KCE PROCESS BOOK
METHOD

Federaal Kenniscentrum voor de Gezondheidszorg
Centre Fédéral d'Expertise des Soins de Santé
Belgian Health Care Knowledge Center

HSR PROCESS NOTES: LITERATURE REVIEW AND INTERNATIONAL COMPARISON

PATRIEK MISTIAEN, ROOS LEROY, CARINE VAN DE VOORDE, SABINE STORDEUR, KOEN VAN DEN HEEDE

2016 www.kce.fgov.be

COLOPHON

Title:	HSR Process Notes: Literature review and international comparison
Authors:	Patriek Mistiaen, Roos Leroy, Carine Van de Voorde, Sabine Stordeur, Koen Van den Heede
Reviewers:	Raf Mertens, Dominique Paulus, Dominique Roberfroid
Other reported interests:	None declared
Layout:	Ine Verhulst
Publication date:	15 December 2016
Domain:	Method
Language:	English
Format:	Adobe® PDF™ (A4)
Legal depot:	D/2016/10.273/107

How to refer to this document? Mistiaen P, Leroy R, Van de Voorde C, Stordeur S, Van den Heede K. HSR Process Notes: Literature review and international comparison. Method Brussels: Belgian Health Care Knowledge Centre (KCE). 2016.

This document is available on the website of the Belgian Health Care Knowledge Centre.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABBREVIATION	DEFINITION
ECLIPSE	Expectations, client group, location, impact, professionals involved, service
GCP	Good clinical practice
HiT	Health Systems in Transition
HSR	Health services research
HTA	Health technology assessment
KCE	Belgian Health care Knowledge Centre
NHS	National Health Service
OECD	Organisation for Economic Co-operation and Development
PECODR	Patient or problem Exposure Comparison Outcome Duration Results
PICOS	Population, intervention, comparator, outcome and study design
PIRT	Population, index test, reference test, target disorder
PF	Project form
RAND	RAND Corporation (research and development)
SPICE	Setting, perspective, intervention, comparison, evaluation
WHO	World Health Organisation
SPIDER	Sample, Phenomenon of interest, Design, Evaluation, Research type

KCE PROCESS BOOK

1 INTRODUCTION

1.1 Health services research: concept description

Health services research (HSR) is the multidisciplinary field of scientific investigation that studies how social factors, financing systems, organisational structures and processes, health technologies, and personal behaviours affect access to health care, the quality and cost of health care, and ultimately our health and well-being. Its research domains are individuals, families, organisations, institutions, communities, and populations.[1]

Health services research aims to answer questions about optimal delivery modes of health care interventions and organisation of and payment for these within a given national or regional structure. And with 'optimal' referring to the combination of most quality of life for patients, least workforce requirements of professionals and least costs for society.

Health services research explores the impact and outcomes of the health care system on people's health and well-being. Researchers in this field often investigate access to health care, organisation of care, coordination between health care practitioners, the quality of care provided, and the cost of services. With such a broad area of study, health services researchers come from a wide variety of disciplines that not only include individuals from the health care field, such as medical doctors, nurses, and other (public) health practitioners, but also professionals from fields such as business, engineering, economics, public policy, sociology, psychology and social work.

1.2 Health services research typically requires a mixed-methods approach

HSR requires a wide range of investigation strategies, preferably executed by a multidisciplinary research team. Health services research includes literature reviews, analysis of health care utilization data sources, analysis of policy and legal documents, comparison between countries, qualitative methods to find facilitators and/or barriers, support for new organisational models, and others. Health services research is typically composed of 'mixed-methods'.

1.3 Scope: focus on literature review and international comparison

The HSR process notes have not the ambition to encompass the wide range of methods potentially used in HSR research. On the contrary, we focus in this document only on "literature review" and "international comparison", the reason being that these two approaches are commonly used by the KCE to answer HSR questions. As a consequence, they require particular attention.

A literature review on a HSR topic requires a particular approach

The basic question for performing a literature review is "what is already known about something?" and in this way a HSR literature review does not differ from a literature review for good clinical practice (GCP) or health technology assessment (HTA). However, research topics in the domain of HSR are often "complex and multidimensional" topics: many HSR interventions are implemented in a "context" and features of this context may interact with the intervention. Therefore it is very important to understand and enumerate those aspects that may modify the intervention's effects and synthesize the available interventions in context of those factors.[2] Moreover, HSR interventions are studied with varying research methodologies, and valuable information for HSR studies is often published as grey documents, which are not adopted in the classical scientific literature databases. These issues give particular challenges to a HSR literature review.

International comparison: little guidance available

Although international comparisons are widely published in HSR journals, the methods are often poorly documented. In addition, to the best of our knowledge, little methodological guidance exists (e.g. published by international organisations that conduct HSR research; methodological papers) in how to conduct international comparisons. In the KCE process notes we will elaborate on several key decisions that have to be taken in the process of conducting an international comparison in the context of KCE studies.

Other guidance can be found in existing KCE process notes

We refer the reader also to the process notes on "qualitative research"[3] and "stakeholder involvement"[4] and the "inventory of Belgian administrative databases and registers"[5] which are of particular interest when conducting HSR research.

1.4 How to use these HSR process notes?

The HSR process notes are nothing more or less than **an aid to draft the study protocol** or "Project Form (PF)". We should acknowledge that complex HSR questions can be treated using a broad variety of study methods, which means that a relatively complex decision process intervenes between the selection of the study subject and the actual start of the study. During this process it is crucial:

- To define the dimensions of the problem on which the KCE will have to issue an advice (e.g. decentralisation vs. concentration of services; type of financing; workforce implications, etc.), without at that stage knowing the options the KCE will finally propose. This thinking and discussion process will result in a selection of relevant research questions and the required pieces of evidence the KCE needs to produce to underpin the actual recommendations that will eventually be published;
- To draft a protocol that is realistic in terms of timing and costs. The study protocol should stay within the limits of the provisional budget and timing unless there are good reasons not to adhere to these.

The consequence is that sufficient time and effort should be devoted to the development and writing of the PF where methodological choices are balanced with time and budgetary constraints without jeopardising the scientific quality standards. This process will also include extensive stakeholder interaction to get a good grasp of the issues at stake and to refine the scope of the research (we refer to the KCE guidance on "stakeholder involvement"[4]).

This process should be performed in a very critical way: for each research question and anticipated methodological approach, the following questions should be asked:

- **What is the added value of this piece of research for the validity, credibility and authority of the KCE advice?**
- **What is the cost and the feasibility of the research?**
- **What is "a must have" and what is "nice to have"?**
- **Couldn't we just refer to work of others?**

The above-mentioned critical appraisal can be formalised in a systematic **"Study protocol challenge"** for HSR projects (especially when they concern complex topics). Before starting the implementation of the study protocol, it needs to be discussed in a small internal discussion meeting, involving the research team, the management, and a few KCE experts not involved in the project.

2 LITERATURE REVIEW

2.1 Introduction

Terminology: "literature review" as umbrella

We use the term "literature review" throughout this chapter; however, many other terms are used and it is not always clear to what extent they differ from each other (a list of terms we encountered related to literature reviews is in Appendix 1). They may differ on the "something" (the topic/content), on the methods they apply to find the answer as well as on the sources for searching the answer and on the type of evidence they include and synthesize. Sometimes differences are small and rather semantic, sometimes the differences are large.

Focus on main issues

This chapter is based on the chapter concerning literature reviews in the general KCE process book,[6] the earlier KCE process notes on health services research,[7] and on the documents retrieved from an exploratory literature review undertaken by the KCE concerning HSR-methods for literature reviews, performed in Spring 2015 and shortly described in Appendix 2.

This chapter gives only headlines for doing a (HSR) literature review; extensive guidance can be found from, among others:

- The Cochrane Handbook for Systematic Reviews of Interventions;[8]
- Searching for studies: A guide to information retrieval for Campbell Systematic Reviews;[9]
- The Centre for Reviews and Dissemination Guidance for undertaking reviews in health care;[10]
- The National Institute for Health Research Development of methodological guidance, publication standards and training materials for realist and meta-narrative reviews;[11]
- The Canadian McMaster University;

- The USA AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews;
- USA Institute of Medicine Finding What Works in Health Care: Standards for Systematic Reviews;[12]

Throughout this chapter additional guidance references will be mentioned for specific topics.

Review of primary studies or reviews: different approaches are needed

Literature reviews can be performed on basis of primary research or on basis of earlier reviews or on a combination of both. The choice depends on the amount of (aggregated) evidence available. The description in section 2.2 is mainly suitable for reviews based on primary research. The methodology for doing a review of reviews is still (rapidly) evolving; therefore a separate section (see section 2.3) is written about this.

Rapid reviews

Of attention, especially for topics on which policymakers want urgent advice, there is currently a lot going on about the methodology to perform rapid reviews.[13-19] This will be further elaborated in specific KCE process notes.

A domain in evolution

It is important to know that methods to perform (HSR) literature reviews are rapidly evolving and new methodological evidence is coming at great speed. As such these process notes will have to be adapted on a regular basis.

2.2 General steps in a literature review process

General steps for a HSR literature review do not differ from those stated in the general KCE process notes:[6]

- Introduction;
- Building a search question;
- Searching electronic sources;
- Searching supplementary sources;
- Searching for evidence on adverse effects;
- Selecting studies;
- Quality assessment of studies;
- Data extraction;
- Analysing and interpreting results;
- Reporting of the literature review.

However, the way steps are performed for HSR-topics may differ, e.g. search sources, type of studies to be included and data-synthesis. Henceforth, above steps will be followed and for each the typical HSR-elements will be illuminated.

It is important that all steps are thought about before starting a literature review and all considerations for each step are written down in a priori research protocol that gives guidance throughout the process. This protocol is integrated in the PF and it can be considered (e.g. during the “study protocol challenge”, see section 1.4) to make the protocol publicly available before the study is started and to submit it to a register of intended literature reviews like Prospero as is internationally advised.[20-24]

2.2.1 Introduction

The introduction is the place to describe the background of the research topic, the rationale and aim(s) of the intended research. The introduction should end with clearly formulated research questions that will be addressed in the literature review.

2.2.2 Building a search question

Constructing an effective combination of search terms for searching electronic databases requires a structured approach. One approach involves breaking down the review question into "facets". Several generic templates exist, e.g. PICOS (Population, Intervention, Comparator, Outcome and Study design), PIRT (Population, Index test, Reference test, Target disorder), SPICE (Setting, perspective, intervention, comparison, evaluation) (see Appendix 3). The best suitable acronym to use is dependent on the research question.

Most often used is the PICO(S); an example is given in Box 1.

Box 1 – Example of a PICOS strategy

Research aim: to investigate the effect(s) of introducing advanced nurse practitioners (ANPs) for complex wound care

P: patients with complex wounds (e.g. ulcerated diabetic foot, burn wounds, grade 4 pressure ulcers, etc.);

I: wound care delivered by an ANP;

C: wound care delivered by a regular nurse OR wound care delivered by a dermatologist;

O: primary outcome: wound healing time // secondary outcomes: patient satisfaction, costs;

S: individual patient randomized trials OR clustered randomized trials.

The next stage is to identify the search terms in each "facet" which best capture the subject. The group of search terms covering each facet of the review question should include a range of standard subject headings (MeSH) and a range of text words (free text to be searched in the title or abstract of studies). Subject headings and text words and their variants can be identified by reading relevant reviews and primary studies identified during earlier searches or a pre-assessment of the literature. Information on the subject indexing used by databases can be found by consulting the relevant indexing manuals and by noting the manner in which key retrieved articles have been indexed by a given database.

The final search strategy will be developed by **an iterative process** in which groups of terms are used, sometimes in several permutations, to identify the combination of terms that seems most sensitive in identifying relevant studies, and to identify what components/facets have to be combined (in most cases it is not advised to combine *all* facets of the PICOS in your search, since there is too much chance relevant studies may be missed). If key-publications are known, you always have to check if these are in the intended search strategy. Always a careful balance must be made between sensitivity (i.e. a sensitive search attempts to retrieve all relevant documents by using a broad search) and specificity (i.e. a specific search attempts to retrieve only relevant documents in a small precise search).[25] This requires skilled adaptation of search strategies based on knowledge of the subject area, the subject headings and the combination of "facets" which best capture the topic.

The KCE information specialist is available to help in designing a search strategy.

2.2.3 Searching sources

It is important to document all search sources and strategies you applied to enable reproducibility. Guidance on how to do this can be found in Atkinson et al., 2015; Rader et al., 2013.[26, 27] Below we describe how to search classic electronic literature sources, additional sources and adverse effects sources.

2.2.3.1 *Searching classic electronic literature sources*

The decision on which literature sources to use depends on the research question.

Of course, also for HSR projects the classical scientific health care literature databases (PubMed/Medline[a], Embase, Cochrane Database for Systematic Reviews, CINAHL, PsycINFO and others) have to be searched. However, many other relevant electronic bibliographic databases exist. Providing an exhaustive list of all potential sources is not possible here. The KCE library catalogue provides a list of such sources and the KCE information specialist is available to help. It is advised that when you have a key systematic review on your topic, you look at the databases they searched in.

Since HSR studies generally cover a large variety of research methods, it is advised to avoid the use of study design search filters as a default option. However, when the number of hits is too high (e.g. several thousands of hits) or the research question corresponds evidently with a particular type of study design (e.g. comparative study design in case of research questions about effectiveness), it can be considered as part of the iterative process to use search filters, e.g. to limit the search results to a particular research design as RCT or as review).

2.2.3.2 *Searching supplementary sources*

HSR studies are often published as policy documents or as large reports by HSR institutions and are often not recorded in the classical scientific literature databases. Therefore, ask experts in the field where valuable information on the topic(s) can be found. In addition, it is necessary that extra searches are done in other sources, such as:

- University and institutional repositories; main repositories are listed in Opendoar, and can also be searched from there;
- Grey literature sources:
  - https://www.openaire.eu/;
  - http://oaister.worldcat.org/;
  - http://www.opengrey.eu/;
  - CADTH: Grey Matters: a practical search tool for evidence-based medicine;
  - Health Services and Sciences Research Resources Data Sets;
- Databases of theses and dissertations like Proquest-dissertations;
- Checking reference lists of relevant documents you included earlier;
- Using related citation tools to find out who has cited a relevant document and thus could also be a relevant (and more recent) document for your topic. Hereto citation indexes can be used as Web of Knowledge, Scopus and Harzing's publish or perish;
- Special topic libraries (e.g. the VAD library for alcohol and drugs topics or the library of FOD/SPF justice for legal matters). The librarians of such libraries have also great knowledge of relevant databases and other sources for specific topics;
- Highly specialized databases (e.g. http://www.jurisquare.com/en/index.html on Belgian legislative publications);
- Google searches (see tips for searching in http://jwebnet.net/advancedgooglesearch.html);
- Websites of other HSR-institutions (see Appendix 4);

[a] PubMed and Medline are both databases from the USA National Library of Medicine; PubMed is searchable for free from any computer with internet, while Medline is only searchable by a paid interface such as OVID. Both cover about the same publications, but the way searches can be performed differ slightly. In general KCE advises to use the OVID-Medline approach. Tips to translate an OVID-MEDLINE search into Pubmed search can be found in Neyt and Chalon, 2013.[28]

- Governmental websites (parliament, ministries, services like FOD/SPF, etc.);
- Websites of international professional organisations;
- Searching content of special HSR journals (Appendix 5);
- Searching trial registers (listed on the KCE library catalogue);
- Conference proceedings (Index of conference proceedings).

The number of extra sources needed to search is dependent on the research topic and time/resources available. Preliminary searches can help to identify the appropriate searches and the final resources that will be consulted will have to be integrated in the research protocol.

To do well-performed grey sources search, mastery of different languages is needed. In general for KCE reports it is sufficient to limit to French, Dutch and English. However, when preliminary searches indicate that a large number of studies is published in another language the inclusion of this language should be considered if relevant in the context of the Belgian health care system. In that case translation support will have to be used (e.g. Google translate as starting point).

Further guidance on finding grey documents can be found in Balshem et al., 2013; Mahood et al., 2014; Stansfield et al., 2014.[29-31]

Searches for HSR topics usually result in large amounts of references that have to be handled which has implications on the timing and resources required to conduct the search.[32]

- Use of Endnote

When searching the classical databases, it is straightforward to export the results to Endnote, where they should be stored.

However, when searching for grey literature and on websites, it is often much more difficult to store your results and to export them to your reference database. Programs as Zotero or Mendeley can be helpful in this; the KCE information specialist can advise on this.

2.2.3.3 Searching for evidence on adverse effects

Although HSR interventions may also have adverse effects, there are no special sources where you can find these which is in contrast to e.g. databases of side-effects of drugs. If you are especially interested in adverse effects of an intervention, it is advised to search the regular sources with no outcome applied in the search strategy and to extract data on adverse effects later on in the review process.

2.2.4 Selecting studies

Study selection is a multi-stage process. The process by which studies will be selected for inclusion in a review should be described in the review protocol.

2.2.4.1 Inclusion and exclusion criteria

The criteria to select studies for inclusion in the review must be clearly formulated in the protocol and pilot-tested by two researchers on clarity before the inclusion process starts. This pilot testing is of utmost importance when the rest of the title and abstract sifting process is conducted by one researcher only. A pilot test could include a 10% sample with two reviewers which is followed by only one researcher sifting on title and abstract (at least when good concordance between the two was reached during the pilot text) (see below section 2.2.4.2).

The criteria depend on the elements in the research question. Those facets of the research question and the several acronyms for these were already mentioned before to design a search strategy. These acronyms are also useful to formulate in-/exclusion criteria. All those acronyms are not to be used in a rigid way and should be considered as supporting tools.

Mostly these elements are the phenomenon you are interested in, the patient categories involved, settings, effects/outcomes, conditions in which the phenomenon happens, type of study, language and date.

Of special attention for HSR topics is the type of studies to be included. Although some HSR studies focus on effectiveness, studied typically in randomized trials, HSR studies use a combination of quantitative and

qualitative approaches to get a comprehensive view of the topic: literature reviews, epidemiological data, qualitative interviews, observational studies and other approaches are combined; HSR studies are generally “mixed-method” studies. As a consequence, no method-search filter can be applied in searching for evidence, but also not with regard to inclusion criteria. These mixed-method approaches also cause challenges in data-extraction and -syntheses (see later).

2.2.4.2 Selection process

As said before, searches for HSR topics usually result in large amounts of references. In case of a very high number of hits (e.g. several thousands), it is important to reconsider your search strategy and look for ways to increase specificity without losing too much of the sensitivity of your search.

The in/exclusion process usually involves three steps:

- First, sifting out the references that are clearly not relevant based on their title;
- Second, inclusion assessment on title/abstract of the references that were kept for potential eligibility (potentially combined with step one);
- Third, inclusion assessment based on the full text.

It is advised that the inclusion process is pilot-tested beforehand to check clarity of inclusion criteria and test reproducibility and interrater agreement. The initial step of sifting out clearly irrelevant references is usually done by one researcher. Step 2 can also be performed by one researcher when a pilot-test was satisfying in terms of concordance. Step 3 is preferably performed by two researchers independently and with a third researcher available to discuss disagreements.

Sometimes it may be decided that assessments of relevance should be made by people who are blind or masked to the journal/source, the authors, the institution, and the magnitude and direction of the results. However, this takes much time, and may not be warranted given the resources required and the uncertain benefit in terms of protecting against bias.

Researchers have to decide if all inclusion criteria are judged simultaneously, resulting in one single judgment inclusion/exclusion” or if each criterion is judged separately. A hierarchical stepwise procedure can also be followed (example in Figure 1), in which documents are first assessed on the first criterion and if passed then on the second criterion, and so on. Whatever approach is followed, it must be explicitly stated in the research protocol and report.

Figure 1 – Example of a stepwise inclusion process

INCLUSION FLOW MEDLINE BASED ON TITLE/ABSTRACT						
						N references
Searched hits						2477
Does the reference concern 'hospital at home'?						
no	doubt	yes				
↓	↓	↓				910
	Does the reference concern a (cluster) randomized trial or a review including randomized trials and reported search strategy and at least two databases (including PubMed/Medline)? *(conference proceedings papers are excluded)*					
	no	doubt	yes			
		↓	↓			161
↓		Does the reference concern pre-, peri or postnatal care? Or psychiatric patients?				
↓	↓	yes	doubt	no		
↓			↓	↓		153
	↓		Is at least one outcome of interest measured (length of stay, readmissions, safety, mortality)?			
	↓		no	doubt	yes	
			↓	↓	↓	145
EXCLUSION	←	←	←	INCLUSION		(of which 82 RCT and 63 SR)

The inclusion procedure can be performed in Endnote or Excel (e.g. see the KCE XLS-document to document your selection process according to the PRISMA-guidelines). Of note, there are currently many projects going on to ease this labour intensive work by automated procedures.[16, 33-38]

The above-described procedure cannot always be followed when searching on the internet, since often the search hits cannot be stored and exported to a reference management programme; in those cases searching, sifting and initial inclusion happen simultaneously. Tools as Zotero and Mendeley can be helpful.

You librarian can assist you in obtaining the full text of potentially relevant documents. For documents that are difficult to obtain, it is advised to specifically search the internet and/or to contact the authors of the document.

2.2.5 Quality assessment of studies

Critical appraisal of included documents is a crucial part of a literature review. It aims at identifying methodological weaknesses and assessing the risk of bias in a coherent way. The methodological assessment is based on a number of key questions that focus on those aspects of the study design that have a significant influence on the validity of the results reported and conclusions drawn. These key questions differ according to the study type, and a range of checklists can be used to bring a degree of consistency to the assessment process.

Hundreds of critical appraisal tools can be found in the literature. These tools have mainly been developed for one specific study design (e.g. randomized controlled trial) or for one category of study designs (e.g. qualitative research).

Since HSR reviews typically include all types of study designs, it may be opted to apply a specific instrument for each study type or to apply a more generic instrument that is applicable to several study types. The first approach is very time-consuming, but the second is less precise.

There are instruments that are generic to multiple designs,[39-42] instruments that have been applied to typical HSR studies[43, 44] and instruments for mixed-method studies.[44-51]

The process of critical appraisal is preferably done by two reviewers independently and eventually be discussed with a third reviewer in case of disagreement. However, because of feasibility it could be acceptable that one reviewer does the quality appraisal and that a second reviewer checks the other's work (for a half of the papers and vice/versa for the other half). The chosen approach should be discussed within the research team (balance between scientific rigour and available time/budget/human resources) and documented in the PF.

If necessary, the authors of the evaluated study should be contacted for additional information.

The results of the critical appraisal should be reported in a transparent way.

2.2.6 Data extraction

Data extraction implies the process of extracting the relevant information from the selected studies that will be ultimately reported. In order to allow an efficient data extraction, the process should be discussed within the research team and detailed in the protocol before the literature search is started.

Data extraction can be very time-consuming and therefore it is important to select beforehand the variables that are needed for your research question. One researcher can perform data extraction, but this is preferably checked by a second researcher.

Key components of the data extraction include:

- Information about study reference(s) and author(s);
- Verification of study eligibility;
- Study characteristics;
- Study methods;
- Participants;
- Interventions;
- Outcome measures, instruments and timing;
- Results;
- Authors' conclusions.

It is also very important to extract data on the context and setting in which a phenomenon or the intervention occurred, since these may explain why particular results are met or not met, and to make an estimation of the transferability to other settings. For instance, if a study reports good effects of a "nurse-led"-intervention, then it is very important to check what type of nurses (e.g. general staff nurses or advanced nurse practitioners) were meant and what type of education they had.

Examples of data extraction templates can be found at the website of the Cochrane Consumers and Communication Group and the Cochrane Effective Practice and Organisation of Care Group.

To describe the interventions from the studies, one could make use of the evidence tables developed for the GCP reports or the TIDIER-checklist;[52] elements are presented in Table 1 below.

Table 1 – Tidier checklist

Item No	Item
Brief name	
1.	Provide the name or a phrase that describes the intervention
Why	
2.	Describe any rationale, theory, or goal of the elements essential to the intervention
What	
3.	Materials: Describe any physical or informational materials used in the intervention, including those provided to participants or used in intervention delivery or in training of intervention providers. Provide information on where the materials can be accessed (such as online appendix, URL)
4.	Procedures: Describe each of the procedures, activities, and/or processes used in the intervention, including any enabling or support activities
Who provided	
5.	For each category of intervention provider (such as psychologist, nursing assistant), describe their expertise, background, and any specific training given
How	
6.	Describe the modes of delivery (such as face to face or by some other mechanism, such as internet or telephone) of the intervention and whether it was provided individually or in a group
Where	
7.	Describe the type(s) of location(s) where the intervention occurred, including any necessary infrastructure or relevant features
When and How Much	
8.	Describe the number of times the intervention was delivered and over what period of time including the number of sessions, their schedule, and their duration, intensity, or dose
Tailoring	
9.	If the intervention was planned to be personalised, titrated or adapted, then describe what, why, when, and how
Modifications	
10.	*If the intervention was modified during the course of the study, describe the changes (what, why, when, and how)
How well	

11.	Planned: If intervention adherence or fidelity was assessed, describe how and by whom, and if any strategies were used to maintain or improve fidelity, describe them
12.	*Actual: If intervention adherence or fidelity was assessed, describe the extent to which the intervention was delivered as planned

**If checklist is completed for a protocol, these items are not relevant to protocol and cannot be described until study is complete. Source: Hoffmann et al. 2014[52]*

Further guidance for data extraction can be found in Hoffmann et al., 2014; Atkins et al., 2011; Fleuren et al., 2014; Flottorp et al., 2013; Hasson et al., 2010; Lewis et al., 2015; Proctor et al., 2013; Rojas Smith et al., 2014; Rycroft-Malone et al., 2013; Stirman et al., 2013; Tong et al., 2012; Little et al., 2015; Elamin et al., 2009.[52-63]

2.2.7 Analysing and interpreting results

All pieces of information have now been gathered and it is time to bring them together in such a way you can learn lessons from it. In contrast to a typical Cochrane review, it is frequently not possible to analyse results of a HSR topic in a quantitative statistical way, e.g. by performing meta-analyses (but if feasible, of course do it). HSR interventions are mostly multi-component interventions and the outcomes are frequently "soft". Almost always results have to be combined in a descriptive narrative way. There is no single, agreed framework for synthesizing complex evidence.

Data-synthesis is an intellectual challenging job. It is preferably done by more than one researcher to check if everyone interprets the data in the same way and to verify that syntheses and conclusions are valid.

To get grasp of formulating the strength of evidence, elements of the GRADE-methodology (see also the use of GRADE in KCE Good Clinical Practice guidelines)[6] can be applied.[64, 65] The GRADE working group (http://www.gradeworkinggroup.org/) has developed specific approaches to present the quality and strength of the available evidence, the judgments that bear on the quality rating, and the effects of alternative management strategies on the outcomes of interest: the GRADE evidence profile (EP) and the Summary of Findings (SoF) table (see also http://processbook.kce.fgov.be/node/120). "Key elements in a GRADE assessment, include the risk of bias in the included studies, the relevance or directness of these studies to the review question, the consistency of results from these studies, the precision of the estimate, and the risk of publication bias in the contributing evidence."

GRADE was originally developed to use in reviews that are based on quantitative studies. Recently a similar approach was developed to use in reviews based on qualitative studies: CERQual.[65] To this aim they discern four main components to assess: methodological limitations, relevance, coherence, and adequacy of data (see Table 2). Concerns about any of these components may lower the confidence in a review finding. Although the CERQual-approach is new and should be further evaluated, it seems promising for HSR studies. Examples where this approach was applied can be found in Bohren et al., 2015[66] and Colvin et al., 2013.[67]

Table 2 – CERQual approach

Component	Definition
Methodological limitations	The extent to which there are problems in the design or conduct of the primary studies that contributed evidence to a review finding
Relevance	The extent to which the body of evidence from the primary studies supporting a review finding is applicable to the context (perspective or population, phenomenon of interest, setting) specified in the review question
Coherence	The extent to which the review finding is well grounded in data from the contributing primary studies and provides a convincing explanation for the patterns found in these data
Adequacy of data	An overall determination of the degree of richness and quantity of data supporting a review finding

Source: Lewin et al., 2015[65]

Further guidance can be found in Guise et al., 2014; Tricco et al., 2011; Tong et al., 2012; Sandelowski et al., 2012; Ludvigsen et al., 2015; Kastner et al., 2013; Hannes et al., 2011; Grimshaw et al., 2010; Evans et al., 2011; Castro et al., 2010; Carroll et al., 2013; Carroll et al., 2011; Candy et al., 2011; Bos et al., 2012; Craig et al., 2008; Craig et al., 2013; Datta et al., 2013; Moore et al., 2015; Petticrew et al., 2013.[2, 49, 50, 62, 68-83]

2.2.8 *Reporting of the literature review*

A literature search should (as much as possible) be reproducible and therefore explicitly documented. The report of a literature search should contain all the sections of the protocol and all information on where and why has been deviated from the protocol, followed by the main results, analyses, strength of evidence and conclusions. In the appendices, search strategies, list of excluded studies and other things can be added.

Further guidance for reporting reviews can be found on the website of the equator network: http://www.equator-network.org/ and in the PRISMA-statement.[84]

It is important to clearly state what the review adds to what is already known on the topic.

2.3 Meta-review: Review of reviews

Since the amount of reviews produced increases very fast[b], it often occurs that several reviews exist on the same (or approximately the same) subject. These reviews can be a good source to get insight into the current state of affairs without necessitating to search yourselves for primary studies.

There are pros and cons for this approach. Main assumed advantage of doing a meta-review is time saving since someone else already searched, sifted, assessed and analysed the available evidence from primary studies. On the other hand reviews are on a higher abstraction level than primary studies and it may be difficult to get grasp on what really happened and how it was studied; details are lost. Moreover, when performing a meta-review, you are dependent on the intervention and outcome criteria that were formulated by others and these may be slightly different from what you want for your research question.

Above-mentioned steps in section 2.2 all apply to a review of reviews, but some adaptations are needed.

In the following the term meta-review will be applied as the name for review of reviews.

2.3.1 *Deciding whether or not to do a review of reviews*

The main factors that have to be taken into account when deciding to perform a meta-review are the **amount of reviews available** on the topic and the **time available for doing a review**. In general a shortage in time and/or a high number of reviews available, may favour the decision to do a meta-review.

To get a first impression on *the amount of reviews* available, you have to do a search in PubMed with MeSH-terms applicable to your topic and restricting the results to reviews by selecting these in the limits options or by adding Review[ptyp] to your search. No general guidelines exist but from a pragmatic point of view it is advisable to discuss in the research team if a meta-review is a viable option, if there are approximately 5-10 or more recent reviews fairly matching your research question.

Also *time constraints* are a factor in the decision. A regular systematic review of primary studies is estimated to take at least 3 months for two reviewers and mostly 1 year[c]. So if only 3 months or less are available, a meta-review might be considered; but it cannot be guaranteed that a meta-review can be accomplished within 3 months.

---

b In the year 2000 PubMed added 65 081 references labeled as ‘review’ and in 2014 there were 116 806 new additions as ‘review’ in the single year.

c On average a Cochrane systematic review takes 23 months from protocol to publication (http://www.cochranelibrary.com/editorial/10.1002/14651858.ED000048).

Other factors to consider in the decision:

- **Applicability** of published reviews to your topic: are the reviews really about the intervention you are interested in and did they extract data on the outcome of your interest?;
- **Recency** of the reviews, especially the search date is essential: it is good to take the search of one of the reviews and rerun it to see how many recent primary studies will be missing and to estimate if these could possibly alter the review conclusions;
- **Methodological quality of reviews**: are the reviews well performed? Hereto an assessment tool can be applied, e.g. AMSTAR[85] as also proposed in the general KCE process notes (other assessment instruments exist and an overview is presented in among others Zeng et al. (2015) and Pieper et al. (2014);[86, 87]
- **Number of included studies** in reviews and overlap between the reviews: if most or all (recent) reviews only include one or two studies, it has probably not much sense to do a new review, but if each review includes tens of primary studies with only a small overlap, this pleas for a new review;
- **Availability of meta-reviews**: it can well be the case that on popular topics, one or more meta-reviews already exist, and if so the decision to do a new one should be weighted with the same factors as mentioned above. It can also be considered to do a review of meta-reviews, but in general this is not advised because the primary evidence is getting out of sight.

No clear-cut decision process can be presented here; it really needs discussion in the research team. There are many trade-offs in determining whether it is more efficient, and a methodologically sound process to rely on a prior review or to start from scratch.[88] And keep in mind that a meta-review can only be as good as the reviews and primary studies on which it is based.

The Effective Health Care Program form the USA AHQR recently developed a list of 12 recommendations (Table 3) that can be useful to guide you in deciding and executing a meta-review.[88]

Table 3 – Recommendations for integrating systematic reviews into new reviews, EPC Guidance

	Recommendations
1	*Existing reviews should be confirmed as systematic reviews through the application of a minimum set of eligibility criteria. We propose that the minimum eligibility criteria for systematic reviews include an explicit and adequate search, application of predefined eligibility criteria to select studies, risk of bias assessment for included studies, and synthesis of results.*
2	*Criteria to assess the relevance, in terms of question elements and currency, and quality of existing systematic reviews under consideration for inclusion in reviews should be predefined.*
3	*The quality of relevant existing systematic reviews should be assessed in an explicit manner with a minimum set of quality criteria that include search of multiple sources, use of a generally accepted tool for risk of bias assessment, and sufficient information to assess the strength of the body of evidence that includes the major domains of risk of bias, directness, consistency, precision, and reporting bias.*
4	*The risk of bias assessments from the existing systematic review may be used when the review described an explicit process, including the use of a tool or method that is compatible with the approach of the current review and that assessed the key sources of potential bias.*
5	*We suggest that risk of bias assessment be repeated in a sample of studies from an existing review under consideration for inclusion in a new review to confirm concordance with current review team approach.*

	Recommendations
6	*We recommend that at a minimum, reviews should narratively describe findings of the prior review(s), including the number and types of studies included, and the overall findings.*
7	*We recommend that newly identified studies be clearly distinguished from studies in the existing review(s) when presented in the narrative and any tables (e.g. separate tables).*
8	*Summary tables should include sufficient information to support ratings for overall strength of evidence, including ratings for individual strength of evidence domains (study limitations, consistency, precision, directness, reporting bias). The strength of evidence ratings should be based on the underlying primary evidence, not the number or quality of existing systematic reviews.*
9	*Using strength of evidence domains as a framework (study limitations, consistency, precision, directness, and reporting bias), review authors should consider how new evidence would change estimates of effect or ratings for strength of evidence. A new quantitative synthesis (i.e. pooled estimate) is needed if new studies would change conclusions or strength of evidence judgements, or to obtain a more precise or more up-to-date estimate.*
10	*In cases where the existing systematic review(s) did not complete strength of evidence grading for a comparison and outcome of interest, the strength of evidence should be assessed for the body of evidence, considering primary studies from prior review(s) and any new studies identified.*
11	*In cases where no new studies are added to the body of evidence, the strength of evidence assessment from the existing systematic review may be used if conducted using an acceptable grading approach consistent with current review context. In these cases, we suggest that the overall strength of evidence assessment be reviewed, considering the strength of evidence domains, to confirm consistency with current review team assessments.*
12	*In cases where new studies are added to the body of evidence, the strength of evidence may need to be reassessed on the basis of all studies/evidence.*

Source: Robinson et al., 2015[88]

2.3.2 Building a search question

All elements of a PICO search strategy have to be followed, but for a review of reviews a search filter for reviews can be added. There are several review search filters, each with a different balance of sensitivity/specificity; according to Lee et al. the filter of healthevidence.org performs well.[89] A very simple filter for PubMed is already mentioned above; also most other databases have a simple filter to restrict the search results to reviews (e.g. OVID-Medline: *"limit X to systematic reviews"*).

2.3.3 Searching electronic sources

All regular sources named for a normal review also apply for a meta-review.

2.3.4 Searching supplementary sources

For a meta-review it is of utmost importance to check websites and repositories of other HSR institutions (Appendix 4), since these frequently produce reviews, which are often published as institutional reports. Also the other 'grey' sources as mentioned in section 2.2.3.2 apply.

McMaster University developed two databases in which they collect and rate reviews: https://www.healthsystemsevidence.org/#!/ and http://healthevidence.org/search.aspx.

It is also advised to check the review register Prospero for ongoing reviews.

2.3.5 Selecting studies

Except that in a meta-review only reviews are included, all other elements apply.

2.3.6 Quality assessment of studies

To get insight in the risk of bias and strength of evidence, two things have to be checked in a meta-review:

- Methodological quality / risk of bias of the included reviews themselves;
- Methodological quality / risk of bias of the primary studies included in the reviews and the way the review-authors have checked that.

There are several instruments available to assess methodological quality of reviews; overviews can be found in among others in Zeng et al., 2015 and Pieper et al., 2014.[86, 87] The most frequently used instrument is AMSTAR;[85] however, this instrument is much directed to reviews of RCTs, and has been criticized [90]. A more general instrument such as the one used by healthevidence.org could sometimes be a better choice.[91] Recently a large international group developed a new instrument, the ROBIS-tool, that also may be considered. [92]

To get insight in the methodological quality / risk of bias of the primary studies included in the reviews, check what instrument the review-authors applied, the results of the scoring and the conclusions of the review-authors.

It is very well possible that the reviews were well performed, but if the review authors could only identify primary studies with high risk of bias, the strength of the evidence is still low. The opposite may also occur: primary studies were well-performed, but the reviews have many weaknesses. So it is important to assess both kinds of methodological quality / risk of bias.

2.3.7 Data extraction

Main elements to extract are (non-exhaustive list):

- Bibliographic information of the reviews (authors, year, title, etc.);
- Review aim;
- Intervention(s) of interest (as worded in the review);
- Outcome(s) of interest;
- Type of studies included;
- Search sources;
- Search date;
- Risk of bias instrument applied;
- Number of studies included;
- Review authors' conclusions.

We advise to extract data as much as possible as worded in the review.

It is important to make a clear distinction between data abstracted by yourself and information copied from the existing review(s).

In addition (and depending on the time available) we advise to extract data for each of the included primary studies as mentioned in paragraph 2.8. Hereto, the wordings in the reviews can be used and you do not have to go back to the primary studies themselves. If several reviews included the same primary study, and if the reviews are comparable in their data-extraction and methodological assessment, contradictions have to be reported. In the case of contradictions, it is advised to obtain the primary study itself and to re-do the data-extraction.

2.3.8 Analysing and interpreting results

Several options are possible, depending on the time available. The **least labour intensive** way is to make a table of the conclusions of all included reviews and the directions of effect found and check if the reviews point the same direction and have similar conclusions. These may lead to conclusions as "*we included 4 reviews and all 4 have similar conclusions that the intervention has a favourable effect.*" In case of contradictory results a statement such as: "*we included 4 reviews, of which 2 found no evidence of effect, 1 found a favourable effect of the intervention and one concludes that the intervention has an opposite effect; therefore, based on the retrieved reviews it was not possible to formulate firm conclusions on the effect of X on Y*." In both cases these statements should be linked to the quality assessment (of both the systematic reviews and included primary studies) and the potential risk of bias. In case of contradictory results it is advised to analyse the primary studies in order to explain the reasons of these

differences and to report which confounding factors explain the different results.

More laborious, is to make a 2X2 table of the included primary studies by review and check per primary study the size and direction of effect as noted by the review-authors; this gives a much more detailed picture.

The most arduous is to extract data of all primary studies across all included reviews and to analyse the primary data (eventually by meta-analysis) as done in a regular systematic review. In this case all advantages and time savings of a meta-review disappear of course.

Elements of GRADE/CERQual can also be applied here to formulate the strength of evidence, as was explained in section 2.2.7.[64, 65]

2.3.9 Reporting of the meta-review

For reasons of transparency and clarity it is advisable to add in the title of the Chapter, the intervention and outcomes studied, followed by 'a meta-review' in order that readers know immediately that the results are derived from a meta-review (see also section 2.10).
Finally, keep in mind that the methodology for doing a meta-review is still (rapidly) evolving. More guidance to the above can be found in among others: Pieper et al., 2014; Robinson et al., 2015; Robinson et al., 2014; Smith et al., 2011; Thomson et al., 2010; Pham et al., 2014; Robinson et al., 2015.[87, 88, 93-97]
Those who want to keep up to date with the methodological developments in (meta-)reviews can subscribe at the weekly email alert of the Effective Health Care Program Scientific Resource Center.

3 INTERNATIONAL COMPARISON

Disclaimer:

When using the term "international comparison", instinctively one thinks about comparing a certain topic between countries. In reality however, one may be more interested in comparing and evaluating a certain topic between institutions, organisations, regions, health care systems or other settings. For the ease of reading, we will consistently use the term "countries" as setting in the following paragraphs, which should of course be replaced by the setting of interest.

3.1 Rationales for international comparisons in HSR research: does an international comparison serve your problem?

Before starting an international comparison, it is advised to consider what the added value of an international comparison for your HSR study may be. Will an "international comparison" shed a light on (part) of the problem(s) you want to study? Below we list "problems" (or phenomena) for which an "international comparison" might be appropriate:

1. Is the problem (or phenomenon) we think we have, and for which we try to formulate recommendations for a reform also seen in other countries? **How do we compare to these countries?** What countries perform much better, and, hence, could yield useful insights?

   Examples:

   - *What is the magnitude of bed blocking (patients who are occupying a hospital bed that they don't strictly need) in acute hospitals in countries A, B and C?*;
   - *What is the "use of emergency departments by patients with primary care problems" in other Western countries?;*
   - *Does nurse prescribing (i.e. the prescription of medication by nurses) exist somewhere else?*

2. **What are the possible and realistic alternative options for a reform?** Why were reform efforts in other countries introduced and do the same reasons apply to Belgium?

   Examples:

   - *How were the problems with regard to bed blocking tackled elsewhere?;*
   - *Which interventions were developed to substitute emergency care use by primary care for patients with acute, unscheduled primary care problems?;*
   - *Why was nurse prescribing (not) implemented in certain countries?;*

3. Can some of these options be categorised as "**best practice**" on the basis of a formal evaluation or of other strong indications?

   Examples:

   - *Which measures against bed-blocking in acute hospitals can be considered as successful or promising?;*
   - *Which remuneration systems for physicians resulted in the most succesful collaboration between hospital physicians and hospital managers?*

4. Are the **contextual factors** that are necessary for it to work in that setting or to support successful implementation sufficiently present in our own setting for policy transfer to take place, or are the policy reforms that would be required to create the right context sufficiently plausible?

   Examples:

   - *Why were certain measures against bed-blocking in acute hospitals in country X, Y and Z (not) successful?;*
   - *Which difficulties were observed when implementing nurse prescribing?*

3.2 Adapt the "set-up" of the international comparison to the problem you want to address

Defining the problem you want to address and the associated research questions you want to answer will help you to identify how and to what extent you will perform an international comparison. Indeed, performing an "international comparison" is not an aim by itself. It is a "set of methods" for gathering useful insights from other countries/settings. Each of the identified problems or phenomena (and related research questions) can be addressed with a different selection of countries, information gathering approach, etc.

In the first step[d], the emphasis is on comparison: **exploring differences and similarities**. Such analyses remain largely at a descriptive level, although they frequently form the basis for more analytical evaluations. Two commonly approaches are the:

- Use of international available data (e.g. OECD Health data[98]);
- Use of descriptions of international health systems (e.g. the Health Systems in Transition (HiT) series of the European Observatory on Health Systems and Policies and the related tool: "the Health System and Policy Monitor"[e]). In many instances the description of the international health system(s) under study may well be the first section of an international comparison to provide context for interpretation or it can be used as part of the preparatory phase (see 3.2.1) to draft a longlist of countries from which a short-list of cases can be selected.

[d] Is the problem we think we have, and for which we try to formulate recommendations for a reform also seen in other countries?

[e] This engine allows you to select different countries and compare their health systems. The system will automatically extract and collate the content from the published HiT for the selected countries and the selected topic. http://www.hspm.org/searchandcompare.aspx

The extent to which this step is done depends on the objective. Benchmarking with international data can, for instance, be the primary objective or it can be very limited to know the position of Belgium and to identify countries that perform better on which we could zoom in:

- An example of the former is the KCE report "Performance of the Belgian Health System"[99] where the performance of Belgium on 106 indicators was compared (when data were available) in an international context (see Papanicolas et al., 2013[100] for additional information about performance measurement);
- An example of the latter is the selection of Denmark as a case study to learn about the barriers, facilitators, prerequisites and (un-)intended effects, etc. of drastically reducing the number of emergency departments. Indeed, based on a comparison of international data about the number of emergency departments per 100 000 population it was observed that Denmark drastically reduced its number of emergency departments during the last decade. This was identified as an interesting case for Belgium since compared to neighbouring countries Belgium had an exceptionally high number of emergency departments per 100 000 inhabitants.

The rather descriptive international comparison (e.g. benchmarking data or description of health systems under study) is mostly only a first step in HSR projects at the KCE. They are followed by an analysis of "practice examples" or wherever possible from **"best practices"** to draw policy lessons for local application. Concrete questions may be: what are the strengths and weaknesses of, what are the lessons learnt from, what are the challenges in e.g. changing something? The aim is to draw lessons about why some policies seem promising and doable, promising but impossible, or doable but not promising.[101] One of the strengths of this approach stems from the observation that political events and processes in one context can often be clarified and illuminated by comparing them with similar events and processes in other contexts. The focus is often on a particular policy challenge common across countries and on how different systems address this issue so as to identify "best practice" and/or the potential to transfer policy or practice from one country to another. Comparator countries can be seen as "laboratories" for experimentation, and experiences can be very useful to develop policies and system solutions for domestic policy problems.[102]

In the next paragraphs we aim to give researchers guidance for each of the steps that have to be undertaken when an international comparison is performed:

- Preparatory work;
- Selection of countries;
- Collecting the evidence;
- Data extraction and presentation;
- Validation of information.

3.2.1 Preparatory work

Before starting the international comparison itself, some preparatory steps should be taken, in which you:

- **Explore the topic by performing a horizon scan** (e.g. Which remuneration schemes for medical specialists exist?). This includes an assessment of the grey and peer-reviewed literature, a consultation of the websites of concerned organisations (e.g. OECD, RAND, WHO, relevant medical and professional associations) as well as contacting key informants in your professional network (e.g. European Observatory on Health Systems and Policies; Belgian experts on the topic). At this stage, the scanning of the literature by means of a "quick and dirty" search in e.g. PubMed and/or Google (or other more topic-specific databases) may suffice. The number of hits may give you a hint on the amount of information you can find on the topic. Retrieved (systematic) reviews will not only render background information, but may also help you find (international) experts on the topic. It may be informative to verify the countries where the studies were performed, to check if there are authors with several publications on the topic (who can be contacted as expert) and to see what the first publication date is (are you dealing with a very old or a recent topic/problem?). Look also for existing international comparisons on the topic of interest;

- **Defining the topic should be integrated in this exploration phase** (e.g. what is bed blocking actually? What is "hospital at home"?). Often, the same words/terms do not mean the same things across countries. And different words may denote similar phenomena. Make sure you do not compare apples and oranges. Also, provide sufficient information for the reader;
- **Evaluate the anticipated amount of information**. Evaluate if there is (will be) enough information to perform an international comparison. Get an idea of which countries have (e.g. the most, little) experience on the topic of interest. See if the phenomenon is widespread. Find out if the majority of information is retrieved from peer-reviewed journals or rather from grey literature (e.g. policy documents, legislative material). Do you anticipate a language problem (e.g. if most of the experience is concentrated in the Scandinavian countries and none of the publications is available in English)? If the harvest of background information is meagre, are there other ways for obtaining information for an international comparison, e.g. by surveys, interviews?

If it is decided by the team to actually perform an international comparison and hence to continue:

- Discuss **which aspects** of the phenomenon you want to analyse:

  Examples: Do you want to focus on barriers and facilitators when implementing nurse prescribing, on the legal aspects, on the economic impact, on patient safety issues, on the opinions of the physicians, nurses or patients? Do you want to concentrate on the treatment level (health promotion, prevention or curation) of burnout in general practitioners, or on the approach (consult, one day, inpatient stay in a facility), or on the philosophy behind the approach (spontaneous and free entrance, possibility to refer, forced entrance), the target audience, the financial aspects or on some of these issues or on all of the above?

- Determine **at which level** the comparison should be done. Should the planned comparison be an entire health system comparison (i.e. a broad comparison of overall health systems), a sector-specific comparison (i.e. a comparison of segments of the health care system, e.g. primary care), a domain-based comparison (i.e. a comparison among components of the health care system, e.g. waiting times, patient experiences), an intervention/technology-based comparison (e.g. to evaluate if that new intervention/technology is as promising in all settings/countries where it is applied)?
- Assess if there is a need to **(re)phrase the research question(s).** It may, for instance, be indicated that they are made more specific (e.g. because the (anticipated) amount of information is too much to handle). Report clearly in the methods section how and why research questions were rephrased;
- Reflect on **other (broader) aspects related to the primary question** of interest that might come into play when countries have to be selected for the international comparison.

  Example: For KCE report 229 "Conceptual framework for the reform of the Belgian hospital payment system" and more precisely for the chapter on remuneration of medical specialists, it had to be decided by the team whether other health system characteristics should play a role in the selection process. For example, it had to be decided if it was necessary that some selected countries had a National Health Service (NHS) system and others a social security system. Note that if the research team is not familiar with the possible relation between the research topic (remuneration of medical specialists) and broader health system characteristics, a (second) scanning of the literature should be performed.

The decisions taken in the previous steps, should be taken into account when the following documents are being prepared (after the methods have been selected – see section 3.2.3.2):

- The literature review protocol (see section 2);
- An interview guide;
- A questionnaire for a survey.

Note: At this stage, it is of utmost importance that the research team is familiar with the subject and has a global idea of the main characteristics of the research topic in the selected countries. Otherwise, if for instance the questions in the survey are only based on the Belgian system, one runs the risk that certain parts of the questionnaire/survey are irrelevant for the selected countries and that important issues (not (yet) applicable in the Belgian context) are not surveyed.

- Evaluate **how many and which (type of) countries** you want/need to select for each of the questions/problems you want to address. As this is a very important aspect of performing an international comparison, it is further elaborated in section 3.2.2;
- Determine **which sources of information you want to use** and **which method(s)** you will apply to perform the international comparison; this topic is further described in section 3.2.3.

Based on this evaluation, a decision should be taken whether an international comparison is doable and which **balance between in-depth analysis and number of countries is most appropriate**. It is possible that some aspects (e.g. benchmarking of data) are done for a larger set of countries than others (e.g. a description of facilitators and barriers of a policy reform in a selection of countries, a selection that can be based on the benchmarking of data). The key message is that an international comparison should follow a flexible approach that is tailored to the question(s) you want to answer (e.g. describing a number of policy options taken by several countries while zooming in on lessons learned from the most succesful countries or identifying the barriers that prohibited succesful implementation in countries that envisaged similar policy reforms).

In this phase of the study it is recommended to **collaborate** closely with **Belgian and international key stakeholders/experts on the topic** (e.g. European Observatory on Health Systems and Policies; OECD; European Commission; partner organisations of the KCE). They may not only provide you with important content and context input, they can help you orientate the search for information (e.g. terms usually applied for a phenomenon), refer you to other (e.g. more experienced) experts, but may also advise you on countries that can be selected for the international comparison and give suggestions for other Belgian (e.g. from Regions or Communities) and/or international key informants.

Several ways may lead you to (international) key informants, e.g. personal network, KCE-network, scanning of the scientific literature, contacting scientific organisations or government agencies, research institutes, etc. in each country, social media.

Hints & pitfalls

Make a point of recording precisely all steps performed, all decisions taken and the rationale for those decisions; they have to be clearly described in your report.

3.2.2 Selection of countries

One of the quandaries in performing international comparisons is the selection of countries (institutions/organisations/regions/health care systems): how to select them? Which ones should be selected? How many should be selected? Should we select the same countries for all research questions?

The **number of countries** selected for comparison is likely to depend on several factors, such as the purpose of the study or the extent to which a country is seen as an entire macro-social unit, which will impact on the number of countries that can be compared in a reliable way. In some cases, it may be more illuminating to select individual regions (or even individual institutions) for comparison across countries rather than entire countries, in particular if a large degree of in-country variation is observed.[102]

Neither in the scientific literature, nor on the website of institutes performing international comparisons (e.g. WHO, OECD, European Observatory) guidance on this issue is available. Even more, not all authors performing cross-country comparisons inform the reader about the rationale for the countries they selected. Yet, what is stressed by many:

- **Predefined criteria**, clearly described in the study protocol, should be applied;
- The choices made when selecting countries should be explicit and relevant (e.g. not simply reflecting convenience) and made **transparent** to the reader; [102]
- If an initially selected country is eventually not retained, this should be clearly described just like the reason for finally not having retained that country.

In addition we suggest:

- To draft a **longlist of countries** based on the preparatory work;
- To reduce the longlist of 'candidate countries' to a **shortlist for each of the research questions**. Two researchers should conduct this process independently on predefined criteria. In a second step, the shortlists should be compared and the argumentation for the choices made reviewed. A a **flow chart** can be drawn to make this decision process more transparent.

The following non-exhaustive list of areas of special attention may guide you in the careful selection of countries and/or may be contributory in the selection process:

- The selection of countries for comparison should **reflect the aims of the cross-country comparison**, i.e. the purpose of the comparison and the question to be addressed;
- The **availability of (free) information on the topics and parameters** listed in the preparatory step, available in the **languages** mastered by the research team (or approaches to deal with languages not mastered by the team: google-translate, key informants, etc.);
- **Policies that are already in operation.** Too frequently, lessons are drawn from concepts that have yet to be put into practice, on the basis of belief about what they might achieve if ever implemented.
- If the objective of the international comparison is to identify facilitators and barriers to implement a certain policy, the availability of **evaluative research** may be an asset;
    - *Example: In KCE Report 219 "Organisation of care for adults with a rare or complex cancer" the Netherlands was chosen because centralisation of care for patients with rare cancers had been implemented earlier and evaluative information was readily available.*
- The (free) **availability of (accurate) data** may further stipulate the selection of countries;
- **Geographical proximity** may be a criterion, e.g. when field visits are envisaged;
- **Time and budgetary constraints** also play a role in the selection (and should hence also be reported);
- **Recommendations** by well-informed stakeholders/experts in the field, for example based on personal contacts, current and past collaborations, etc.

In most instances **multiple reasons** may apply for the selection of countries. Yet, they all have to be carefully described in the methods section. In most HSR projects conducted by the KCE, the scoping of the subject will result in strong hints about potential solution elements for the problem at hand. Therefore, the international comparison in most HSR studies aims to give insights from other countries as a way to test or refute our working hypotheses, and to choose the countries for that purpose. E.g. for the study on the organisation of care for adults with rare and complex cancers, it does not make sense to elaborate on the variability in country policies. What you need is information from countries where they explicitly installed a policy to address the dispersion of care.

3.2.3 Collecting the evidence

3.2.3.1 Sources of information

International comparisons should be based on a multitude of information sources. The use of multiple sources of information allows the topic of investigation being addressed more broadly. Also, any finding or conclusion is likely to be much more convincing and accurate if it is based on several different sources of information, following a corroboratory mode ("the converging lines of inquiry"). With data triangulation, the potential problems of construct validity can also be addressed because the multiple sources of evidence essentially provide multiple measures of the same phenomenon.[103]

Depending on the purpose of the cross-country comparison, the following sources can be used:

- **Peer-reviewed scientific literature**, retrieved from peer-reviewed journals and **grey literature**, literature that is not formally published in scientific journals, e.g. technical reports from government agencies or scientific research groups, working papers from research groups or committees, white papers, and preprints (see section "Literature search"), letters, minutes of meetings, administrative documents, documents from websites of government institutions and professional associations. A problem often encountered with grey literature is that it is very often written in the national language without a translation in English, Dutch or French. As such it may be required to use key informants or translation aids (e.g. google translate) (see also section 2 "Literature search");
- **Laws and regulations;**
    - *Example: The way nurse prescribing is actually implemented in a country largely depends on the country's legal framework.*
- **Websites of professional and scientific associations;**
    - *Example: For KCE report 165 "Management of burn-out in general practitioners" the websites of professional associations of physicians were consulted to retrieve extra information on planned and ongoing projects in the selected countries.*
- As stated earlier **key informants** with large experience in the topic under investigation (e.g. field actors who have a large knowledge of the legislation or who worked as policy advisors in the sector) from the selected countries should be asked to extend the knowledge base by providing supplementary information. For instance, with regard to the correct interpretation of information (e.g. legal rules) their feedback may be very useful;
- **Administrative or routinely collected data** and (**national or regional) registries:** for KCE reports it is advised to first explore data from international organisations such as Eurostat and OECD before the use of country-specific (un-)published data based publications is used. Yet, it is important to evualuate the validity and accuracy of these data sources in the context of the relevant research questions;
- Also social media (e.g. Twitter) can be used as a way to get access to information or get in touch with appropriate key informants.

Note:

It is impossible to add here an order of relative importance of the different sources of information, as that will depend largely on the research question to be answered. For instance, in situations without much information readily available – nor in the peer-reviewed scientific literature nor in the grey literature - key informants from a country may be a guide to ("hidden") information. In other instances, key informants will only be convinced to spend time and energy in your research if you first demonstrate that you have already done some efforts yourself.

3.2.3.2 Methods to obtain information

- **Literature review**. It is important to realise that a review of the literature is for many HSR topics not sufficient to find the latest information on the topic of interest (see also section 2 "Literature search");
- **Questionnaires/surveys;**

- **Interviews**, where it is important to follow the line of inquiry, as reflected by the protocol and to ask actual conversational questions in an unbiased manner[103];
- **Site visits** can be considered to make the concept more tangible and to have access to several key informants in a short time period (e.g. site visit of a "major trauma centre" in the Netherlands and Germany);
- **International conferences;**
    - *Example: For KCE report 165 "Management of burn-out in general practitioners" the KCE team attended the congress of the European Association for Physician Health in Barcelona where they met several key-informants/initiators of programmes to prevent/manage burn-out in general practitioners.*
- **Other …**

Note:

Again, it is impossible to give advice on the selection of methods to be used or on the order of using them, as this will depend heavily on the research question. But, for instance, if you plan to perform a survey it is recommended first to perform the literature review as that will give you more insight into the topic, highlight issues/problems you have not thought of and thus enrich your questionnaire. It can also be useful to check with Belgian stakeholders if all relevant questions for the problem under study are covered by the questionnaire.

3.2.4 Data extraction and presentation

- Make sure you have an analysis matrix or a data-extraction form before you start with the data extraction. This form facilitates a consistent data extraction throughout the process.
- After data extraction of a couple of countries, it may be necessary to adapt the analysis matrix or the data-extraction form (e.g. leave out some less important aspects, add some others);
- Present the data so that it facilitates the critical evaluation of your data, e.g. tables by topic for a comparative analysis;
    - *Example: In KCE report 165 "Management of burn-out in general practitioners" a table is provided with a transversal overview of the facilities, which were categorised in 3 groups: facilities with a low threshold, the facilities for care coordination and the facilities for treatment.*
- In some instances (when it helps to underpin the recommendations) a narrative description may be useful, e.g. country by country;
    - *Example: In KCE report 72 "Physician workforce supply in Belgium. Current situation and challenges" a case study report was written for each country.*

3.2.5 Validation

It is recommended to ask in a final step key informants (e.g. those informants who delivered the data) of the selected countries/regions to review and validate the sections you have written about their country/region. In this way misinterpretations may be put right, important details and the most recent information can be added (e.g. implementation/abandonment of certain policies, first evaluative data). It is not always necessary to let them review the entire chapter, it may suffice to ask them to review a couple of tables and key messages.

APPENDICES

APPENDIX 1. TERMS ENCOUNTERED FOR LITERATURE REVIEWS

- Bayesian network meta-analysis;
- Comparative effectiveness reviews;
- Comprehensive reviews;
- Critical interpretive synthesis;
- Evidence briefs;
- Evidence summaries;
- Evidence synthesis;
- Framework synthesis;
- Health impact assessment;
- Health policy assessment;
- Health technology assessment;
- Individual patient data meta- analysis;
- Integrative (literature) review;
- Knowledge synthesis;
- Mega-analysis;
- Meta-analysis;
- Meta-ethnography;
- Meta-narrative Evidence Syntheses;
- Meta-narrative reviews;
- Meta-review;

- Meta-meta-review;
- Meta-summary;
- Meta-synthesis;
- Mixed method review;
- Mixed methods-mixed research synthesis;
- Mixed-model reviews;
- Mixed studies review;
- Network meta-analysis;
- Overview of reviews;
- Rapid Evidence Assessment;
- Rapid knowledge synthesis;
- Rapid realist review;
- Rapid review;
- Realist synthesis;
- Research synthesis;
- Review of reviews;
- Scoping meta-review;
- Scoping review;
- Systematic literature review;
- Systematic review;
- Systematic review of systematic reviews;
- Umbrella review;
- ….

APPENDIX 2. EXPLORATORY LITERATURE REVIEW FOR METHODS ON HSR LITERATURE REVIEWS

In order to update KCE process notes concerning HSR literature reviews, an exploratory literature review was undertaken.

First, main *health literature databases* were searched:

- PubMed
- Embase
- CINAHL

Next to those general sources, the websites of well-known institutes performing HSR-studies were checked and content list of typical HSR-journals were screened. In addition, websites of organisations concerning reporting policies and research appraisal were checked. Also the library science database LISTA (Library, Information Science & Technology Abstracts) was searched since the focus of this project is on methodology for systematic reviews what is an area of research for librarians. As final source, the included relevant articles from above steps, were entered in “Publish or Perish” (www.harzing.com) to find out who has cited this article and these citing articles will be checked on inclusion criteria.

The applied search strategies were

- For PUBMED: "Health Services Research/methods"[Majr] AND ("2010/01/01"[PDAT] : "2015/12/31"[PDAT]) (hits 1194 on 110315)
- For EMBASE: 'health services research'/mj AND [2010-2015]/py (hits 2009 on 110315)
- For CINAHL: (MM "Health Services Research/MT") (2010-2015) (hits 397 on 110315)
- For LISTA: (systematic review or meta-analysis) OR TI grey literature OR DE "GREY literature" (2010-2015 AND academic journals) (hits 451 on 160415)

Articles concerning methodological approaches to perform (grey) literature searches in the HSR-field, either with regard to data sources, search strategies, methodological assessment, strength of evidence, data-extraction, data-synthesis or reporting, were included. Articles that only discussed methodological approaches to reviews of randomized clinical trials were excluded, since this is already well described by others, such as the Cochrane Collaboration.

Included manuscripts were categorized into one or more groups:

- Recommendations for data sources (with special attention to grey literature)
- Recommendations for search strategies (with special attention to grey literature)
- Approaches for methodological assessment (with special attention to instruments that are applicable to a diversity of research methods/designs)
- Approaches for measuring strength of evidence (with special attention to instruments that are applicable to a diversity of research methods/designs)
- Recommendations for data-extraction (with special attention to mixed-methods approaches and to reviews of reviews and reviews of meta-reviews)
- Recommendations for data-synthesis (with special attention to mixed-methods approaches and to reviews of reviews and reviews of meta-reviews)

All searches done (except for forward citation search), totalling 17 146 references (more were seen on websites, but not recorded if not relevant). One researcher checked potential relevance and reduced to 835 references.

APPENDIX 3. SEARCH ACRONYMS

PICOS

- Population
- Intervention
- Comparator
- Outcome
- Study design

PIRT

- Population
- Index test
- Reference test
- Target disorder

ECLIPSE

- Expectations (About improvement or innovation or information)
- Client Group (At whom is the service aimed? e.g. persons above 65)
- Location (Where is the service sited? e.g. primary care, hospital)
- Impact (What is the change in the service which is being looked for? What would constitute success? How is this being measured?
- Professionals Involved
- Service (e.g. outpatient services)

SPICE

- Setting (What is the context of the question?)
- Perspective (Who are the users/potential users of the outcomes?)
- Intervention (What is being done to them?)

- Comparison (What are the alternatives?)
- Evaluation (How will you measure if the intervention is successful?)

PECODR

- Patient or problem
- Exposure
- Comparison
- Outcome
- Duration
- Results

Further reading: Methley et al., 2014; Stern et al., 2014; Kloda et al., 2013.[104-106]

APPENDIX 4. HSR INSTITUTIONS

- AcademyHealth http://www.academyhealth.org/ and http://www.hsrmethods.org/
- Agency for Healthcare Research and Quality (AHRQ) http://www.ahrq.gov/ ("methods future research needs report"-series)
- Australian Health Services Research Institute (AHSRI) http://ahsri.uow.edu.au/index.html
- Belgian Healthcare Knowledge Centre (KCE) https://kce.fgov.be/
- Campbell Collaboration http://www.campbellcollaboration.org/
- Canadian Institute of Health Services and Policy Research http://www.cihr-irsc.gc.ca/e/13733.html
- Centre for Reviews and Dissemination (CRD) http://www.york.ac.uk/inst/crd/
- Cochrane Collaboration http://www.cochrane.org/about-us
- Cochrane Effective Practice and Organisation of Care (EPOC) http://epoc.cochrane.org/
- Cochrane Methodology Review Group http://www.cochranelibrary.com/app/content/browse/page/?context=editorial-group/Methodology%20Review%20Group
- Commonwealth Fund http://www.commonwealthfund.org/
- Evidence for Policy and Practice Information and Co-ordinating Centre http://eppi.ioe.ac.uk/cms/
- European Network for Health Technology Assessment (EUnetHTA) http://www.eunethta.eu/
- European Observatory on Health Systems and Policies http://www.euro.who.int/en/about-us/partners/observatory
- Health Services Research Europe http://www.healthservicesresearch.eu/

- Institute for Clinical and Economic Review http://www.icer-review.org/
- Institute of Medicine, http://www.iom.edu/
- MacMaster University, http://www.mcmasterhealthforum.org/
- Medical Research Council http://www.mrc.ac.uk/?nav=main
- National Collaborating Centre for Methods and Tools http://www.nccmt.ca/index-eng.html
- National Information Center on Health Services Research and Health Care Technology (NICHSR) http://www.nlm.nih.gov/nichsr/
- National Institute for Health and Care Excellence (NICE) https://www.nice.org.uk/
- National Institute for Health Research http://www.nihr.ac.uk/
- National Institutes of Health http://www.nih.gov/
- Netherlands institute for health services research (NIVEL) www.nivel.nl
- Patient-Centered Outcomes Research Institute http://www.pcori.org/
- RAND Corporation http://www.rand.org/topics/health-and-health-care.html
- Robert Wood Johnson Foundation http://www.rwjf.org/en/about-rwjf.html
- Scientific Center for Quality of Healthcare (IQ healthcare) http://www.iqhealthcare.nl/en/
- The Netherlands Organisation for Health Research and Development ZonMw, http://www.zonmw.nl/en/

APPENDIX 5. HSR JOURNALS

- Advances in Health Economics and Health Services Research http://www.emeraldinsight.com/series/ahes
- BMC Health Services Research (http://www.biomedcentral.com/bmchealthservres)
- BMC Medical Research Methodology http://www.biomedcentral.com.vdicp.health.fgov.be:8080/bmcmedresmethodol/
- Evaluation and the health professions http://ehp.sagepub.com/
- Health Affairs http://www.healthaffairs.org/
- Health Care Management Review http://journals.lww.com/hcmrjournal/pages/aboutthejournal.aspx
- Health Policy http://www.journals.elsevier.com/health-policy/
- Health policy and planning http://heapol.oxfordjournals.org/
- Health Research Policy and Systems http://www.health-policy-systems.com/
- Health Services and Outcomes Research Methodology http://www.springer.com/public+health/journal/10742
- Health Services Management Research http://hsm.sagepub.com/
- Health Services Research (http://www.hsr.org/)
- Implementation science http://www.implementationscience.com/content/4/1/50
- International journal for quality in health care http://intqhc.oxfordjournals.org/
- International Journal of Health Services http://www.sagepub.com/journals/Journal202388

- International Journal of Multiple Research Approaches http://www.tandfonline.com/loi/rmra20
- Journal of Comparative Effectiveness Research http://www.futuremedicine.com/loi/cer
- Journal of Health Politics, Policy and Law http://jhppl.dukejournals.org/
- Journal of Health Services Research & Policy (http://hsr.sagepub.com/)
- Journal of mixed methods research http://mmr.sagepub.com/
- Medical care http://journals.lww.com/lww-medicalcare/Pages/default.aspx
- Medical care research and review http://mcr.sagepub.com/
- Quality and safety in health care http://scimagojr.com/journalsearch.php?q=4800154004&tip=sid
- Research Synthesis Methods http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1759-2887
- Systematic reviews http://www.systematicreviewsjournal.com/content
- Value in Health http://www.valueinhealthjournal.com/

APPENDIX 6. METHODOLOGICAL QUALITY ASSESSMENT INSTRUMENTS

- Systematic reviews: http://processbook.kce.fgov.be/node/153
- Randomized trials: http://processbook.kce.fgov.be/node/154
- Diagnostic studies: http://processbook.kce.fgov.be/node/155
- Observational studies: http://processbook.kce.fgov.be/node/156
- Guidelines: http://processbook.kce.fgov.be/node/157
- Qualitative studies: see KCE Report 187[3] and Reynolds et al., 2011; Walsh et al., 2006; Schou et al., 2012; O'Brien, 2014; Hannes et al., 2013; Cunningham et al., 2011; Campbell et al., 2011; Dixon-Woods et al., 2007[107-114]
- Mixed-method studies: Sandelowski et al., 2012; Castro et al., 2010; Pace et al., 2012; Pluye et al., 2009; Sirriyeh et al., 2012; Souto et al., 2015; Heyvaert et al., 2013; Wisdom et al., 2012[44-51]
- Generic instruments: Crowe et al., 2011a; Crowe et al., 2011b; Katrak et al., 2004; van der Graaf et al., 2015[39-42]
- HSR-studies: O'Cathain et al., 2008; Wisdom et al., 2012[43, 44]

REFERENCES

1. Lohr KN, Steinwachs DM. Health services research: an evolving definition of the field. Health Serv Res. 2002;37(1):7-9.
2. Guise JM, Chang C, Viswanathan M, Glick S, Treadwell J, Umscheid CA, et al. Agency for Healthcare Research and Quality Evidence-based Practice Center methods for systematically reviewing complex multicomponent health care interventions. Journal of clinical epidemiology. 2014;67(11):1181-91.
3. Kohn L, Christiaens W. The use of Qualitative Research Methods in KCE studies. Method. Brussels: Belgian Health Care Knowledge Centre (KCE); 2012. KCE Reports 187 (D/2012/10.273/68) Available from: https://kce.fgov.be/sites/default/files/page_documents/KCE_187C_qualitative_research_methods.pdf
4. Piérart J, Léonard C, Chalon PX, Daue F, Mertens R. Stakeholder Involvement in KCE working processes. Method. Brussels: Belgian Health Care Knowledge Centre (KCE); 2012 20120402. KCE Reports 174 (D/2012/10.273/11) Available from: https://kce.fgov.be/sites/default/files/page_documents/KCE_174C_stakeholder%20involvement_in_KCE_working_processes.pdf
5. WIV. Healthstat - Inventaris registers [Web page].Brussel;2015 [cited 10/05/2016]. Available from: http://www.healthstat.be/web/
6. Vlayen J, Chalon P. KCE Process Book. Brussel: Belgian Health Care Knowledge Centre (KCE); 2012.
7. Van de Voorde C, Léonard L. Search for Evidence and Critical Appraisal: Health Services Research (HSR). Brussel: Belgian Health Care Knowledge Centre (KCE); 2007.
8. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011.
9. Hammerstrom K, Wade A, Jorgensen AMK. Searching for studies: A guide to information retrieval for Campbell Systematic Reviews. Oslo: The Campbell Collaboration; 2010.

10. Akers J, Aguiar-Ibáñez R, Baba-Akbari Sari A, Beynon S, Booth A, Burch J, et al. Systematic Reviews. CRD's Guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination; 2010.
11. Wong G, Greenhalgh T, Westhorp G, Pawson R. Development of methodological guidance, publication standards and training materials for realist and meta-narrative reviews: the RAMESES (Realist And Meta-narrative Evidence Syntheses - Evolving Standards) project. Health Services and Delivery Research. 2014;2(30).
12. Eden J, Levit L, Berg A, Morton SC. Finding What Works in Health Care: Standards for Systematic Reviews. IOM, Institute of Medicine; 2011.
13. Featherstone R, Dryden D, Foisy M, Guise J-M, Mitchell M, Paynter R, et al. Advancing knowledge of rapid reviews: an analysis of results, conclusions and recommendations from published review articles examining rapid reviews. Systematic Reviews. 2015;4(1):50.
14. Ganann R, Ciliska D, Thomas H. Expediting systematic reviews: methods and implications of rapid reviews. Implement Sci. 2010;5:56.
15. Khangura S, Konnyu K, Cushman R, Grimshaw J, Moher D. Evidence summaries: the evolution of a rapid review approach. Syst Rev. 2012;1:10.
16. Polisena J, Garritty C, Kamel C, Stevens A, Abou-Setta A. Rapid review programs to support health care and policy decision making: a descriptive analysis of processes and methods. Systematic Reviews. 2015;4(1):26.
17. Hayden JA, Killian L, Zygmunt A, Babineau J, Martin-Misener R, Jensen JL, et al. Methods of a multi-faceted rapid knowledge synthesis project to inform the implementation of a new health service model: Collaborative Emergency Centres. Syst Rev. 2015;4:7.
18. Schunemann HJ, Moja L. Reviews: Rapid! Rapid! Rapid! ...and systematic. Syst Rev. 2015;4:4.
19. Saul JE, Willis CD, Bitz J, Best A. A time-responsive tool for informing policy making: rapid realist review. Implement Sci. 2013;8:103.
20. The PLoS Medicine Editors. Best practice in systematic reviews: the importance of protocols and registration. PLoS Med. 2011;8(2):e1001009.
21. Stewart L, Moher D, Shekelle P. Why prospective registration of systematic reviews makes sense. Syst Rev. 2012;1:7.
22. Van der Wees P, Qaseem A, Kaila M, Ollenschlaeger G, Rosenfeld R. Prospective systematic review registration: perspective from the Guidelines International Network (G-I-N). Syst Rev. 2012;1:3.
23. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. Syst Rev. 2012;1:2.
24. Booth A, Clarke M, Ghersi D, Moher D, Petticrew M, Stewart L. An international registry of systematic-review protocols. The Lancet. 2011;377(9760):108-9.
25. Pimental S. Acquiring evidence-tips for effective literature searching. Perm J. 2005;9(2):58-60.
26. Atkinson KM, Koenka AC, Sanchez CE, Moshontz H, Cooper H. Reporting standards for literature searches and report inclusion criteria: making research syntheses more transparent and easy to replicate. Research Synthesis Methods. 2015;6(1):87-95.
27. Rader T, Mann M, Stansfield C, Cooper C, Sampson M. Methods for documenting systematic review searches: a discussion of common issues. Research Synthesis Methods. 2014;5(2):98-115.
28. Neyt M, Chalon PX. Search MEDLINE for economic evaluations: tips to translate an OVID strategy into a PubMed one. Pharmacoeconomics. 2013;31(12):1087-90.

29. Balshem H, Stevens A, Ansari M, Norris S, Kansagara D, Shamliyan T, et al. Finding Grey Literature Evidence and Assessing for Outcome and Analysis Reporting Biases When Comparing Medical Interventions: AHRQ and the Effective Health Care Program. Agency for Healthcare Research and Quality; 2013.

30. Mahood Q, Eerd DV, Irvin E. Searching for grey literature for systematic reviews: challenges and benefits. Research Synthesis Methods. 2014;5(3):221-34.

31. Stansfield C, Brunton G, Rees R. Search wide, dig deep: literature searching for qualitative research. An analysis of the publication formats and information sources used for four systematic reviews in public health. Research Synthesis Methods. 2014;5(2):142-51.

32. Havill NL, Leeman J, Shaw-Kokot J, Knafl K, Crandell J, Sandelowski M. Managing large-volume literature searches in research synthesis studies. Nurs Outlook. 2014;62(2):112-8.

33. Choong MK, Galgani F, Dunn AG, Tsafnat G. Automatic Evidence Retrieval for Systematic Reviews. Journal of Medical Internet Research. 2014;16(10):1-.

34. Lorenzetti DL, Ghali WA. Reference management software for systematic reviews and meta-analyses: an exploration of usage and usability. BMC Med Res Methodol. 2013;13:141.

35. O'Mara-Eves A, Thomas J, McNaught J, Miwa M, Ananiadou S. Using text mining for study identification in systematic reviews: a systematic review of current approaches. Syst Rev. 2015;4:5.

36. Rathbone J, Carter M, Hoffmann T, Glasziou P. Better duplicate detection for systematic reviewers: evaluation of Systematic Review Assistant-Deduplication Module. Syst Rev. 2015;4:6.

37. Stansfield C, Thomas J, Kavanagh J. 'Clustering' documents automatically to support scoping reviews of research: a case study. Research Synthesis Methods. 2013;4(3):230-41.

38. Tsafnat G, Glasziou P, Choong MK, Dunn A, Galgani F, Coiera E. Systematic review automation technologies. Syst Rev. 2014;3:74.

39. Crowe M, Sheppard L. A review of critical appraisal tools show they lack rigor: Alternative tool structure is proposed. Journal of Clinical Epidemiology. 2011;64(1):79-89.

40. Crowe M, Sheppard L. A general critical appraisal tool: An evaluation of construct validity. International Journal of Nursing Studies. 2011;48(12):1505-16.

41. Katrak P, Bialocerkowski AE, Massy-Westropp N, Kumar S, Grimmer KA. A systematic review of the content of critical appraisal tools. BMC Med Res Methodol. 2004;4:22.

42. van der Graaf Y, Zaat J. Onderzoeksartikelen lezen. Nederlands Tijdschrift voor Geneeskunde. 2015;159(A9004).

43. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy. 2008;13(2):92-8.

44. Wisdom JP, Cavaleri MA, Onwuegbuzie AJ, Green CA. Methodological reporting in qualitative, quantitative, and mixed methods health services research articles. Health Serv Res. 2012;47(2):721-45.

45. Pluye P, Gagnon MP, Griffiths F, Johnson-Lafleur J. A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in Mixed Studies Reviews. Int J Nurs Stud. 2009;46(4):529-46.

46. Pace R, Pluye P, Bartlett G, Macaulay AC, Salsberg J, Jagosh J, et al. Testing the reliability and efficiency of the pilot Mixed Methods Appraisal Tool (MMAT) for systematic mixed studies review. Int J Nurs Stud. 2012;49(1):47-53.

47. Sirriyeh R, Lawton R, Gardner P, Armitage G. Reviewing studies with diverse designs: the development and evaluation of a new tool. J Eval Clin Pract. 2012;18(4):746-52.

48. Souto RQ, Khanassov V, Hong QN, Bush PL, Vedel I, Pluye P. Systematic mixed studies reviews: Updating results on the reliability

and efficiency of the mixed methods appraisal tool. International Journal of Nursing Studies. 2015;52(1):500-1.

49. Sandelowski M, Voils CI, Leeman J, Crandell JL. Mapping the Mixed Methods-Mixed Research Synthesis Terrain. J Mix Methods Res. 2012;6(4):317-31.

50. Castro FG, Kellison JG, Boyd SJ, Kopak A. A Methodology for Conducting Integrative Mixed Methods Research and Data Analyses. J Mix Methods Res. 2010;4(4):342-60.

51. Heyvaert M, Hannes K, Maes B, Onghena P. Critical Appraisal of Mixed Methods Studies. Journal of Mixed Methods Research. 2013;7(4):302-27.

52. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. Bmj. 2014;348:g1687.

53. Atkins D, Chang SM, Gartlehner G, Buckley DI, Whitlock EP, Berliner E, et al. Assessing applicability when comparing medical interventions: AHRQ and the Effective Health Care Program. Journal of clinical epidemiology. 2011;64(11):1198-207.

54. Fleuren MA, Paulussen TG, Van Dommelen P, Van Buuren S. Towards a measurement instrument for determinants of innovations. Int J Qual Health Care. 2014;26(5):501-10.

55. Flottorp SA, Oxman AD, Krause J, Musila NR, Wensing M, Godycki-Cwirko M, et al. A checklist for identifying determinants of practice: a systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. Implement Sci. 2013;8:35.

56. Hasson H. Systematic evaluation of implementation fidelity of complex interventions in health and social care. Implement Sci. 2010;5:67.

57. Lewis CC, Stanick CF, Martinez RG, Weiner BJ, Kim M, Barwick M, et al. The Society for Implementation Research Collaboration Instrument Review Project: a methodology to promote rigorous evaluation. Implement Sci. 2015;10(1):2.

58. Proctor EK, Powell BJ, McMillen JC. Implementation strategies: recommendations for specifying and reporting. Implement Sci. 2013;8:139.

59. Rojas Smith L, Ashok M, Dy SM, Wines RC, Teixeira-Poit S. Contextual Frameworks for Research on the Implementation of Complex System Interventions. Methods Research Report.: Agency for Healthcare Research and Quality;; 2014.

60. Rycroft-Malone J, Seers K, Chandler J, Hawkes CA, Crichton N, Allen C, et al. The role of evidence, context, and facilitation in an implementation trial: implications for the development of the PARIHS framework. Implement Sci. 2013;8:28.

61. Stirman SW, Miller CJ, Toder K, Calloway A. Development of a framework and coding system for modifications and adaptations of evidence-based interventions. Implement Sci. 2013;8:65.

62. Tong A, Flemming K, McInnes E, Oliver S, Craig J. Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. BMC Med Res Methodol. 2012;12:181.

63. Little EA, Presseau J, Eccles MP. Understanding effects in reviews of implementation interventions using the Theoretical Domains Framework. Implement Sci. 2015;10:90.

64. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. Bmj. 2008;336(7650):924-6.

65. Lewin S, Glenton C, Munthe-Kaas H, Carlsen B, Colvin CJ, Gulmezoglu M, et al. Using Qualitative Evidence in Decision Making for Health and Social Interventions: An Approach to Assess Confidence in Findings from Qualitative Evidence Syntheses (GRADE-CERQual). PLoS Med. 2015;12(10):e1001895.

66. Bohren MA, Vogel JP, Hunter EC, Lutsiv O, Makh SK, Souza JP, et al. The Mistreatment of Women during Childbirth in Health Facilities Globally: A Mixed-Methods Systematic Review. PLoS Med. 2015;12(6):e1001847; discussion e.

67. Colvin CJ, de Heer J, Winterton L, Mellenkamp M, Glenton C, Noyes J, et al. A systematic review of qualitative evidence on barriers and facilitators to the implementation of task-shifting in midwifery services. Midwifery. 2013;29(10):1211-21.

68. Tricco AC, Tetzlaff J, Moher D. The art and science of knowledge synthesis. Journal of clinical epidemiology. 2011;64(1):11-20.

69. Ludvigsen MS, Hall EO, Meyer G, Fegran L, Aagaard H, Uhrenfeldt L. Using Sandelowski and Barroso's Metasynthesis Method in Advancing Qualitative Evidence. Qual Health Res. 2015.

70. Kastner M, Makarski J, Hayden L, Durocher L, Chatterjee A, Brouwers M, et al. Making sense of complex data: a mapping process for analyzing findings of a realist review on guideline implementability. BMC Med Res Methodol. 2013;13:112.

71. Hannes K, Harden A. Multi-context versus context-specific qualitative evidence syntheses: combining the best of both. Research Synthesis Methods. 2011;2(4):271-8.

72. Grimshaw J. A Guide to Knowledge Synthesis. Canadian Institutes of Health Research (CIHR); 2010.

73. Evans BC, Coon DW, Ume E. Use of Theoretical Frameworks as a Pragmatic Guide for Mixed Methods Studies: A Methodological Necessity? J Mix Methods Res. 2011;5(4):276-92.

74. Carroll C, Booth A, Leaviss J, Rick J. "Best fit" framework synthesis: refining the method. BMC Med Res Methodol. 2013;13:37.

75. Carroll C, Booth A, Cooper K. A worked example of "best fit" framework synthesis: a systematic review of views concerning the taking of some potential chemopreventive agents. BMC Med Res Methodol. 2011;11:29.

76. Candy B, King M, Jones L, Oliver S. Using qualitative synthesis to explore heterogeneity of complex interventions. BMC Med Res Methodol. 2011;11:124.

77. Bos V, van Kammen J. Kennissynthese; een handleiding. Den Haag: ZonMw; 2012.

78. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: new guidance. London: Medical Research Council; 2008.

79. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. Bmj. 2008;337:a1655.

80. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. Int J Nurs Stud. 2013;50(5):587-92.

81. Datta J, Petticrew M. Challenges to evaluating complex interventions: a content analysis of published papers. BMC Public Health. 2013;13:568.

82. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. Bmj. 2015;350:h1258.

83. Petticrew M, Rehfuess E, Noyes J, Higgins JP, Mayhew A, Pantoja T, et al. Synthesizing evidence on complex interventions: how meta-analytical, qualitative, and mixed-method approaches can contribute. J Clin Epidemiol. 2013;66(11):1230-43.

84. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4:1.

85. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol. 2007;7:10.

86. Zeng X, Zhang Y, Kwong JSW, Zhang C, Li S, Sun F, et al. The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review. Journal of Evidence-Based Medicine. 2015;8(1):2-10.

87. Pieper D, Antoine S-L, Morfeld J-C, Mathes T, Eikermann M. Methodological approaches in conducting overviews: current state in HTA agencies. Research Synthesis Methods. 2014;5(3):187-99.

88. Robinson KA, Chou R, Berkman ND, Newberry SJ, Fu R, Hartling L, et al. Twelve recommendations for integrating existing systematic reviews into new reviews: EPC guidance. Journal of Clinical Epidemiology. 2015.

89. Lee E, Dobbins M, Decorby K, McRae L, Tirilis D, Husson H. An optimal search filter for retrieving systematic reviews and meta-analyses. BMC Med Res Methodol. 2012;12:51.

90. Burda BU, Holmer HK, Norris SL. Limitations of A Measurement Tool to Assess Systematic Reviews (AMSTAR) and suggestions for improvement. Syst Rev. 2016;5(1):58.

91. Healthevidence.org. Quality Assessment Tool – Review Articles [Web page].2013 [cited 22/04/2016]. Available from: http://healthevidence.org/documents/our-appraisal-tools/QA_tool&dictionary_18.Mar.2013.pdf

92. Whiting P, Savovic J, Higgins JP, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol. 2016;69:225-34.

93. Robinson K, Chou R, Berkman N, Newberry S, Fu R, Hartling L, et al. Integrating Bodies of Evidence: Existing Systematic Reviews and Primary Studies. 2015.

94. Robinson KA, Whitlock EP, Oneil ME, Anderson JK, Hartling L, Dryden DM, et al. Integration of existing systematic reviews into new reviews: identification of guidance needs. Syst Rev. 2014;3:60.

95. Smith V, Devane D, Begley CM, Clarke M. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. BMC Med Res Methodol. 2011;11(1):15.

96. Thomson D, Russell K, Becker L, Klassen T, Hartling L. The evolution of a new publication type: Steps and challenges of producing overviews of reviews. Research Synthesis Methods. 2010;1(3-4):198-211.

97. Pham MT, Rajić A, Greig JD, Sargeant JM, Papadopoulos A, McEwen SA. A scoping review of scoping reviews: advancing the approach and enhancing the consistency. Research Synthesis Methods. 2014;5(4):371-85.

98. OECD. OECD Health Statistics 2015 [Web page].2015 [cited 26/04/2016]. Available from: http://www.oecd.org/health/health-data.htm

99. Vrijens F, Renard F, Camberlin C, Desomer A, Dubois C, Jonckheer P, et al. Performance of the Belgian Health System - Report 2015. Health Services Research (HSR). Brussels: Belgian Health Care Knowledge Centre (KCE); 2016 12/1/2016. KCE Reports 259 Available from: http://kce.fgov.be/sites/default/files/page_documents/KCE_259C_performancereport2015.pdf

100. Papanicolas I, Smith PC. Health system performance comparison An agenda for policy, information and research. Policies EOoHSa, editor. Berkshire, England: Open University Press; 2013.

101. Marmor T, Freeman R, Okma K. Comparative Perspectives and Policy Learning in the World of Health Care. Journal of Comparative Policy Analysis: Research and Practice. 2005 7(4).

102. Cacace M, Ettelt S, Mays N, Nolte E. Assessing quality in cross-country comparisons of health systems and policies: towards a set of generic quality criteria. Health Policy. 2013;112(1-2):156-62.

103. Yin RK. Case study research. Design and Methods. 3rd ed. Thousand Oaks. London. New Delhi: Sage Publications; 2003.

104. Methley AM, Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. BMC Health Serv Res. 2014;14(1):579.

105. Stern C, Jordan Z, McArthur A. Developing the Review Question and Inclusion Criteria. AJN The American Journal of Nursing. 2014;114(4):53-6.

106. Kloda LA, Bartlett JC. Formulating answerable questions: question negotiation in evidence-based practice. Journal of the Canadian Health Libraries Association. 2013;34(2):55-60.

107. Reynolds J, Kizito J, Ezumah N, Mangesho P, Allen E, Chandler C. Quality assurance of qualitative research: a review of the discourse. Health Res Policy Syst. 2011;9:43.

108. Walsh D, Downe S. Appraising the quality of qualitative research. Midwifery. 2006;22(2):108-19.

109. Schou L, Hostrup H, Lyngso EE, Larsen S, Poulsen I. Validation of a new assessment tool for qualitative research articles. Journal of Advanced Nursing. 2012;68(9):2086-94.

110. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med. 2014;89(9):1245-51.

111. Hannes K, Booth A, Harris J, Noyes J. Celebrating methodological challenges and changes: reflecting on the emergence and importance of the role of qualitative evidence in Cochrane reviews. Syst Rev. 2013;2:84.

112. Cunningham PJ, Felland LE, Ginsburg PB, Pham HH. Qualitative methods: a crucial tool for understanding changes in health systems and health care delivery. Medical care research and review : MCRR. 2011;68(1):34-40.

113. Campbell R, Pound P, Morgan M, Daker-White G, Britten N, Pill R, et al. Evaluating meta-ethnography: systematic analysis and synthesis of qualitative research. Health technology assessment (Winchester, England). 2011;15(43):1-164.

114. Dixon-Woods M, Sutton A, Shaw R, Miller T, Smith J, Young B, et al. Appraising qualitative research for inclusion in systematic reviews: a quantitative and qualitative comparison of three methods. J Health Serv Res Policy. 2007;12(1):42-7.