How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: a five-step approach (part 1/3)

Ghislaine A.P.G. van Mastrigt, Mickaël Hiligsmann, Jacobus J.C. Arts, Pieter H. Broos, Jos Kleijnen, Silvia M.A.A. Evers & Marian H.J.M. Majoie


To link to this article: http://dx.doi.org/10.1080/14737167.2016.1246960

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: a five-step approach (part 1/3)

Ghislaine A.P.G. van Mastrigt, Mickaël Hiligsmann, Jacobus J.C. Arts, Pieter H. Broos, Jos Kleijnen, Silvia M. A.A. Evers and Marian H.J.M. Majoie

ABSTRACT
Introduction: Systematic reviews of economic evaluations are useful for synthesizing economic evidence about health interventions and for informing evidence-based decisions. Areas covered: As there is no detailed description of the methods for performing a systematic review of economic evidence, this paper aims to provide an overview of state-of-the-art methodology. This is laid out in a 5-step approach, as follows: step 1) initiating a systematic review; step 2) identifying (full) economic evaluations; step 3) data extraction, risk of bias and transferability assessment; step 4) reporting results; step 5) discussion and interpretation of findings. Expert commentary: The paper aims to help inexperienced reviewers and clinical practice guideline developers, but also to be a resource for experts in the field who want to check on current methodological developments.

1. Introduction

Constant and worldwide tension between clinical opportunities and financial possibilities necessitates a critical approach to health-care expenditure. An increasing number of economic evaluations (EEs) are being performed, but many decisions are still based on effectiveness data. When those decisions are highly focused on evidence-based methodologies that consider only one dimension of relevant evidence (i.e. whether the intervention works), this may contribute to inefficient policy and practice decisions [1]. Systematic reviews (SRs) are the reference standard for synthesizing data because of their methodological rigor [2]. Systematic reviews of economic evaluations (SR-EEs) are useful for synthesizing economic evidence about health interventions. Using the Kleijnen Systematic Reviews (KSR) Evidence database which aims to include all SRs of the medical literature from 2015 onwards, we estimate that 35–50 SR-EEs are being performed yearly. SR-EEs can be categorized roughly into three groups: (1) multipurpose reviews, (2) reviews for informing the development of clinical practice guidelines (CPGs), and/or (3) reviews for developing decision analytical models. Both multipurpose SR-EEs and SR-EEs for guideline development aim to synthesize and critically appraise existing EEs of a health-care intervention or disease area in order to inform policy decisions. In other words, they provide information on what is known, what remains unknown, and can reveal the knowledge gaps about EE for that specific topic [3]. In addition, recommendations are written based on the findings of SR-EEs which have been performed for guideline development. In the third group, SR-EE can be used to support the development of a decision analytical model. While the authors acknowledge the merit of these models and their contribution in promoting evidence-based decisions in health care, the guidance in this paper primarily covers only the first two types of SR categories.

The methods for SR-EEs overlap in part and share similarities with the methods of effectiveness reviews [4,5]. Various leading organizations and collaborations of experts, including the Cochrane Collaboration, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR), the Centre for Reviews and Dissemination of the University of York (CRD), the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group, Guidelines International Network (GIN), the Agency for Healthcare Research and Quality in the United States (AHRQ),...
1.1. The five-step approach for preparing an SR-EE

Based on a five-step approach, the methods for preparing SR-EEs for informing evidence-based health-care decisions are presented. The consecutive steps are described in detail in the paragraphs corresponding to the headings noted in Figure 1. This paper provides an overview of all five steps and detailed information on Step 1) how to initiate SR-EEs, Step 4) report results, and Step 5) discuss and interpret results. As the topics on identifying (full) EEs (Step 2) and performing data extraction (including risk of bias and transferability assessment – Step 3) in SR-EEs need more explanation, a detailed overview of these topics is provided in two separate papers [18, 19].

1.2. Basic knowledge of economic evaluations

Two types of EE can be distinguished: full EEs and partial EEs. Full EEs are the preferred type of EE for both multipurpose SR-EEs and SR-EEs for CPG development [20]. Full EEs are defined as studies in which (1) two or more alternative interventions are compared, and (2) both costs and effects (consequences and benefits) of the compared treatments are taken into account [20]. Full EEs are regarded as the optimal type of EE. Overall, EEs are specifically designed to inform policy decisions, and this is the main purpose of performing both types of SR-EEs. In a partial EE, the two noted requirements (comparison two treatments and measurement of both costs and consequences) are not met.

Three types of full EEs can be distinguished: cost-effectiveness analysis (CEA), cost-utility analysis (CUA), and cost–benefit analysis (CBA). In a CEA, costs are expressed in monetary units and effects in natural units, such as reduction in blood pressure, or decrease in the number of myocardial infarctions. In a CUA, a special type of CEA, the costs and consequences (comparison two treatments and measurement of both costs and consequences) are not met.

To our knowledge, there is no up-to-date, detailed, and practical overview of the consecutive steps to follow in preparing an SR-EE. This paper aims to help inexperienced reviewers and CPG developers, but also to assist experts in the field who want to check on current methodological developments. This guidance paper is also specifically written to guide CPG developers in performing preparing SR-EEs. Moreover, the proposed overall review approach can be helpful for all researchers and students who want to improve and standardize their approach to SRs.
For most countries, partial EEs are not the recommended analytical perspective [6], although these studies might be included in SR-EEs in a certain field of research where there is a lack of knowledge on a specific decision topic—e.g., when an SR-EE is performed to inform CPG development and no relevant full EEs have been identified for that specific topic. Five types of partial EEs can be distinguished: outcome description, cost description, cost-outcome description, effectiveness (or efficacy) evaluation, and cost analysis [20]. For further information on the details of the methodology of EEs, we refer to the standard textbooks [20,21].

Furthermore, it is also important to know which analytical approach has been used. First, EEs can be based on a modeling study, in which various (literature) sources of data have been used to build a model, and this is used to estimate the cost-effectiveness of a particular health-care intervention. Second, in trial-based EEs, the data for EEs are collected alongside data from a clinical trial (e.g., a randomized controlled trial, pragmatic trial, or quasi experiment) [20]. Third, other approaches are also possible, such as an analysis of real-world data based on patient registries, or it is also possible to use a combination of analytical approaches. Budget impact analysis (BIA) is often mentioned in relation to CPG development and is required for informing formulary approval or reimbursement decisions [22,23]. BIA can be either freestanding or part of a comprehensive economic assessment alongside EEs [23]; it addresses the expected changes in expenditure of a health-care system after a new intervention has been implemented [23]. Details on budget impact analyses are beyond the scope of this paper [20,22,23].

As already mentioned in the introduction, the findings of SR-EEs can also be used for the development of an economic decision model, but as the methods for this type of SR-EEs differ from SR-EEs used for multipurpose and CPG development, they are beyond the scope of this paper. For further reading on this topic, we refer to Shemilt et al. [4].
information specialist (librarian) [7,9,12,15] should be included. When conducting an SR-EE for CPG development, it is highly recommended that a patient representative \([9,12,15,16]\) and a person with experience in CPG development also be included to complete the team \([9]\).

Conflicts of interest or disclosure of interests should be handled in an appropriate way \([30]\). These can be defined as ‘a divergence between an individual’s private interests and his or her professional obligations such that an independent observer might reasonably question whether the individual’s professional actions or decisions are motivated by personal gain, such as direct financial, academic advancement, clinical revenue streams, or community standing \([31]\)’. GIN has published a recent paper on the principles of managing conflicts of interest in guideline development \([30]\).

### 2.2. Step 1.2: identify and define a relevant topic

When no up-to-date SR with evidence of acceptable quality on the topic of interest is available, an SR-EE should be initiated. For both multipurpose SRs and CPGs, a scoping review can be prepared to get a sense of existing economic evidence on a specific topic \([13]\). An important first step in preparing such SR is to perform a scoping review. The purpose of this scoping review is to obtain more knowledge on the topic of interest and to investigate whether any recent similar SRs in the international literature have been performed or are underway \([32,33]\). The results of this review can be used for writing the protocol of the SR-EE. The main reason for performing SR-EEs is to support practice (e.g. guideline development) and policy, and to direct new research efforts \([34]\).

When the final goal of the SR-EEs is to write recommendations for CPGs, we identified two important issues which should be considered before starting the SR.

First, due to time and financial restrictions, it is often not possible to add economic evidence to every clinical question in every newly developed or existing CPG. Accordingly, priorities need to be set, for which consultation with different stakeholders (e.g. health-care providers, patients, payers, purchasers, policymakers, and guideline developers) is highly recommended at the beginning of the project \([25]\). This can be done by (independently prioritizing) the relevant research topics.

Second, based on our own experience, it is preferable to start SR-EEs for a specific CPG whenever there is already an SR on clinical effectiveness being finished. The results of the effectiveness review can be used by stakeholders to prioritize the topics for the SR-EEs. For instance, in an effectiveness SR, only relevant treatment options for clinical practice were included. For the SR-EEs, this means that studies not investigating these treatment arms can be excluded from the SR-EEs. An SR-EE can be done before the effectiveness SR but is less optimal in terms of efficiency.

### 2.3. Step 1.3: write a protocol of SR-EEs

The next step is to write a protocol containing all methods for the planned SR-EEs. The preparation of a protocol is an essential component of the SR process because it ensures that the SR is carefully designed and that what is planned is explicitly documented before the review starts. In other words, it supports consistent conduct and accountability of the project team, integrity of the research, and transparency of the applied SR methods \([27]\). We recommend using the PRISMA-P checklist, a guideline to help prepare protocols for SRs which provides a minimum set of items to be included in the protocol \([27]\). Using the P (population or patient), I (intervention or comparator), C (control/comparator/comparison), O (outcome) mnemonic, eligibility criteria and the purpose of the review can be specified. See Table 1 for an example of inclusion and exclusion criteria for an SR-EE for incorporating EEs in the Dutch CPG for the treatment of epilepsy in the Netherlands \([35]\). It should be noted that although complex interventions such as clinical pathways or complete care chains can pose difficulties in interpreting the study findings, they need to be taken into account in SR-EEs.

The following protocol items can be included: review title, timescale (starting date and date of completion), project team details (names, contact details, funding sources, COIs), background, purpose, review question(s) and searches performed, data extraction (selection and coding), transferability assessment, risk of bias assessment, strategy for data synthesis and reporting \([27]\). More specifically, the background should include the following:

- A definition of concepts, including a detailed explanation of the intervention whose costs/cost-effectiveness are being examined. An ‘event pathway diagram’, is recommended, as it provides a systematic and explicit method for presenting the pathway of events and distinct resource implications and the associated outcome values \([12,13]\);
- The review topic and motivations \([13]\);
- Introduction to the health-care setting, population, and outcomes of interest \([13]\);
- A motivation for the review, with the reference to current debates among policy makers and/or clinicians relating to the topic, to gaps in the evidence base and user needs \([13]\).

### 2.4. Step 1.4: publish protocol of SR-EEs

Publication of the study protocol of SR-EEs is becoming a more common practice and is highly recommended as it avoids unnecessary duplication of studies, reduces publication biases, and increases availability and accessibility of a priori methods \([27,38–40]\). It also increases efficiency, as others are able to check what your work in progress is, and avoid duplication \([13,41]\). A peer review process is part of the protocol by Cochrane \([12]\) and also part of the publication policy of the journal Systematic Reviews. This journal is an open access journal and encompasses all aspects of the design, conduct, and reporting of SRs \([42]\). Registration of SR protocols can be done at either the International Prospective Register of Systematic Review (PROSPERO) \([32]\) or, when an SR follows the Cochrane protocol, at the Cochrane website \([12]\). PROSPERO is an international open access database launched
in 2011 by the CRD of the University of York which prospectively registers SRs. Currently, it contains data for 10,000 SRs in health and social care. The Cochrane Collaboration is a global independent network of researchers, professionals, patients, and careers which provides high-quality SRs for making health decisions. There are also several journals which publish SR protocols – for instance, Systematic Reviews and British Medical Journal (BMJ) Open.

3. Step 2: identifying (full) EEs

3.1. Step 2.1: select relevant data sources

Searching for relevant EEs is an iterative process, and several data sources need to be checked in order to identify all of them [43]. The main sources for identifying full EEs are general databases (or basic databases), such as Medline (freely accessible on the Internet through PubMed) [44,45], Embase [45], Econlit, and Web of Science. The two databases specifically developed for EEs of health-care interventions, the NHS EED and the HEED, are no longer publishing; NHS EED can be used for searches of full EEs up to March 2015, but HEED is no longer accessible. Furthermore, depending on the topic, other more specific databases can be selected: e.g. especially for SR-EEs for CPG development (guideline clearinghouse, NICE). Finally, a third group of databases can be used for identifying studies. These so-called ‘optional databases and web pages’ may hold additional information relevant for a more comprehensive SR-EEs [18]. For example, for grey literature and for conference proceedings, the ISPOR [46] and the Cochrane Colloquium [47].

Searching for relevant citations (studies) can also be done by checking the references in known publications [7,12,48] or by searching for additional studies that are cited in articles known to be relevant (such as Web of Science) [11]. In addition, it can be helpful to contact the authors or experts cited in a study (HEALTHECON-ALL) [49] to obtain additional information [12]. This could be, for instance, data on resource use if only costs are reported, or more information on variance (standard deviations or minimum or maximum scores) as only measures of central tendency are reported.

3.2. Step 2.2: development of search strategies

Developing new search strategies (strings of search terms) is time consuming and highly dependent on the reviewer’s experience; it is estimated to take up to 20 h for an experienced reviewer [50]. However, it is not always necessary to develop new search strings for every new SR-EE. It is recommended to use existing validated search filters as much as possible; these offer an optimal balance between sensitivity
and specificity in relation to the aims of the SR [11,44]. In general, a successful search strategy is regarded as one that delivers a manageable amount of references with a searcher-specified balance of sensitivity and specificity [51].

The InterTASC Information Specialists’ Sub-Group (ISSG) provides a list of such filters, which is updated monthly [52]. The appendices of Cochrane SRs or other high-quality SRs are also good sources for filters. A checklist (the PRESS checklist [53] or the CADTH checklist [54]) can be used to ensure quality when peer reviewing search strategies. These two checklists have been developed to identify and assess the impact of errors in the elements of electronic search strategies associated with accuracy and completeness of the evidence base.

For researchers interested in designing their own search strategy, we suggest including all the relevant concepts of every research question can be identified using the PICO scheme [37] (see also eligibility criteria, step 1.3). Different concepts and different filters can be combined into one search strategy with the Boolean operator ‘AND’. For each concept, it is advised to include a wide range of free-text terms separated by the Boolean operator OR. Free-text terms may include synonyms, acronyms and abbreviations, spelling variants, old and new terminology, brand and generic medicine names, and lay and medical terminology [11]. Furthermore, it is recommended that possible truncations be used as much as possible (e.g. for both the concepts “cost-effectiveness and costs use cost”), as well as wildcards (e.g. to include both women and woman use wom?n) and proximity operators (e.g. for near use NEAR) [12]. Finally, try to restrict your search as little as possible; accordingly, it is not recommended to restrict for language [12,36] or to choose too narrow a time frame.

3.3. Step 2.3: perform searches
As a rule of thumb, information specialists and experienced reviewers find it feasible to screen between 100 and 150 abstracts within 1 h [55]; for inexperienced researchers, this would be a lower number. Clear documentation of both manually or electronically preformed searches is essential for the reproducibility of your study findings [56]. Details on the searches performed in the databases and websites (e.g. dates covered, database host systems and database coverage dates, concepts and search strategies used, number of records [hits] retrieved, details of any supplementary searches undertaken, including the rationale and language restrictions) should be systematically collected [7,11,12,15] and added to the appendices of publications. Bibliographic details of the references identified and the pdfs of the papers can be merged, using reference software (e.g. EndNote or RefWorks) [11]. Duplicates need to be removed by means of the reference software and also by hand, as reference software is not always reliable [57].

3.4. Step 2.4: selection of studies
Screening of potential relevant studies needs to be conducted in two stages [7,10]. First, the records need to be screened on review title and abstract. Studies should be selected based on the eligibility criteria stated in the published protocol (steps 1.3 and 1.4). Second, the full text records must be screened for compliance with eligibility criteria. Ideally, all steps critical for study selection (steps 2.3 and 2.4) and also those for data extraction (step 3) should be done by two reviewers independently [7,12,15]. However, this is not always feasible due to financial or time constraints. An alternative method could be for reviewers to perform a double check by consultation in case of doubt. Any discrepancies between the two reviewers should be resolved by consensus [7,10,11].

Whenever there are multiple publications of the same study, they need to be linked [7,10,12]. This can be done by reporting in the results section – when discussing the flowchart of study selection – information on the various records of the same study. To increase the transparency of the excluded studies for the SR-EEs, a list of excluded studies that appear to meet eligibility criteria but were nevertheless excluded can be provided in the appendices [7,12]. This list needs to contain bibliographic details of the excluded studies and the reason for exclusion [7,10,12]. A flowchart of the PRISMA statement on study inclusion can be used to show all details of the selection process in a systematic way [7,10,58]. Step 2 will be discussed in more detail in Thielen et al. [18].

4. Step 3: data extraction, risk of bias assessment, and transferability

4.1. Step 3.1 data extraction
The next step in the process is the extraction of all relevant data from the included studies. A data extraction sheet is developed based on the study design, study objective, and predefined outcomes as described in the study protocol (steps 1.3 and 1.4). There are several examples of data extraction forms available in the literature [7,12,59], typically containing many common items. A data extraction sheet includes both the general study characteristics (for instance, author, year of publication, type of intervention, control treatment, type of EE), details on study methods and outcomes (e.g. resource use, costs, effects, measurement, valuation methods, incremental cost-effectiveness ratios). For modeling studies, special attention needs to be paid to aspects such as model structure, key assumptions, input data values, and uncertainty analyses [60]. Disaggregated presentation of the results (including resource use in natural as well as monetary units) is advised in order to facilitate interpretation of the results [7]. It is highly recommended to pilot this sheet on user-friendliness and completeness, using a few sample studies [7,10,12]. Furthermore, we recommend using a picklist to record the various response options, and to decrease possible ambiguity when using several reviewers.

4.2. Step 3.2: risk of bias assessment
The next step in the development of SR-EEs is critical appraisal, or a quality check of the included studies. In other words, are there any possible biases which may impact study outcomes. A ‘bias’ can be described as the difference between the true value (of the population) and the observed value (that of
the sample) from any other course, as a sampling variation [61]. An example of a bias is, for instance, using a perspective that is too narrow (a hospital perspective instead of societal perspective) for the EE analyses, with the consequence that not all relevant costs and outcomes are taken into account. A list of other possible biases for both trial and model-based EEs can be found elsewhere [62]. Risk of bias assessment for an SR-EE means that the chosen study design, methods, assumptions, models, and possible biases are critically appraised [63]. This needs to be done in a way that is transparent and fully supported by available evidence, the strength of which is made easily accessible to any critical reader [64]. It is important to keep in mind that the quality of EEs can be only as good as that of the trials on which they are based [65]. The choice of a specific risk of bias assessment checklist depends on the purpose of the SR. The same list can be used for both multipurpose SR-EEs and SR-EEs for CPG development; however, in addition to those checklists, for CPG development, a special checklist needs to be used. Furthermore, the initial study characteristics (type of EEs, analytical approach) are also critical in choosing a specific checklist. The following checklists are recommended, based on best practices [7,10–12,66].

1. For multipurpose SR-EEs and SR-EEs for CPG development, if you want to use the same checklist for the appraisal of both trial-based and model-based EEs, the CHEC-extended [67–69] or BMJ checklist [70] are the preferred options. In cases where one is specifically interested in model-based EEs and if the expected number of included studies is low (e.g., <10 studies), based on a pragmatic decision the Phillips checklist [71] is recommended. In cases in which the number of included model-based EEs is high (e.g., >10 studies; also a pragmatic decision), the ISPOR checklist [71] is the preferred list.

2. To incorporate economic evidence in developing CPGs, the GRADE approach [72] should be used. For trial-based EEs, the ‘GRADE evidence profile’ (a specific form of balance sheet) and in ‘Summary of Findings tables’ are the preferred way for summarizing the data [73]. For model-based EEs, the GRADE approach is not applicable, and therefore, the NICE checklist can be used [11,74].

The above-mentioned recommendations are general ones, but it is important to realize that there can be aspects which are also important to take into account when choosing one or more of these checklists. For instance, the time available for preparing the review, the audience and type of publication (report, paper or CPG), and the experience of the reviewers also need to be considered. For a complete overview of all available risk of bias assessment checklists for EEs, see the paper of Wijnen et al. [19].

4.3. Step 3.3: Transferability assessment

Transferability can be defined as the ability to extrapolate results obtained from one setting or context to another [75]. It can be an issue in interpreting the results of an EE when the study has been performed in a country other than decision country. For instance, if a study has been done in the United States, the clinical setting maybe not be transferable to Dutch care. Accordingly, the transferability of study findings needs to be assessed [19]. Nine different checklists are available for this [19]. Recommendations regarding the use of transferability checklists are based on the same criteria used for selecting the risk of bias assessment checklists (step 3.2). Taking these criteria into account, the Welte checklist [78] is recommended for both multipurpose SR-EEs and SR-EEs for informing CPGs (for more details, see Step 5.1 factor six). The transferability issue will also be discussed in Section 5.2.1 of this paper. In addition, it is advisable to check for country-specific guidelines for EEs [6], as they provide background information on what the main differences between countries regarding guidance for designing EEs are.

Step 3 will be discussed in more detail in Wijnen et al. [19], where specific examples will be provided on data extraction (trial-based and model-based EEs), risk of bias, and transferability assessment.

5. Step 4: Reporting results

5.1. Step 4.1 Result section and data synthesis

The next step of the review process is the presentation of the findings, including data syntheses, in a result section (step 4.1).

5.2. Results and data synthesis for multipurpose SR-EEs

When the results of SR-EEs are presented, the reader should be able to understand the results and the major conclusions. Accordingly, all relevant findings of the studies need to be presented in detail in summary tables and also summarized in the text [1,7,11]. Ranking the studies by means of a league table based on the costs per QALY can be very useful [79]. A Dominance Ranking Matrix, a simple classification system for summarizing and interpreting the results of various EEs in an SR-EE can also be used for the same purpose [13]. In addition, in order to make comparison of different study results possible, it is preferred to convert all different currencies reported to a one common currency (e.g., US dollar, Euro) and to use the same year as a reference [12]. There is a free web-based tool [80] developed by the Campbell and Cochrane Economics Methods Group (CCEMG) and the Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre) which automatically adjusts estimates for costs and price year, taking purchasing power parities between countries into account. Furthermore, as mentioned before, disaggregated presentation of the results (resource use and costs) is advised in order to make proper judgments about the transferability of the evidence [7,81]. There are currently no agreed-upon methods for pooling combined estimates of cost-effectiveness (e.g., incremental cost-effectiveness, cost–utility, or cost–benefit ratios), extracted from multiple EEs, using meta-analysis or other quantitative synthesis methods [12]. Based on this, and due to possible various sources of heterogeneity (patients, study design, and outcomes) [7], pooling of different
EEs is not recommended. Graphic presentation of the data can be used if common metrics (e.g. costs and QALYs) for each study are applied. Examples of this are cost-effectiveness planes [24] or a hierarchical matrix, which summarizes the findings of EEs for interventions versus comparators [82].

5.3. Results and data synthesis for SR-EEs in guideline development

For trial-based EEs, the GRADE approach is recommended, as explained in step 3.2 (Risk of Bias assessment). The findings of EEs (incremental costs/effects, cost-effectiveness ratios and uncertainty) can be added to the GRADE evidence profiles, in the same table which describes quality assessments and effectiveness results [8,17]. For model-based EEs, using the same tables as for multipurpose SR-EEs (discussed in the previous section) is advised.

6. Step 5: discussion and interpretation of results

6.1. Step 5.1 discussion section for both multipurpose SR-EEs and SR-EEs in guideline development

In the discussion section of both multipurpose SR-EEs and SR-EEs for guideline development, it is recommended that the following general topics [16,29,38] be addressed: summary of evidence, heterogeneity, study strengths and limitations, time frame for updating the review, general conclusions, source of funding, conflict of interest, and discussion of findings in relation to earlier SRs. Other specific topics related to the discussion of EEs should also be addressed, e.g. transferability, generalizability, and implementation problems or budget impact. For a detailed description of the different topics, see Table 2.

6.2. Step 5.2 development of recommendations for SR-EEs in guideline development

The final step of the five-step approach is to develop recommendations based on the retrieved, extracted, and appraised results and syntheses of the results (steps 3.1, 3.2, and step 4.1). This is applicable only for SR-EEs which are being performed to inform CPG development. For the development of recommendations, the following general approach (steps 5.1–5.2–5.3) is advised.

6.3. Step 5.2.1 discuss SR-EE findings with project team

First, discuss the most important results (from steps 3.1, 3.2, and 4.1) with the multidisciplinary project team (composed in step 1.1.). More specifically, for each identified EE in the SR, the following seven factors can be discussed:

1. For all EEs, it is recommended that the quality of the study be discussed. Four categories of quality can be distinguished: High, Moderate, Low and Very Low [26]. Consider EEs for inclusion only if they have moderate or high quality.

(2) Discuss whether the findings of the study show that the experimental or new intervention is cost-effective. This is the case when both costs and effects of the new intervention compare favorably with both the costs and effects of the control intervention (i.e. costs are lower and effects are better). However, as EEs are always based on both cost and effects, a new intervention which is more expensive but results in higher outcomes in comparison with an existing intervention can also be considered cost-effective. This depends on the threshold values being used (see also section on country-specific guidelines for EEs).

(3) It is also recommended that the variability and uncertainty of studies be discussed. For instance, do all the conclusions from the base case analyses still hold after the sensitivity analyses have been performed?

(4) Take into account the balance between health benefits, side effects, and risks [26]. For instance, when a new treatment regime with medication is more effective but serious side effects are more frequent in comparison with an existing one, is this still the preferred therapy?

(5) Discuss whether the study results are generalizable, probably generalizable or not generalizable to the setting of the CPG (see also Table 2).

(6) Discuss whether the study results are transferable, probably transferable or not transferable in the context of the CPG. Determine whether the following three statements are true or false: (1) The relevant technology is comparable to the one that shall be used in the decision country, (2) The comparator is relevant to the one that is relevant in the decision country, and (3) The study is of an acceptable quality (outcome general factor 1). If one of these three statements can be termed ‘false’, the EE cannot be used for CPG development, as it is not transferable to the guideline context [78]. For further assessment of the transferability, check the specific knockout criteria defined by Welte, such as the study perspective, discount rate, and medical cost approach used [78].

(7) Discuss whether the incorporation of the EEs into a specific CPG poses any implementation problems, e.g. if the new intervention has a large impact on the total budget because the disease is highly prevalent and therefore many patients will get the new treatment option. The following categories can be used: unlikely, likely and will not pose implementation problems.

To structure the discussion regarding these seven factors, the form presented in Table 3 can be used for every study. The last part of the form, on the overall conclusion of the findings, the overall strength of evidence, and possible research gaps, can be filled out based on the seven factors discussed.

6.4. Step 5.2.2 formulate and present recommendations

The conclusions of the SR-EEs for every research question may not always be directly applicable as recommendations in CPGs, although, together with the scientific evidence from the effect studies and supplementary considerations, they
must lead to the final recommendations. The supplementary considerations which besides all the scientific evidence can be taken into account are more specific; expert opinion, patient preferences, costs (budget impact or practicability), clinical relevance, legal consequences, availability of facilities, organization of care, and security. In a paragraph of the CPG, there should be a discussion of how the results of the body of evidence and considerations are assessed and weighed. Several methods can be used to make the final decisions on the formulation of the recommendations for CPG development [9,16]: voting system, informal consensus (e.g. expert opinions), and formal consensus techniques (e.g. Delphi, nominal group techniques). Furthermore, in general, a recommendation should provide a concrete and precise description of which option is appropriate in which situation and in what population group, as well as being informed by the body of evidence [16]. Furthermore, the considerations should be described explicitly and systematically, using arguments both for and against the options for the care being examined. The advantages and disadvantages of an intervention should be described, as well as the alternatives. In addition, the most important results should be presented in tables (see step 3.2: risk of bias assessment) or explained in the text. In this way, the (guideline) user is able to identify easily which components of the body of evidence are relevant for each recommendation [16].

Based on results and discussion of all studies on a specific research question, a strong or a conditional recommendation can be formulated [26]. A strong recommendation formulated from EEs is based on high-quality or moderate-quality studies which are transferable, generalizable, and can be implemented into the intended setting. In addition, these EEs cannot show more than acceptable variability and uncertainty in the study results. If all these factors can be applied, the recommendations based on EEs can be used for CPG development. However, in some instances, evidence is not always clear, and there may be uncertainty about the best care option(s). If there is uncertainty in the evidence, this should be stated in the CPG [16], and in those circumstances conditional recommendations should be defined. No recommendations for CPG should be based on low-quality studies, which are probably not transferable or not generalizable, and cannot be implemented into the intended setting if they pose high variability and/or high levels of uncertainty.

When there is no evidence available on a specific research question, this needs to be stated in the text – for example, ‘no recommendation can be made based on insufficient evidence’ or ‘this recommendation was formulated based on expert opinion, as no literature on EEs was available on the topic’. In the considerations section, the arguments for giving these conditional recommendations should be provided.

In general, the presentation of the recommendations should be clear [16]; specifically, users should be able to find the most relevant recommendations easily. The recommendations must address the main research question(s) that have been covered by the guideline and can be identified and highlighted in different ways [16].

### 6.5. Step 5.2.3 external review before publication

Finally, a standard part of the guideline development process is an external review of the newly developed guideline by stakeholders before actual publication. These reviewers should not be involved in the development of the guideline and can for instance be experts in the clinical area, methodological or HTA experts, or represent the target population (patients,
More research is also needed on the most optimal way to incorporate model-EEs in CPGs. How to incorporate model-EEs in CPGs is important to increase transparency and standardize the methodology used for SR-Ees; there is a constant need for more detailed and up-to-date guidance.

The following topics should receive extra attention in the coming years.

- An extension of the PRISMA statement for SRs of EEs should be developed.
- A consensus-based and comprehensive checklist should be developed for quality assessment of both model-based and trial-based EEs, especially for incorporating economic evidence into CPGs.
- More guidance is needed on how to incorporate model-based EEs in a systematic way into CPGs, using the GRADE approach.
- A consensus-based and comprehensive checklist should be developed which the project team can use to discuss the findings of every EE, so recommendations based on economic evidence can be more easily implemented into CPGs.
- Alternatives for specialist economic databases (NHS EED or HEED databases) must be considered. As these specialist databases are no longer available (HEED) and up-to-date (NHS EED), research is needed on what bibliographic databases are preferred for searching out and identifying EEs.
- More research is also needed on the most optimal way to identify EEs (focusing more on appropriate indexing of EEs or developing more sensitive search strategies).
- To increase the transparency of study findings, open access publication of papers and journals should be promoted by governments and universities. In addition, online publication of study protocols, databases of study findings, and models should be encouraged. EEs with negative study findings should also be published with open access.

### Table 3. Decision table to support the development of recommendations for SR-Ees for guideline development, including an example of a study considered for the Dutch National Guideline for epilepsy.

<table>
<thead>
<tr>
<th>Study (author, year)</th>
<th>Ketogenic diet</th>
<th>Children and adolescents with intractable epilepsy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor</strong></td>
<td><strong>Decision</strong></td>
<td><strong>Comments</strong></td>
</tr>
<tr>
<td>Based on the quality appraisal, what is quality of the EE?[^a]</td>
<td>High quality[^b]</td>
<td>High quality</td>
</tr>
<tr>
<td>Is the new intervention cost-effective?</td>
<td>Yes or No</td>
<td>No</td>
</tr>
<tr>
<td>Are the outcomes of the EE uncertain and/or variable?[^c]</td>
<td>Acceptable variability/uncertainty</td>
<td>High variability/uncertainty</td>
</tr>
<tr>
<td>Balance between health benefits, side effects, and risk[^d]</td>
<td>Health benefits clearly outweigh side effects, and risk</td>
<td>No</td>
</tr>
<tr>
<td><strong>Transferability of the study results</strong></td>
<td>Transferable</td>
<td>Transferable</td>
</tr>
<tr>
<td><strong>Generalizability of the study results.</strong></td>
<td>Generalizable</td>
<td>Generalizable</td>
</tr>
<tr>
<td>Implementation problems</td>
<td>Unlikely</td>
<td>Likely</td>
</tr>
<tr>
<td>Overall conclusion of the findings</td>
<td>A high quality Dutch trial-based EE showing nondominance of the intervention in comparison with care as usual. Findings are on based on interim analysis at 4 months.</td>
<td></td>
</tr>
<tr>
<td>Overall strength of evidence[^e]</td>
<td>Not enough evidence</td>
<td></td>
</tr>
<tr>
<td>Research gaps</td>
<td>Longer follow-up time is needed, &gt;12 months</td>
<td></td>
</tr>
</tbody>
</table>

[^a]: High quality: further research is very unlikely to change confidence in the estimate of effect; Moderate quality: further research is likely to have an important impact on confidence and is likely to change the estimate of effect; Low quality: further research is very likely to have an important impact on confidence and is likely to change the estimate of effect; Very low quality: any estimate of effect is very uncertain. ^[^b]: High quality based on interim analysis at 4 months. |
<table>
<thead>
<tr>
<th>Type of review</th>
<th>General recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both review types</td>
<td>Follow the 5-step approach. For basic knowledge of EE methods look at Drummond et al. [20]. Full EEs are the preferred type of EEs for both multipurpose SR-EEs and SR-EEs for CPG development [16].</td>
</tr>
<tr>
<td>SR-EEs for CPG development</td>
<td>- Basic knowledge on CPG development. Check ISPOR website for country-specific EE guidelines [8]. Use AGREE II tool for CPG development [16]. Use PRISMA statement for SRs [2]. Check CHEERS for reporting EEs [29].</td>
</tr>
<tr>
<td>Both review types</td>
<td>- Minimize biases. Ideally, all steps critical for study selection (2.3 and 2.4) but also those for data extraction and appraisal (3.1, 3.2 and 3.3) should be performed independently by two reviewers.</td>
</tr>
<tr>
<td>Reporting of SR-EEs</td>
<td>- Use PRISMA-P checklist [27]. Publish the protocol on the PROSPERO website [32] or, if there is a Cochrane review, on the Cochrane Website [33].</td>
</tr>
</tbody>
</table>

**Step 1: Initiating a SR-EEs**

| Multipurpose SR-EEs    | 1.1 Compose project team. Compose a multidisciplinary project team to include the following expertise: clinical, SR methods, quantitative methods and health technology assessment methods, and library and information science. |
| SR-EEs for CPG development | 1.1 Compose project team. Compose a multidisciplinary project team to include the following expertise: clinical, SR methods, quantitative methods and health technology assessment methods, CPG development, library and information science, and patient and public views. |

| Both review types      | 1.1 Manage any conflicts of interests. When preparing SR-EEs, conflicts of interests should be handled in an appropriate way. |
| Multipurpose SR-EEs    | 1.2 Identify and define a relevant topic or research questions. Perform a scoping review to identify the most relevant research questions. |
| SR-EEs for CPG development | 1.2 Identify and define a relevant topic or research questions. Consult different stakeholders and perform a scoping review to identify the most relevant research questions. Start the SR-EE after the effectiveness SR is finished. |

| Both review types      | 1.3 Write a protocol of SR-EEs. Write a protocol of the SR-EEs by using the PRISMA-P checklist [27]. |
| Both review types      | 1.4 Publish protocol of SR-EEs. Publish the protocol on the PROSPERO website [32] or, if there is a Cochrane review, on the Cochrane Website [33]. |

**Step 2: Identifying full EEs**

| Both review types      | 2.1 Select relevant data sources. Use list of databases to select those relevant for SR-EEs (see Appendix Thiel et al. [18]). General databases. Select at least Medline (freely accessible on the internet via PubMed), Embase, Econlit, and Web of Science. Additional searches can be performed in NHS EED (although this has not been updated since March 2015). |

(Continued)
Table 4. (Continued).

<table>
<thead>
<tr>
<th>Type of review</th>
<th>General recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Specific and optional databases</strong></td>
</tr>
<tr>
<td>Both review types</td>
<td>Select specific databases according to your topic (if applicable). Search optional database(s) and websites for HTA reports and conference proceedings.</td>
</tr>
<tr>
<td></td>
<td><strong>Grey literature</strong></td>
</tr>
<tr>
<td></td>
<td>Consider including grey literature. Search in known publications for relevant citations. Make use of citation searching (i.e. identify articles that have cited a set of relevant articles); use Web of Science or Google scholar for this.</td>
</tr>
<tr>
<td></td>
<td><strong>Citation searching</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Search terms</strong></td>
</tr>
<tr>
<td>Both review types</td>
<td>Make use of the PICO (Patient, Intervention, Comparator, Outcome) scheme to find relevant search terms for all important concepts/aspects of the research question. Include a wide range of free-text terms. Use proximity operators (e.g. NEAR) if possible. Employ thesauri and synonyms. Use truncation options for your search term. For English, use British or American spelling.</td>
</tr>
<tr>
<td></td>
<td><strong>Search filters</strong></td>
</tr>
<tr>
<td>Both review types</td>
<td>Determine whether you want to use a more sensitive or precise search filter. SRs will profit from sensitive filters because precise filters will miss some articles. Look for search filters that filter for publication types (e.g. economic or trial publications). Choose validated filters as much as possible. Check The InterTASC Information Specialists’ Sub-Group website for validated search filters [85]. Appendices of Cochrane SRs or other high quality SRs are also good sources for filters.</td>
</tr>
<tr>
<td></td>
<td><strong>Combine search terms and filters with Boolean (AND, OR, NOT) operators</strong></td>
</tr>
<tr>
<td>Both review types</td>
<td>Carefully consider on what, and if at all, you want to restrict your search results. It is not recommended to restrict on language or to choose too narrow a time frame.</td>
</tr>
<tr>
<td></td>
<td><strong>Document the search process</strong></td>
</tr>
<tr>
<td>Both review types</td>
<td>Document and report all steps of the search, including the complete search strategy for every database.</td>
</tr>
<tr>
<td></td>
<td><strong>Handle references</strong></td>
</tr>
<tr>
<td>Both review types</td>
<td>Use bibliographic software to keep track of downloaded references and publications. De-duplicate the downloaded records by using a reference management software program.</td>
</tr>
<tr>
<td></td>
<td><strong>Screen references</strong></td>
</tr>
<tr>
<td>Both review types</td>
<td>Ideally, two reviewers should screen the references independently. Screen titles and abstracts of the downloaded studies based on the eligibility criteria that were set in the protocol.</td>
</tr>
<tr>
<td></td>
<td><strong>Step 3: Data extraction, risk of bias assessment and transferability</strong></td>
</tr>
<tr>
<td>Both review types</td>
<td>Adapt the data extraction sheet for every specific study; include all relevant items from the list in Table 1 in paper by Wijnen et al. [19]. Include the selected risk of bias and transferability checklists in the data extraction sheet. It is recommended that a picklist be used to choose the different response options.</td>
</tr>
<tr>
<td>Type of review</td>
<td>General recommendations</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td><strong>3.2 Risk of bias assessment</strong></td>
<td>It is recommended that a few studies (i.e. two or three) should be used to pilot the assessment among multiple raters, after which discrepancies should be discussed to ensure a more uniform assessment strategy.</td>
</tr>
<tr>
<td>Both review types</td>
<td>For trial-based and model-based EEs use the CHEC-extended [67–69] or BMJ checklist [86]. In cases where one is specifically interested in model-based EEs and if the expected number of included studies is low (e.g. &lt;10 studies: pragmatic decision), the Phillips checklist [71] could be used. In cases in which the number of included model-based EEs is high (e.g. &gt;10 studies; pragmatic decision), the ISPOR checklist is likely to be more practical for reviewing purposes [66].</td>
</tr>
<tr>
<td>SR-EEs for CPG development</td>
<td>Full EEs should be preferred over partial EEs at all times. In the absence of full EEs, partial EEs may represent important intermediate stages in our understanding of the costs and consequences of health services programs and therefore might be convenient. For trial-based EEs use the GRADE approach [8]. For model-based use EEs use the NICE checklist [11].</td>
</tr>
<tr>
<td>Both review types</td>
<td>The Welte checklist [55] is recommended for all trial-based and model-based EEs.</td>
</tr>
<tr>
<td><strong>3.3 Transferability assessment</strong></td>
<td></td>
</tr>
<tr>
<td>Both review types</td>
<td>Step 4: Reporting of results Convert all different currencies reported to a common currency (e.g. US dollar, Euro) and use the same year as a reference [80]. Graphic presentation of the data can be used if common metrics (e.g. costs and QALYs) for each study are applied [13,83]. Due to the methodological and study-specific heterogeneity issues of EEs, a meta-analysis is not recommended. SR-EEs for CPG development For trial-based EEs: use GRADE evidence profiles and Summary of Findings tables [87]. For model-based EEs: see recommendations for multipurpose SR-EEs. Multipurpose SR-EEs For multipurpose SR-EEs (trial-based and model-based): the findings of step 3 can be presented in self-developed summary tables and also summarized in the text.</td>
</tr>
<tr>
<td>Both review types</td>
<td>Step 5: Discussion and interpretation of results General topics for SR-EEs: summary of evidence, heterogeneity, study limitations, study strengths, time frame for update of review, previous SR-EE findings in relation to current SR-EE findings, general conclusions, recommendations for further research, source of funding and conflict of interest. Specific topics for SR-EEs: transferability, generalizability and implementation problems or budget impact. SR-EEs for CPG development Discuss the following seven factors for each study with the project team: study quality, cost-effectiveness, variability and uncertainty, balance between health benefits, side effects and risk, generalizability, transferability and expected implementation problems. Fill in Table 3 to record the group’s discussion for each EE on all seven factors. SR-EEs for CPG development In addition to the seven factors discussed, important other considerations which need to be taken into account before recommendations for CPG can be written are: expert opinions, patient preferences, costs (budget impact or practicability), clinical relevance, legal consequences, availability of facilities, organization of care and security. Use one of the following methods for structuring the discussion: voting system, informal consensus (e.g. expert opinions), and formal consensus techniques (e.g. Delphi, nominal group techniques) to formulate the final recommendations. When there is no evidence available on a specific research question, a statement needs to be added on this. In general, the presentation of the recommendations should also be clear. More specifically, users should be able to find the most relevant recommendations easily. Use one of the following methods for structuring the discussion: voting system, informal consensus (e.g. expert opinions), and formal consensus techniques (e.g. Delphi, nominal group techniques) to formulate the final recommendations. When there is no evidence available on a specific research question, a statement needs to be added on this. In general, the presentation of the recommendations should also be clear. More specifically, users should be able to find the most relevant recommendations easily. SR-EEs for CPG development The CPG should be reviewed externally before it is published.</td>
</tr>
</tbody>
</table>
Key issues

- We recommend using the PRISMA statement for SR preparation [2], the AGREE II tool for CPG development [16] and CHEERS for EE reporting [29] whenever possible.
- We advise following the 5-step approach in preparing both multi-purpose and CPG SR-EEs (see Table 4 for a summary of the most important topics).
- We recommend composing a project team, identifying and defining a relevant research question, to write and publish a protocol for all the SR-EEs.
- We recommend using at least Medline, Embase, Econlit, and the Web of Science databases for study selection. Additional searches can be performed in NHS EED (although this specialist database has not been updated since March 2015). For the complete list of databases see Thielen et al. [18].
- We advise using validated search filters as much as possible. They can be found on the ISSG website [85], in the appendices of Cochrane SRs or other high quality SRs.
- We recommend using a data-extraction sheet, performing risk of bias assessment and a check on transferability when preparing a SR-EE.
- We advise discussing the following aspects for every study before providing recommendations: study quality, transferability, variability and uncertainty, generalizability, expected implementation problems and the balance between health benefits, side effects and risk.
- An extension of the PRISMA statement should be developed especially for systematic reviews of EEs.

Funding

This project received financial support from the Dutch Society of Medical Specialists, (project number: 329666498, Utrecht) and from ZonMw (the Dutch Organization for Health Research and Development, project number: 836022001, The Hague).

Declaration of interest

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

ORCID

Ghislaine A.P.G. van Mastrigt http://orcid.org/0000-0001-8053-5512
Mickaël Hiligsmann http://orcid.org/0000-0003-4274-9258
Jos Kleijnen http://orcid.org/0000-0003-2787-7091
Silvia M.A.A. Evers http://orcid.org/0000-0003-1026-570X

References

Papers of special note have been highlighted as either of interest (+) or of considerable interest (+++) to readers.

20. Wijnen et al. describe the second step of the SR-EE review process in detail. More specific, how to select the relevant databases, how to develop search strategies, how to perform the searches and how to select the studies.
data-extraction sheet, how to select an appropriate risk of bias assessment checklist and how to check on transferability.


32. International prospective registry of systematic reviews. [http://www.crd.york.ac.uk/PROSPERO/]

33. Cochrane collaboration. [www.cochrane.org]


• Synopsis: Liberati et al. discuss the development of the PRISMA statement which consists of a 27-item checklist and a four-phase flow diagram. The checklist includes items deemed essential for transparent reporting of a systematic review


43. Critical Reviews Advisory Group. Introduction to systematic reviews. [http://www.shef.ac.uk/scharr]


46. [www.ispor.org/]

47. [https://colloquium.cochrane.org/]


49. HEALTHCON-ALL. [https://www.jiscmail.ac.uk/cgi-bin/wrmin? A1=ind16048&l=HEALTHCON-ALL]


52. The InterTASC Information Specialists’ Sub-Group Search Filter Resource [https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home]


Appendix 1. Examples of global and country-specific sources for clinical practice development.

<table>
<thead>
<tr>
<th>Global</th>
<th>Name organization</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Guidelines International Network (GIN)</td>
<td><a href="http://www.g-i-n.net">http://www.g-i-n.net</a></td>
</tr>
<tr>
<td></td>
<td>World Health Organization (WHO)</td>
<td><a href="http://apps.who.int/iris/bitstream/10665/75146/1/9789241548441_eng.pdf">http://apps.who.int/iris/bitstream/10665/75146/1/9789241548441_eng.pdf</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Country specific</th>
<th>Name organization</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Netherlands</td>
<td>Knowledge Institute of Medical Specialists</td>
<td><a href="http://www.kennisinstituut.nl">http://www.kennisinstituut.nl</a></td>
</tr>
<tr>
<td>Scotland</td>
<td>Scotland Intercollegiate Guidelines Network (SIGN)</td>
<td><a href="http://www.sign.ac.uk/guidelines/fulltext/50/">http://www.sign.ac.uk/guidelines/fulltext/50/</a></td>
</tr>
<tr>
<td>Canada</td>
<td>Cancer Care Ontario Program in evidence-based care handbook</td>
<td><a href="https://www.cancercare.on.ca/about/programs/pceb/">https://www.cancercare.on.ca/about/programs/pceb/</a></td>
</tr>
</tbody>
</table>