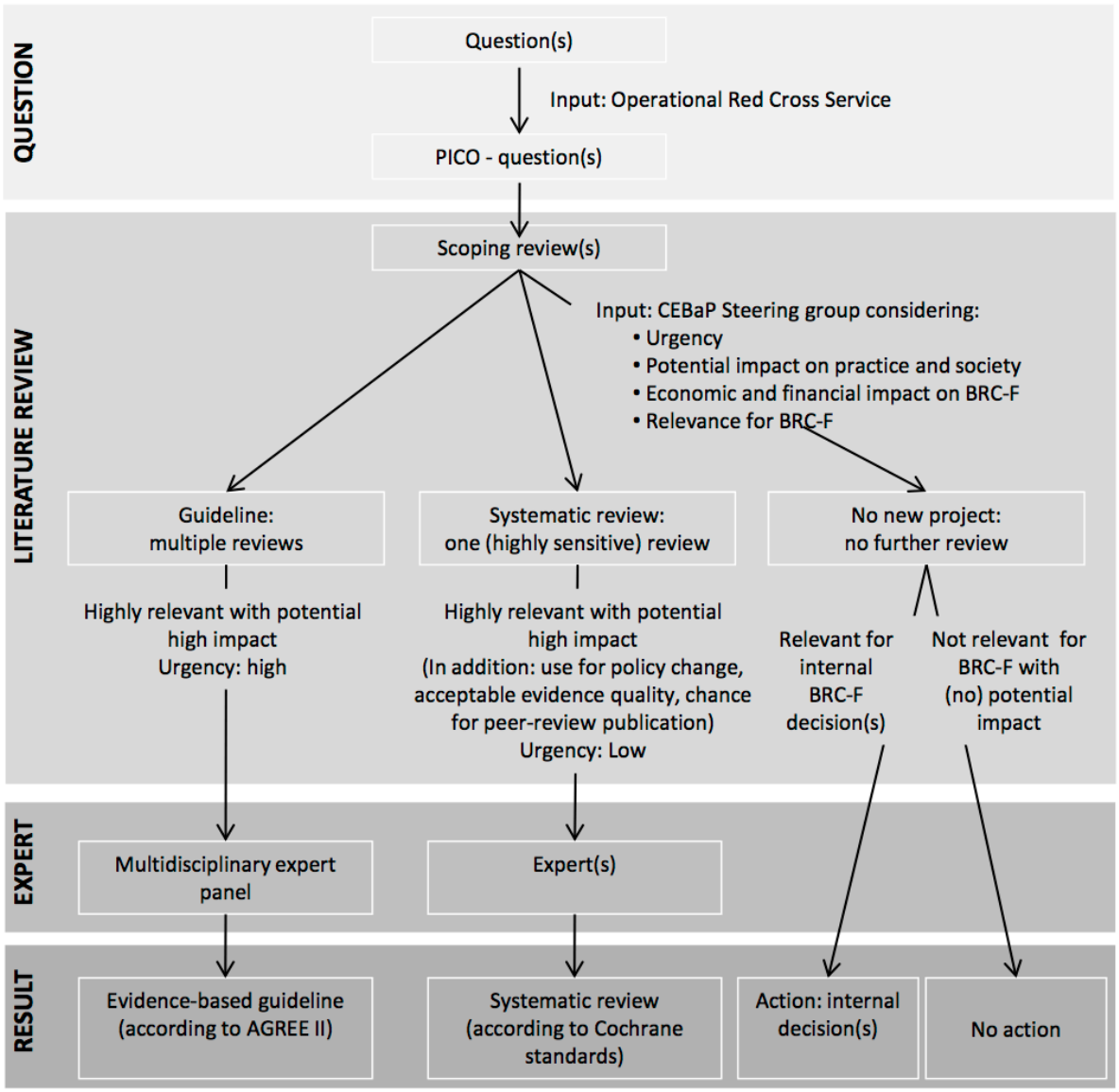
Section/topic	#	Checklist item
Title		
Title	1	Identify the report as a systematic review, meta-analysis, or both.
Abstract		
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.
Introduction		
Rationale	3	Describe the rationale for the review in the context of what is already known.
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).
Methods		
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative *evidence* (e.g., publication bias, selective reporting within studies).
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.
Results		
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).
Discussion		
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.
Funding		
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.

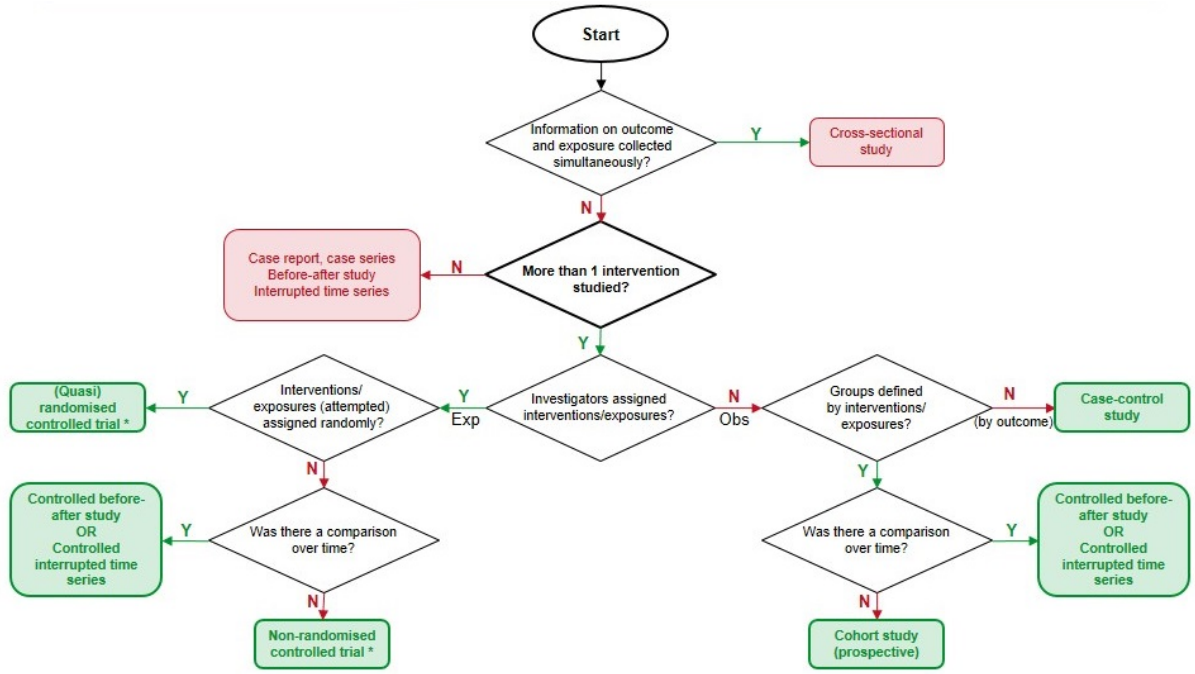
APPENDIX 3: TABLE OF METHODOLOGICAL PRINCIPLES OF GUIDELINE DEVELOPMENT VERSUS SYSTEMATIC REVIEW DEVELOPMENT (BRCF)

Methodology	Practice guideline	Systematic review
Number of reviewers: -building search strategy -literature search + selection articles -data extraction -quality assessment	 -1 reviewer -1 reviewer -1 reviewer -1 reviewer	 -2 independent reviewers -2 independent reviewers -2 independent reviewers -2 independent reviewers
Search formula	Specific search formula; use of methodological filters if necessary	Sensitive search formula
Databases	GIN, NGC, The Cochrane Library, BestBETs, MEDLINE, Embase. If necessary topic-specific databases can be added. No unpublished studies, conference abstracts, grey literature and ongoing trials will be identified.	At least: Central, MEDLINE, Embase Efforts are made to identify conference abstracts, grey literature and ongoing trials.
Selection criteria with respect to study design	Guidelines, systematic reviews, experimental studies and observational studies We only go to lower type of study design if no evidence is found or if the evidence cannot be included based on the inclusion and exclusion criteria.	Experimental studies or observational studies In making this choice it is considered a priori what study designs are likely to provide reliable data with which to address the objectives of the review.
Selection criteria with respect to content	Only the most direct and important factors are considered in the inclusion criteria	Depending on the research question
Quality assessment	Of the 'body of evidence'	Of each outcome separately
Meta-analysis	no	If possible
Involvement of experts	yes	yes
Formulating recommendations	yes	no
Assessment grades of recommendation	yes	no
External peer review	yes	no

APPENDIX 4: PROJECT INITIATION AND PROJECT FLOW



APPENDIX 5: FLOWCHART STUDY DESIGNS



* when all subjects received all interventions, “within subjects design” should be added to the evidence summary

APPENDIX 6: STANDARD WORDING OF EVIDENCE CONCLUSIONS

A) No Evidence

No studies on the effect of the intervention on the outcome, and fulfilling the selection criteria, were found: "*No relevant studies were identified using the above search strategy and criteria."*

A recommendation will not be included in the guideline, unless it is a Good Practice Point (GPP).

B) Conflicting evidence

The results of the different studies are not consistent and no explanation for the differences can be found: "*There is conflicting evidence from # experimental studies and/or # observational studies…"* (only name the conflicting studies). Detailed information can be given below, based on the sentences of the table below (except the first sentence with general conclusion).

A recommendation will not be included in the guideline.

C) Inconclusive evidence

In a limited number of cases no "overall conclusion" can be made, because there is no decisive argument to value the conclusions of one study more than those of another study. This is what we call "inconclusive evidence": *"There is inconclusive evidence concerning…"*.

D) (Limited) evidence (except conflicting evidence):

P- value	Level of evidence	*Imprecision* (GRADE methodology)	Evidence synthesis
< α (significant)	*A or B*	*No*	**<u>Intervention</u>** There is evidence in favour of [intervention]: . (In making this evidence conclusion, we place a higher value on … over …. [only in case conclusions are not straightforward]) It was shown that <intervention> resulted in a statistically significant increase/decrease of <outcome>, compared to <comparison> (<Author> <year>). Ψ Evidence is of high/moderate quality.
		Yes	**<u>Intervention</u>** There is limited evidence in favour of [intervention]: . (In making this evidence conclusion, we place a higher value on … over …. [only in case conclusions are not straightforward]) It was shown that <intervention> resulted in a statistically significant increase/decrease of <outcome>, compared to <comparison> (<Author> <year>). Ψ Evidence is of moderate quality and results cannot be considered precise due to limited sample size, lack of data and/or large variability of results.
	C or D	*No*	**<u>Intervention</u>** There is limited evidence in favour of [intervention]:. (In making this evidence conclusion, we place a higher value on … over …. [only in case conclusions are not straightforward]) It was shown that <intervention> resulted in a statistically significant increase/decrease of <outcome>, compared to <comparison> (<Author> <year>). Ψ Evidence is of low/very low quality. **<u>Risk factor</u>** There is limited evidence with benefit/harm for [risk factor]:. (In making this evidence conclusion, we place a higher value on … over …. [only in case conclusions are not straightforward]) It was shown that <risk factor> resulted in a statistically significant increased/decreased risk of <outcome>, compared to <comparison> (<Author> <year>). Ψ Evidence is of low/very low quality.

		Yes	**<u>Intervention</u>** There is limited evidence in favour of [intervention]:. (In making this evidence conclusion, we place a higher value on … over …. [only in case conclusions are not straightforward]) It was shown that <intervention> resulted in a statistically significant increase/decrease of <outcome>, compared to <comparison> (<Author> <year>). Ψ Evidence is of low/very low quality and results cannot be considered precise due to limited sample size, lack of data and/or large variability of results. **<u>Risk factor</u>** There is limited evidence with benefit/harm for <risk factor>:. (In making this evidence conclusion, we place a higher value on … over …. [only in case conclusions are not straightforward]) It was shown that <risk factor> resulted in a statistically significant increased/decreased risk of <outcome>, compared to <comparison> (<Author> <year>). Ψ Evidence is of low/very low quality and results cannot be considered precise due to limited sample size, lack of data and/or large variability of results.

Ψ Repeat this sentence for every "subconclusion"; in case of only one study, do not repeat <author> <year>

P- value	Level of evidence	*Imprecision* (GRADE methodology)	Evidence synthesis
> α (not-significant)	*A or B*	*No*	**Intervention** There is evidence showing no difference between intervention and control:. (In making this evidence conclusion, we place a higher value on … over …. [only in case conclusions are not straightforward]) It was shown that <intervention> did not result in a statistically significant difference of <outcome>, compared to <comparison> (<Author> <year>). Ψ Evidence is of high/moderate quality.
		Yes	**Intervention** There is limited evidence neither in favour of the intervention nor the control:. (In making this evidence conclusion, we place a higher value on … over …. [only in case conclusions are not straightforward]) A statistically significant increase/decrease of <outcome>, using <intervention> compared to <comparison>, could not be demonstrated (<Author> <year>). Ψ Evidence is of moderate quality and results of this study/these studies are imprecise due to limited sample size, lack of data and/or large variability of results.
	C or D	*No*	**Intervention** There is limited evidence showing no difference between intervention and control:. (In making this evidence conclusion, we place a higher value on … over …. [only in case conclusions are not straightforward]) It was shown that <intervention> did not result in a statistically significant increase/decrease of <outcome>, compared to <comparison> (<Author> <year>). Ψ Evidence is of low/very low quality. **Risk factor** There is limited evidence showing no correlation between the <risk factor> and the <outcome>:. (In making this evidence conclusion, we place a higher value on … over …. [only in case conclusions are not straightforward]) It was shown that <risk factor> did not result in a statistically significant increased/decreased risk of <outcome>, compared to <comparison> (<Author>

			<year>). Ψ Evidence is of low/very low quality.
		Yes	**<u>Intervention</u>** There is limited evidence neither in favour of the intervention nor the control:. (In making this evidence conclusion, we place a higher value on … over …. [only in case conclusions are not straightforward]) A statistically significant increase/decrease of<outcome>, using <intervention> compared to <comparison>, could not be demonstrated (<Author> <year>). Ψ Evidence is of low/very low quality and results of this study/these studies are imprecise due to limited sample size, lack of data and/or large variability of results. **<u>Risk factor</u>** There is limited evidence concerning the risk of <outcome> in presence of/in case of/when <risk factor> (compared to <control>). (In making this evidence conclusion, we place a higher value on … over …. [only in case conclusions are not straightforward]) A statistically significant increased/decreased risk of <outcome> in case/presence of <risk factor> compared to <comparison> could not be demonstrated (<Author> <year>). Ψ Evidence is of low/very low quality and results of this study/these studies are imprecise due to limited sample size, lack of data and/or large variability of results.

Ψ Repeat this sentence for every "subconclusion"; in case of only one study, do not repeat <author> <year>

Study Design	**Initial quality of evidence**	**Lower if**	**Higher if**	**Quality of a body of evidence**
Experimental studies	High	Limitations in design -1 Serious -2 Very serious Inconsistency -1 Serious -2 Very serious Indirectness -1 Serious -2 Very serious Imprecision -1 Serious -2 Very serious Publication Bias -1 Likely -2 Very likely	Large effect +1 Large +2 Very large Dose response +1 Evidence of a gradient All plausible residual confounding +1 Would reduce a demonstrated effect +1 Would suggest a spurious effect if no effect was observed	High Moderate Low Very low
Observational studies	Low			

Balshem H., Helfand M., Schünemann H., Oxman A., Kunz R., Brozek J., Vist G., Falck-Ytter Y., Meerpohl J., Norris S., Guyatt G., GRADE guidelines: 3. Rating the quality of evidence, Journal of Clinical Epidemiology (2011) 64, 401-406.