How to prepare a systematic review of economic evaluations for clinical practice guidelines: database selection and search strategy development (part 2/3)

FW Thielen, GAPG Van Mastrigt, LT Burgers, WM Bramer, HJM Majoie, SMAA Evers, J Kleijnen

Expert Review of Pharmacoeconomics & Outcomes Research. 2016;16(6):705-721

ABSTRACT

Introduction: This article is part of the series "How to prepare a systematic review of economic evaluations (EES) for informing evidence-based healthcare decisions", in which a five-step approach is proposed.

Areas covered: This paper focuses on the selection of relevant databases and developing a search strategy for detecting EEs, as well as on how to perform the search and how to extract relevant data from retrieved records.

Expert commentary: Thus far, little has been published on how to conduct systematic review EEs. Moreover, reliable sources of information, such as the Health Economic Evaluation Database, have ceased to publish updates. Researchers are thus left without authoritative guidance on how to conduct SR-EEs. Together with van Mastrigt et al. we seek to fill this gap.

INTRODUCTION

To support their decisions in health care, policy and decision makers need reliable information on the cost-effectiveness of health care interventions.[56] Systematic reviews of economic evaluations (SR-EEs) are a source of this information.[57] However, although these reviews have become increasingly important, little has been published on how to perform SR-EEs.[58] Without such guidance, those who wish to perform SR-EEs are left with practice guidance and recommendations that focus solely on medical efficacy research, which is usually concerned only superficially – if at all – with economic outcomes.

The vast amount of publications and their widely differing quality, together with subjective components that may guide a searcher's decision, call for standardized methods.[59] Therefore, a carefully planned strategy is essential when a thoroughly conducted SR is the goal.[60] Moreover, SRs should be reproducible, verifiable, efficient, and accountable.[57,61,62]

With a five-step approach on how to perform SR-EEs of health-care interventions, van Mastrigt and colleagues make a first attempt to fill the gap that has occurred in the absence of both guidance and reliable and comprehensive economic databases.[30] Their goal is to pave the way in establishing future guidance for SR-EEs. In the meantime, their approach can be used as a preliminary manual for performing SR-EEs in a sound scientific way. Their guidance aids users in employing efficient and transparent methods, which are central to any SR.[57] Just as for part 1/3 of this paper series, this article's main target audience is developers of clinical practice guidelines (CPGs) who need a point of reference on how to perform SR-EEs. Similarly, it can be a helpful tool for researchers in health technology assessment, systematic reviewers, and for students who seek to prepare an SR-EE. To illustrate the case, we will discuss our theoretical considerations alongside a recent example of an SR-EE that was part of developing a CPG for the treatment of epilepsy in The Netherlands.[63]

BACKGROUND

Typically, evidence for a CPG is gathered by systematically reviewing publications that are concerned with the effectiveness of different treatment options.[64] In addition, it has become increasingly acknowledged that CPGs should also entail economic evidence.[65,66(p7),67] This can be done in two, not necessarily independent, ways: (1) an SR and critical summary of the economic evidence already published is undertaken or (2) a decision analytic model is built to model economic effects.[57] This article will focus solely on the former approach.

In general, most steps of an SR-EE involve the same stages that are needed to conduct an SR of evidence for clinical effectiveness.[57] More specifically, any SR-EE will be based on the same two-stage process that has become the established standard for SRs of effects,[57] namely: (1) developing a search strategy and (2) applying the search strategy to a set of specified

databases.[68] However, some methods of SR-EEs diverge significantly as economic outcomes replace effectiveness or safety outcomes that would be detected in SRs.[57] As a result, database selection as well as the identification of search terms and filters differs. However, guidance on how to extend a search strategy and what databases to use when seeking to incorporate EEs is scarce, fragmented, or not applicable to all cases. In this article, we will present solutions for overcoming these issues, based on published guidance in the field and our experience.

THE FIVE-STEP APPROACH FOR PREPARING AN SR-EE

Following van Mastrigt's approach for conducting SR-EEs, the first step is to compose a multidisciplinary project team, frame the study, prioritize the topics, and write and publish the protocol. With regard to the subsequent steps, it should be noted that adding a medical information specialist or librarian to the search team adds great value to the quality of the searches.[69] In the second step, EEs need to be identified; this includes (1) selecting relevant databases, (2) developing an adequate search strategy, (3) performing the searches, and (4) selecting the relevant studies. This article will provide a more detailed description of these four parts of the second step, while step 3 is described by Wijnen et al.[70] in more detail. An overview of all other steps and a detailed description of steps 1, 4, and 5 can be found in van Mastrigt et al.[30] For an overview of the five-step approach, see Figure 1.

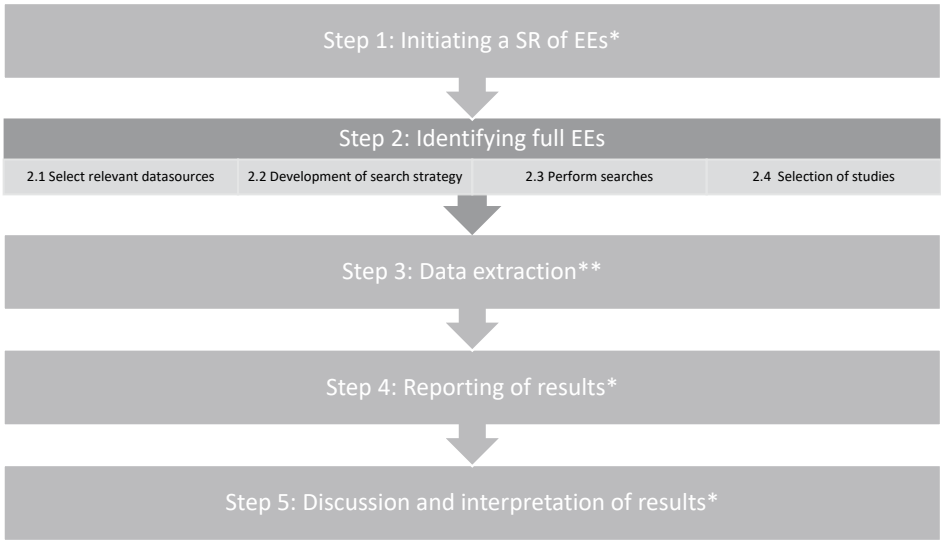

Figure 1 - An overview of the 5-step approach for preparing a systematic review of economic evaluations to inform evidence-based decisions. *Described in detail by van Mastrigt et al.,[30] **Described in detail by Wijnen et al.[70]

STEP 2.1 OF THE OVERALL FRAMEWORK: SELECTION OF RELEVANT DATA SOURCES

Until recently, a large part of EEs in health care could be detected by searching databases that specifically focus on these evaluations, such as the U.K. National Health Service Economic Evaluation Database (NHS EED) and the Health Economic Evaluation Database (HEED). However, HEED ceased publication at the end of 2014 and is no longer accessible for searches.[71] And, although still accessible through the Cochrane Library and the Centre for Reviews and Dissemination (CRD) website, the NHS EED has not been updated since March 2015.[72]

Many databases can be accessed via different search providers and platforms, and these pose varying requirements for a search strategy. Most end users will access well-known standard biomedical databases such as MEDLINE or Embase [1]. Apart from the question of whether all EEs are indexed in these databases, records can be indexed inconsistently, and there is no uniform interpretation of the definition of EEs [3]. In addition to electronic bibliographic databases, other resources such as gray literature, research registries, or web pages may contain useful information. Also, registries of unpublished studies can be searched, and researchers can be contacted for additional data.

No database is comprehensive enough to cover all relevant published research.[73] Therefore, the general consensus for effectiveness is that at least several databases need to be searched for a comprehensive result.[74–78] Guidelines for SRs recommend searching at least two bibliographic databases,[79,80] although there is no agreed-on standard for how many should be searched.[31] As the number of searched databases increases, database bias (referred to as the probability that the index of a record in a specific database is dependent on its results) and potential language bias can be reduced.[81] Which databases should be selected for a review depends heavily on the study objectives,[31] and there is no consensus about this either.[82] Being aware of how each interface for searching databases works is essential, since search results might well vary if the same database is searched through different interfaces (e.g. searching MEDLINE via PubMed or via OVID).[82]

Electronic databases for searching EEs

Backed by an extensive amount of evidence,[83–92] Mathes et al.[93] recommend searching at least MEDLINE and Embase for SR-EEs. In addition, they suggest searching one health economic database, such as HEED or NHS EED. Also, the Cochrane Handbook[92] and the manual for developing the National Institute for Health and Care Excellence (NICE) guidelines,[64] together with the Campbell and Cochrane Economics Methods Group (CCEMG),[94] emphasize the use of the NHS EED on their website when searching for economic evidence for SRs. However, as HEED is no longer available and the NHS EED is no longer updated, this advice is obsolete.

Gray literature

Gray literature (i.e. technical reports, studies, or essays that are unpublished, have restricted distribution, and are therefore rarely included in bibliographic retrieval systems)[95] has the potential to add valuable information to an SR-EE, especially when little is known about the topic under study. Although finding and including gray literature is particularly time-consuming and difficult, it is regarded as necessary for minimizing bias in reviews.[96] When possibly including gray trials, Hopewell et al.[96] recommend contacting the authors of these trials for more information. Examples of missing information could, for instance, be values for the standard deviation or variance when only the mean or median is reported.

The CRD health technology assessment database identifies gray literature.[97]

Citation searching

In citation searching, the reviewers search for articles that have cited a set of relevant articles which have already been detected.[31] For example, this can be done on the Science Citation Index Expanded™ (Thomson Reuters, United States),[98] via the Web of Science™. Citation searching can also include reference checking. Here, the reviewers can scan the reference lists of useful records previously identified to see if they refer to as yet unknown articles.

Database selection: a practical example
Wijnen et al.[63] sought to present an overview of published and ongoing full EEs of all health-care interventions for patients with epilepsy. The main search was conducted in March 2015. The following databases were searched: MEDLINE (via PubMed), Embase, NHS EED, EconLit, Web of Science, the Cost-Effectiveness Analysis Registry, the Cochrane Library of Systematic Reviews, the CRD Database of Abstracts of Reviews of Effects (DARE), and CRD Health Technology Assessment Database. With the first five databases, 'basic databases' were selected. Since the search was conducted up until March 2015, it can be expected that NHS EED was exhaustively searched. All other databases are classified as 'optional database' in this publication. It seems worthwhile mentioning that DARE also stopped its service in 2014.

Classification of databases

We classified several databases and websites into three categories, based on their ability to detect EEs in health care; these three categories are (1) basic, (2) specific, and (3) optional. For a complete but non-exhaustive list, see online Appendix 2A. The choice of databases is independent of whether the purpose is to conduct a multipurpose review or to develop a new CPG.

1) Basic databases: We refer to 'basic databases' as those that are recommended for use in any case when performing SR-EEs. Using a well-constructed search strategy, most relevant EEs will be detected.
2) Specific databases: For an SR on a topic for which a specific database is available, we recommend using it. Specific databases are those that provide information primarily in a particular research field. An SR on a mental health topic for instance would benefit

from searches performed on PsycINFO (American Psychological Association, United States).[99,100]

3) Optional databases and websites: Under the category of 'optional databases,' we grouped databases and web pages that may hold additional information relevant for a more comprehensive SR. For example, optional databases will identify Health Technology Assessment (HTA) reports (Canadian Agency for Drugs and Technologies in Health [CADTH] HTA database) and conference proceedings (International Society For Pharmacoeconomics and Outcomes Research (ISPOR) website or the Cochrane Colloquium). Furthermore, trial registries may provide an outlook on what studies are currently being performed and may provide further evidence in the near future.

Until a new EE database becomes available, we recommend searching at least the basic databases MEDLINE,[101] Embase,[102] NHS EED,[97] EconLit (EBSCO),[103] and Web of Science,[104] bearing in mind that the NHS EED stopped updating in March 2015. If applicable, a search on a more disease-specific database can be necessary. As many optional databases should be added as is feasible.

STEP 2.2 OF THE OVERALL FRAMEWORK: DEVELOPMENT OF A SEARCH STRATEGY

Developing an entirely new comprehensive search strategy (i.e. a string of search terms) is a time-consuming effort which highly depends on the reviewer's experience. The time needed for developing and testing such a strategy is reported to be around 20 h for experienced reviewers.[105] It needs to be noted that these estimates also entail the testing of such a strategy against a so-called 'gold standard' (i.e. a known set that entails all relevant publications).[59] However, it is not necessary to develop and test a search strategy from scratch for every new SR-EE. When designing a comprehensive search strategy, it is advised to ask the help of a biomedical information specialist, available at many universities.[61,69,106] Considerable work has been done to support researchers in detecting relevant articles for SRs concerning the effectiveness of treatment and diagnostics. However, little has been published on empirically validated search strategies for EEs.[56] In general, a successful search strategy is regarded as one that delivers a manageable amount of references with a searcherspecified balance of sensitivity and precision.[76] The definition of what is regarded as being manageable obviously depends on the size and expertise of the review team. When making use of predefined methods for screening, researchers other than information specialists screened a median of 296 articles per hour.[107]

Important elements in a comprehensive search strategy

In searching literature databases, a search strategy typically makes use of different search terms that are related to elements in the research question. With a so-called 'conceptual approach' (also known as a 'conventional approach'[108]), different information sources are used to identify relevant terms and their synonyms.[109] Several databases offer the possibility to employ medical subject headings (referred to as MeSH® terms in e.g. PubMed®), or Emtrees® (Embase®). Both MeSH and Emtrees groups controlled vocabulary and hence serve as thesauri used to index biomedical literature in the respective databases. For a comparison of MeSH® and Emtree®, see [110].

Search filters are defined as a collection of search terms based on research and validated against a so-called 'gold standard' (i.e. a known set of relevant records),[59] used to identify certain types of records, often for very broad topics.[59,111] They are regarded as a time-saving 'ready-made solution', leaving searchers 'free to concentrate on the other aspects of the search'.[73] Hence, they improve both the efficiency and effectiveness of searches.[59]

Although there seems to be no consensus on how to set up a good search filter, filters can be tested for their quality in terms of (1) sensitivity, (2) specificity, (3) precision, and (4) accuracy(see Table 1).[59] Sensitivity is defined as the proportion of relevant citations that were retrieved; specificity is the proportion of low-quality (or off-topic) records not detected; precision is the proportion of articles that are of high quality; accuracy is the proportion of all articles that are correctly classified.[112] While it should be the general aim to maximize sensitivity,[68] a high level of precision is needed to meet the requirements of guideline developers and HTA researches and to prepare scoping or rapid reviews.[113] It should be noted that achieving a high degree of sensitivity is often associated with a lowering of precision and vice versa.[58,68,113–115]

For identification of full EEs, we recommend choosing a sensitive rather than a precise filter.

Once all synonyms, MeSH/Emtree terms, and search filters are detected, they can be connected through the Boolean or proximity operators per Patient, Intervention, Comparator, Outcome (PICO) aspect. All PICO aspects are then combined with AND. Finally, the complete search strategy can be pasted into the database search interface. It needs to be noted that each interface follows specific syntax rules.[116]

Boolean operators

Search terms within a concept (synonyms) should be combined with the Boolean operator OR. Aspects and filters can be combined into a search strategy with the use of the Boolean operator 'AND.' In addition, some search interfaces allow the use of proximity operators such as 'NEAR' or 'ADJ.' By searching for two (groups of) words on a certain internal distance, the search achieves more specificity in comparison with combining terms with 'AND' and more sensitivity in comparison with searching for specific phrases. The proximity

Table 1 - Calculation of sensitivity, precision, and specificity for the evaluation of search filters.

		Manual filter (hand searching)	
		Relevant (gold standard)	Not relevant
Search filter	Retrieved	A	B
	Not retrieved	C	D
		A + C	B + D
	Sensitivity:	$\frac{A}{A+C} \times 100$	
	Precision:	$\frac{A}{A+B} \times 100$	
	Specificity:	$\frac{D}{D+B} \times 100$	

between the words can often be set by the user. This can be of particular value if one search term can be described in several ways. The Cochrane Handbook for Systematic Reviews of Interventions (hereafter: Cochrane Handbook)[62] recommends using the 'NEAR' operator due to its higher degree of sensitivity and precision as opposed to 'NEXT' and 'AND,' respectively. It should be noted, however, that the proximity should be used only to combine words within one aspect (such as the disease or intervention aspect). Accordingly, it cannot replace the 'AND' between aspects. Theoretically, the Boolean operator 'NOT' can be used to exclude specific aspects. It should, however, be avoided in searches for SRs or used with great caution due to the possibility that it could unintentionally remove relevant records.[68]

Truncation

Most databases offer the use of truncation, which is a way to search for multiple words with the same word stem. Usually truncation is indicated with an asterisk (*) at the end of a word stem. Truncating effectiv* would for instance search for effective, effectiveness, effectivity, etc. Likewise, some databases offer a wildcard operator (such as '?' in the Cochrane Library or '#' in Ovid), which is meant to replace one single character Searching for wom?n will in this case search for women and woman.[68] Truncation should be done carefully. Truncation of the word cost* for anything related to costs will for instance also search for costimulants which is not directly related to costs. In this example, truncation took place at a word stem that was too short.

Restrictions

Most databases allow different methods for restricting their search results. It is recommended that language restrictions not be included in the search strategy,[68] although this is not always feasible. Likewise, restrictions on dates should not be applied except for specific reasons, such as when updating earlier reviews or when a certain technique being evaluated was not present before a certain date. Formats such as letters can add relevant additional information

that relates to trial reports; they can update them or may be intended to correct mistakes. Therefore, they should not be excluded per se.[68]

Selection of search terms and filters

Following the first steps of Mastrigt et al.,[30] the eligibility criteria for studies to be included in the SR are already defined. These criteria will inform the four basic components of the PICO scheme: population (or participant, or population), intervention, control or comparator, and outcomes;[117] this is a helpful step in the conceptualization of the research question.[93] Other search tools such as PICOS (where the S refers to study design) seem to be less sensitive in comparison with PICO.[118] Usually, not all PICO aspects are well covered by the title or abstracts or indexed key words of an article, and not all aspects are equally important.[68] Therefore, the final search strategy for SREEs will often consist of the following three main key concepts of interest: (1) health/disease, 2b) intervention, and (3) economics. Search terms for each concept can be derived from the conceptual approach or by using already existing search filters. For each concept, it is advised to include a wide range of free-text terms separated by the Boolean operator OR, to make as much use of truncation and wildcards as possible (see below),[68] and to use proximity operators if they are available in the interfaces used. Specifics of the three concepts will be discussed in the following subsection. Since February 2016, Embase provides a PICO search interface that can be useful for conceptualizing a first search strategy.[119]

Several databases offer the possibility of employing thesauri (also known as MeSH terms in MEDLINE or Emtree in Embase). These thesauri provide additional alternative terms that can be used as synonyms in the creation of the search strategy.

For English, it is recommended using both British and American spellings for the free-term search.[120]

Health/disease and intervention concept

As both health/disease and intervention concepts share many features and are closely related to each other, they are discussed together. For both concepts, making use of an already existing search strategy or filters is recommended. These may be found in the appendices of Cochrane SRs, publications of the NICE,[121] or other high-quality SRs. If the planned SR-EE is part of a CPG development process, information on the health- or disease-specific string can be taken from the search used to detect studies that evaluate the clinical effectiveness of the intervention of interest.

As mentioned earlier, some search filters for specific topics already exist and sometimes are even partially integrated by database providers (e.g. clinical queries in PubMed). The InterTASC Information Specialists' Sub-Group (ISSG) provides a list, updated monthly, for search filters grouped by study design and focus.[122]

Economic concept

Search terms for the economic concept are dependent on the research question and on the type of EEs that are sought to be incorporated. If, for instance, economic modelling studies are considered for the SR, it is not enough to incorporate only economics-related search terms.

Most often, search filters and full search strategies are reported together with their respective sensitivity, specificity, precision, and accuracy. In 2009, Glanville and colleagues found that EEs cannot be identified efficiently using indexing terms provided by most databases.[123] Therefore, they tested the performance of available search filters for their ability to detect EEs in MEDLINE and Embase. They concluded that, while some filters are ble to achieve high levels of sensitivity, precision is usually low.[123]

Since a newly created search filter needs to be validated, its development is a challenging, time-intensive, and resource-consuming task. Some search filters for detecting EEs have been published in the literature. Although these filters have been translated to fit more than one database, the translation is not always optimal, so they are not easily transferrable between databases. The selection of an appropriate search filter depends on the scope set out for the SR, as well as on which databases are to be searched. Therefore, we refer to the regularly updated ISSG website which holds a list of published filters for finding EEs in the databases CINAHL, Embase, MEDLINE, and PsycINFO.[124] If feasible, we advise choosing a sensitive rather than a precise search filter for SRs. This is because the former will most likely detect more records than the latter.

In 2016, the CADTH issued an update to the Peer Review of Electronic Search Strategies (PRESS) guideline that aims to evaluate electronic search strategies.[125] Originally, the PRESS guideline focused on librarians and other information specialists as primary users, but it can also be of great use for researchers undertaking SRs.

Recommendations for a complete search strategy – in a nutshell

When developing the search strategy, it is important to breakdown the research question into its main conceptual elements. The PICO scheme can help with this, although not all PICO elements might be useful.

A search strategy should encompass a wide range of freetext terms, make use of proximity operators when possible, and employ thesauri. Truncation should be used with caution, and for English, British and American spelling should be used. Restrictions of search results (e.g. language and time frame) should be used as little as possible when setting up a search strategy.

Already existing and validated search filters should be selected for being highly sensitive or highly precise or a combination of both. A soundly conducted SR will profit from a sensitive rather than from a precise search filter. Filters to find EEs can be found on the ISSG website.

Developing a search strategy: a practical example
Wijnen et al.[63] constructed a total of eight different search strategies to cover all relevant aspects that the to-be-developed CPG should cover. To keep this example comprehensible, we will focus on the search strategy for detecting publications concerning the ketogenic diet. A schematic overview on this search strategy is depicted in Figure 2.
Applying the PICO strategy to this case would detect "individuals with epilepsy" as patients, "ketogenic diet" as intervention. As no specific comparator is mentioned, it is assumed that the authors searched for any comparator possible. For this part of the CPG development process, only economic evaluations were of interest as outcomes. For studies of effects, this would obviously be different.
For the example at hand, the important aspects for a database search would thus be patient, intervention, and outcomes (since no specific comparator was of interest). For the patient and intervention aspects, an experienced information specialist compiled a broad set of search terms. For the outcome aspect, an already published search filter designed for MEDLINE was used.[126] This filter can be found on the ISSG website.[124]

STEP 2.3 OF THE OVERALL FRAMEWORK: PERFORM SEARCHES

Once the search strategies for the selected databases have been created, the search can be performed. Relevant studies that are already known should be included in the newly retrieved set of articles. If not, it needs to be determined why the search strategy could not detect them. Accordingly, the search strategy might have to be adapted. This triangulation method can serve as a sort of quality check.

A clear documentation of all searches (i.e. electronic database searches and hand and reference searches) is essential for the reproducibility and future updates of the study findings.[31,68,79,80,127] This means that the details of all searches performed (e.g. database selected, time frame covered, key words and restrictions used [i.e. the entire search strategy], number of records retrieved, etc.) should be collected systematically and added to appendices of the report (see online Appendix 2B for an example). Reference managing software (e.g. EndNote, Refworks, etc.) can be used to manage bibliographic details and deduplicate results and prepare references for publications. This will ensure efficient handling of all references retrieved from different databases.[68] The user should, however, be aware of how the reference manager used handles deduplication and the preparation of references for publication.[128,129] Reference information for gray literature and reports can be found on WorldCat®.[130]

After references from all databases have been downloaded into a reference software program, they can be deduplicated. Most reference management software programs have built-in deduplication options, but several methods have been published as well.[131–133] Deduplication is often considered time-consuming, even when using bibliographic software, because users feel the need to check the correctness of the selected duplicates. A safe and fast method has been developed in EndNote, where fields can be set upon which the duplicates are compared.[131,134]

The PICO scheme

Patients		Intervention		Comparator		Outcome
Search terms for the example of Wijnen et al.[70]						
Epilepsy	AND	Ketonic diet	AND	*- not applicable -*	AND	Economic evaluations
Synonyms / alternative keywords						
epilepsy[MeSH] OR epilepsy[TIAB] OR epileps*[TIAB] OR epilept*[TIAB] OR seizures[MeSH] OR seizures[TIAB] OR seizure[TIAB] OR Convulsion OR convulsions[TIAB]		"ketogenic diet"[MeSH] OR ("ketogenic"[TIAB] AND "diet"[TIAB]) OR "ketogenic diet"[TIAB] OR (ketogen*[TIAB] AND diet[TIAB]) OR "diet therapy"[MeSH] OR "diet therapy"[TIAB]				Search filter with best balance of sensitivity and specificity taken from Wilczynski et al. [51]: cost*[Title/Abstract] OR "costs and cost analysis"[MeSH:noexp] OR cost benefit analys*[Title/Abstract] OR cost-benefit analysis[MeSH Term] OR health care costs[MeSH:noexp]

AND, OR = Boolean operators; *MeSH* = Medical Subject Heading (for MEDLINE via PubMed); *TIAB* = abbreviation for Title/Abstract (for MEDLINE via PubMed); * = truncation (for MEDLINE via PubMed); *noexp* = EXPLODE function turned off (for MEDLINE via PubMed)

Figure 2 - Schematic overview on search strategy of Wijnen et al.[70] Per PICO item, all synonyms and MeSH terms were combined with the Boolean operator OR. Truncation (in the form of an *) was used whenever possible. All search terms were restricted to be detected in title and abstracts only (see [TIAB] or [Title/Abstract]). Within one PICO item, different words can be combined with AND. For the intervention aspect, "ketogenic" was combined with "diet". At this place a proximity operator could have been used. The same approach could also have been used for the search term "diet therapy". To detect economic evaluations, a published search filter was copied.[126] Finally, all elements of the PICO scheme were combined with the Boolean operator AND to produce a single search strategy that could then be pasted into a MEDLINE search interface (in this case PubMed).

STEP 2.4 OF THE OVERALL FRAMEWORK: SELECTION OF STUDIES

Screening of potential relevant studies should be conducted in two stages.[31,79] First, after removing the duplicates, all remaining records are screened, preferably by two independent reviewers,[135] on 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 are screened for compliance with eligibility criteria.[135] Often it is recommended that, ideally, all steps critical for study selection (2.3 and 2.4) and for data extraction (3.1 and 3.2) should be done by two reviewers independently.[31,80,135] However, as this is not always

achievable, one reviewer can select and extract the data, with a second one checking this for completeness and accuracy.[31] Pilot testing of these processes should be performed using a representative sample of studies.[31,79,135] Accordingly, the inclusion criteria should be applied to a sample of records.[79] Any discrepancies between the two reviewers should be resolved by consensus.[31,79,135] In addition, a third reviewer may be consulted if any issues need further discussion.[31,135] The review process can be done in different ways. As a formal measure of agreement, Cohen's Kappa can be calculated,[31,135] although not all guidelines regard this as necessary.[79] The review process can be managed through EndNote,[107] but many other programs are available as well. A compendium of different tools that also calculate Cohen's Kappa automatically can be found elsewhere.[136]

All information on the abovementioned processes can be reported in the study protocol and in the methods section of the publication.[31,79,135] If there are multiple records of the same study, these records should be linked together.[68,79,85] This can be done by making a systematic numerical order for the studies and reporting this in the results section. This could be done as follows: for the oldest report, the number '1A' (used further in SR-EE when reporting or discussing this study), '1B' for the second report of that specific study (mentioned only once in the results section when discussing the number of included studies), '1C' for the third publication, etc. A list of studies that were excluded from the SR at the full paper stage should be provided in the online appendices,[31,135] to keep the study transparent and reproducible. This list needs to contain bibliographic details of the excluded studies and the reason for exclusion.[31,79,135]

A flowchart of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement on study inclusion should be used to show all details of the selection process in a systematic way.[31,79,137]

EXPERT COMMENTARY & FIVE-YEAR VIEW

As much as the development of the NHS EED and HEED databases was heralded as an improvement in providing access to EEs,[113] the discontinuation of updating these databases has had a tremendous impact on how to conduct SR-EEs. The cessation of these databases created a gap, with no new database currently capable of replacing them. The scientific community seems to be reacting with procrastination. Renowned practice guidance such as the Cochrane Handbook,[62] the NICE manual for developing NICE guidelines,[64] and other reliable sources of information (e.g. the CCEMG website[94]) need to be revised and updated so that using these databases is no longer recommended. Without comprehensive economic databases, researchers need to rely on other information sources which are not specialized in EEs and must use more complex search strategies with specialized search filters to detect EE literature in available databases. Setting up a new health economic database might seem

Table 2 - Step-by-step plan on how to identify economic literature for a systematic review

Step 2: Identifying full economic evaluations	
Step 2.1 Select relevant data sources	
General databases	Select at least Medline, Embase, NHS EED, Econlit, and Web of Science. Be aware that NHS EED has not been updated since May 2015.
Specific & optional databases	Select specific databases according to your topic (if applicable).
	Search optional databases for HTA reports and conference proceedings.
Grey literature	Consider including grey literature; this can minimize bias and be a valuable source of information. Database noemen worldcat?
Citation searching	Search for relevant citations in already known publications.
	Make use of citation searching (i.e. identify articles that have cited a set of relevant articles already detected).
Step 2.2 Development of a search strategy	
Search terms	Make use of the PICO 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', 'ADJ') if possible.
	Employ thesauruses and synonyms.
	Use truncation options for your search terms (beware not to truncate to short word stems).
	For English, use British and American spelling.
Search filters	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 already developed and validated filters. The ISSG website[122] holds a regularly updated repository of such filters.
Combine search terms and filters with Boolean (AND, OR, NOT) operators	Carefully consider on what basis, and if at all, you want to restrict your search results. It is not recommended that restriction be made on the basis of language or within a narrow time frame.
Step 2.3 Perform searches	
Document the search process	Document and report all steps of the search, including the complete search strategy for every database.
Handle references	Use bibliographic software to keep track of downloaded references and publications.
	Deduplicate the downloaded records by using a reference management software.
Step 2.4 Selection of studies	
Screen references	Two reviewers should screen the references independently.
	Screen titles and abstracts of the downloaded records based on the eligibility criteria that were set earlier.

Abbreviations: NHS EED: National Health Service Economic Evaluation Database; SRs: systematic reviews; ISSG: InterTASC Information Specialists' Sub-Group; HTA: Health Technology Assessment.

like a good solution. However, with regard to the tremendous amount of resources needed to build and maintain such information repositories, it is questionable if this will add value.

Based on several key guidelines for preparing SRs in effectiveness research and on major publications exploring methods for detecting economic publications, we issue our advice on how to identify EEs for SRs in data sources not specializing solely in health economic literature. All recommendations are compiled into a step by step plan that can be used as a checklist (see Table 2).

As yet there is no consensus on how many and which specific databases need to be searched to identify all relevant EEs. Also, there is no unanimous agreement by which methodology a solid search strategy should be developed (see for instance[108,138]). Our contribution can thus be seen as merely temporary guidance until more methodological research on this topic has been published or new databases for EEs have been set up. With an increasing amount of validated, reliable, and user-friendly search filters to detect health economic literature, the creation of a new database specialized on health EEs might become redundant.

Updating new and existing SRs is a key objective for future research in this area,[139] particularly because many reviews are currently outdated or no longer accessible.[140] On the one hand, surveillance systems could assess the need for updating SRs.[141] On the other hand, Elliott et al.[142] suggest initiating living SRs which should be high quality, up-to-date online summaries of health research that are continuously updated with newly available research.

In the years to come, researchers will have the possibility to (1) implement process parallelization, (2) use novel techniques and applications to automate the process, and (3) methodologically modify certain SR processes, in order to address the issue of timeliness in the compilation of SRs.[143] Automation processes seem to be the most promising innovation in this regard,[144] as they would make handcrafted SRs (at least in part) obsolete.[145] The SR toolbox website holds a regularly updated compendium of available software tools to support the process of compiling SRs.[136] With upcoming automation processes and the increasing availability of validated search filters, it is conceivable that the cessation of health economic-specific databases will no longer be a misfortune for the scientific community. For the last decade, it seems that most research concerned with developing search strategies for detecting EEs focuses on the two major players, MEDLINE and Embase anyway.[56,58,113,123,146,147] In the near future, a search of those two databases could possibly be sufficient to detect most EEs. However, an important step for this to become reality is that EEs must be correctly indexed. Concepts related to health economics are often broadly defined, and the mere definition of what constitutes important components of EEs differs among scholars and changes over time (see definitions of costs components in [148] and [149]). Establishing new guidelines to stimulate a uniform use of terms could help overcome this issue.