

Evidence reviews in energy and climate policy

Dr. Peter Warren
University College London, UK
peter.warren@ucl.ac.uk

Abstract

Evidence reviews, commonly confused with literature reviews, are a crucial method not only for collating and synthesising the evidence base, but also for determining what the quality of previous evidence is and extracting the most amount of value from previous studies using systematic techniques. The paper focuses on the use of evidence reviews to inform the development of energy and climate policies, using the UK as a case study. A framework is proposed for understanding the different types of evidence reviews based on government resource constraints. Although the application of evidence reviews is growing, the method has received much less attention in the energy and climate policy field in comparison to other policy areas, such as health policy and social policy. This paper argues that the method (particularly systematic scoping reviews and Rapid Evidence Assessments) is resource-efficient, delivers good value-for-money and is comprehensive for informing the development of energy and climate policy within the timescales and resources of governments. They ensure that only high quality evidence is used (through the use of quality assessment scales), thus helping to ensure that policies maximise positive societal impacts, minimise any negative impacts, are defensible from an expert perspective, and that they learn from past experiences, both domestically and internationally. The paper discusses the practical challenges associated with the four main types of evidence review, and draws on the experiences of four evidence reviews commissioned in 2016–2017 by the UK Government's Department for Business, Energy & Industrial Strategy (BEIS).

Key words

Evidence reviews; evidence quality; policy development; energy and climate policy

Key messages

- Evidence reviews have an important role to play in informing the development of energy and climate policy
- A framework is provided to understand the practicalities of evidence reviews under resource constraints
- A quality assessment scale is presented that was applied in four evidence reviews in 2016-17 and which can account for a diverse evidence base

Introduction

The UK Government employs a large number of scientists and analysts to undertake internal analysis or to commission external research, in order to produce robust evidence to inform the development of energy and climate policies. This is in contrast to the more common model in other countries of governments setting up separate energy (or related) agencies to undertake evidence gathering activities and analysis, and to implement policies once they have been designed by a government Ministry (for example, the Swedish Energy Agency, Japan's Agency for Natural Resources and Energy, and Singapore's Energy Market Authority). In the UK, these activities are conducted in-house

by the Department for Business, Energy and Industrial Strategy (BEIS) (2018). This paper focuses on the application of evidence reviews to inform the development of energy and climate policy from a government perspective, using the UK as a case study.

Evidence is crucial for the development of government policy to ensure that policies maximise positive societal impacts, minimise any negative impacts, are defensible, and to ensure that new policies learn from past experiences, both domestically and internationally. Evidence reviews are a useful method for comprehensively collating and synthesising the evidence base on a particular topic. The method is commonly confused with literature reviews: the former is a method for collecting data for analysis, whereas the latter is not a method, but instead aims to critique the current literature and to identify research gaps. Although the use of evidence reviews is growing in the energy and climate policy field, the method remains underutilised and some authors, such as Sorrell (2007) and Warren (2014), have called for its greater use by adapting techniques from other disciplines that more commonly use the method, such as education, crime and justice, and social welfare.

This paper argues that evidence reviews are a crucial method not only for understanding what has been done before, but also for determining what the quality of previous evidence is and extracting the most amount of value from previous studies using systematic techniques. They are a comprehensive and resource-efficient approach for governments to understand the evidence base and to utilise the results in policy development. This paper aims to review the practicalities of evidence reviews to inform the development of energy and climate policies.

The next section provides background to the different forms of evidence, gives an overview of the different types of evidence review and summarises how evidence reviews are conducted. The following section outlines a framework for understanding the different types and practicalities of evidence reviews from a government perspective. This includes discussion of four examples of evidence reviews commissioned by the UK Government to inform the development of policies to reduce emissions through household energy efficiency and demand-side response (the response of consumers to changes in energy prices or incentive payments for changing when energy is consumed – Albadi and El-Saadany, 2008). The paper then discusses the importance of evidence quality and reviews criteria that are common across fields for assessing evidence quality. The final part of the paper provides the main conclusions.

Evidence Reviews

Evidence refers to: “the available body of facts or information indicating whether a belief or proposition is true or valid” (Oxford Dictionary, 2017). There are various forms of evidence in the energy field across different disciplines, the most common being quantitative modelling (e.g. energy optimisation models, energy accounting models and energy simulation models), qualitative interviews (e.g. structured/semi-structured/unstructured interviews and focus groups) and surveys (e.g. quantitative/qualitative/mixed methods questionnaires). Less common are trials (e.g. randomised control trials and targeted (non-randomised) trials), ethnography (e.g. participant observation and shadowing), quantitative interviews (e.g. multi-criteria decision-making analysis and Q methodology) and evidence reviews (e.g. meta-analyses, realist syntheses and thematic summaries). There are various reasons for this pattern, though many revolve around methodological traditions in particular disciplines, a lack of expertise to apply methods from non-energy fields and resource costs (particularly for randomised control trials).

Although the appropriateness of the method is driven by the research questions, governments often require data and results that are nationally representative or statistically representative of particular targeted groups, such as those in fuel poverty (consumers that are “members of households that are living on a lower income in a home that cannot be kept warm at reasonable cost”, as defined in the UK’s *Warm Homes and Energy Conservation Act*, 2000) or a particular sector. Where the costs of achieving such samples are large, evidence reviews can provide a resource-efficient method for providing comprehensive data and analysis on a particular topic, as they involve collating and synthesising all of the work that has been done on a particular intervention, trial or programme to better understand what works and what does not (Petticrew and Roberts, 2006).

Evidence reviews, particularly systematic reviews (discussed in the next section), are well established in some fields, such as the medical sciences (especially through the Cochrane Collaboration (2018), which provides a database of >5,000 systematic reviews), health and social care (such as through the National Institute for Care and Health Excellence (NICE) (2018)), education, and international development (particularly through the Campbell Collaboration (2018), which was established in 2000 as the non-medical equivalent of the Cochrane Collaboration, and the Critical Appraisal Skills Programme (CASP) (2018)). In the energy field, the method is beginning to develop through initiatives, such as the UK’s Collaboration for Environmental Evidence (2018), the UK Energy Research Centre’s Technology and Policy Assessment group (2018), and the Research Councils UK’s Centre for Energy Epidemiology (2018) (based at University College London). Examples include the impacts of energy systems on marine ecosystems (Papathanasopoulou *et al.*, 2016), the barriers to energy services in the poorest countries (Watson *et al.*, 2012) and bioenergy research (e.g. Muench and Guenther, 2013; Gurwick *et al.*, 2013; Rehfuss *et al.*, 2014). However, the use of evidence reviews in the energy and climate policy sub-field remains limited.

The various different methodological techniques have been comprehensively summarised in Dixon-Woods *et al.* (2005) and Snilstveit *et al.* (2012) from a methodological perspective, and in Sorrell (2007) and Warren (2014) in relation to energy. A review of reviews by Warren (2015) is summarised in Table 1, which shows three over-arching methodological groups: integrative (primarily quantitative), mixed methods and interpretive (primarily qualitative).

Category	Systematic Review Type
Interpretive	Narrative Summaries
	Thematic Summaries
	Grounded Theory
	Meta-Ethnography
Integrative	Content Analysis
	Case Survey
	Qualitative Comparative Analysis Method
	Meta-Analysis
Mixed Methods	Meta-Study
	Realist Synthesis
	Miles and Huberman's Cross-Case Techniques
	Framework Synthesis

Table 1: the different types of evidence review by methodological group (source: collated from Dixon-Woods *et al.*, 2005; Snilstveit *et al.*, 2012; Warren, 2015)

Evidence reviews follow the eight main stages below, which is adapted from Harden and Thomas (2005), Gough *et al.* (2012) and Petticrew and Roberts (2006):

- 1) Review questions and boundaries
- 2) Selection of evidence review type
- 3) Comprehensive search strategy
- 4) Inclusion and exclusion criteria
- 5) Quality assessment scale
- 6) Data and information extraction
- 7) Synthesis of findings
- 8) Dissemination of findings

A crucial part of developing a robust review protocol (a document outlining stages 1-8 above) for evidence reviews is undertaking a pilot study to test the replicability and transparency of the methodological process. A minimum of three reviewers should independently use a small number of search terms in a few databases to ensure that the same documents are being retrieved and deemed relevant after the inclusion and exclusion criteria are applied (Warren, 2015). Following this, the reviewers should independently achieve the same quality assessment score for each document and produce the same final sample of documents. In practice, scores within +/- 1 of each other are acceptable, as the crucial aspect is that the same documents pass or fail the quality assessment stage and do not have scores that vary substantially within the 'pass' range. For example, if one reviewer scores a document maximum marks and another reviewer just passes the same document on the threshold of pass or fail, the quality assessment process adopted needs to be reviewed to ensure its robustness and replicability. An additional step may be undertaken to validate the filtering of higher quality studies and lower quality studies where time and resources allow. For example, the quality assessment scale scores could be compared with independent expert assessments of the quality of a random sample of reports using their own techniques and expert judgements. Quality assessment is discussed in more detail in the second half of the paper.

Evidence Reviews in Government

Evidence Reviews in the UK Government

The Government Social Research profession of the UK Civil Service has acknowledged the use of evidence reviews to inform government policy, but the current guidance is arguably out-of-date and is currently archived, thus limiting the development of evidence review expertise (see UK Civil Service, 2014). There is also a strong focus on just one type of evidence review, the Rapid Evidence Assessment (REA) (discussed below). However, evidence reviews range from reviews of reviews to systematic reviews with the type of review determined by the resources required to undertake it, such as time, the number of reviewers in the review team, the budget and access to relevant evidence sources. This paper presents a new framework for understanding the practicalities of different types of evidence reviews from the perspective of governments that commission them, which is shown in Figure 1. The framework is developed from the experiences of BEIS within the UK Government and is not intended to be representative of governments around the world, as specific costs will vary depending on the local context and currency.

Nevertheless, the suggested timings for undertaking different types of review should be to some degree universal.

Evidence Reviews Framework

Figure 1 also includes other review techniques that do not come under the umbrella term of 'evidence reviews', such as annotated bibliographies, traditional literature reviews and evidence maps, which are not methods for collecting data for analysis and usually do not use systematic techniques. These techniques are often referred to under the umbrella term of 'literature reviews', which instead aim to critique the current literature and to identify research gaps. This paper does not discuss these forms of review techniques but instead focuses on the practicalities of evidence reviews in the energy and climate policy field.

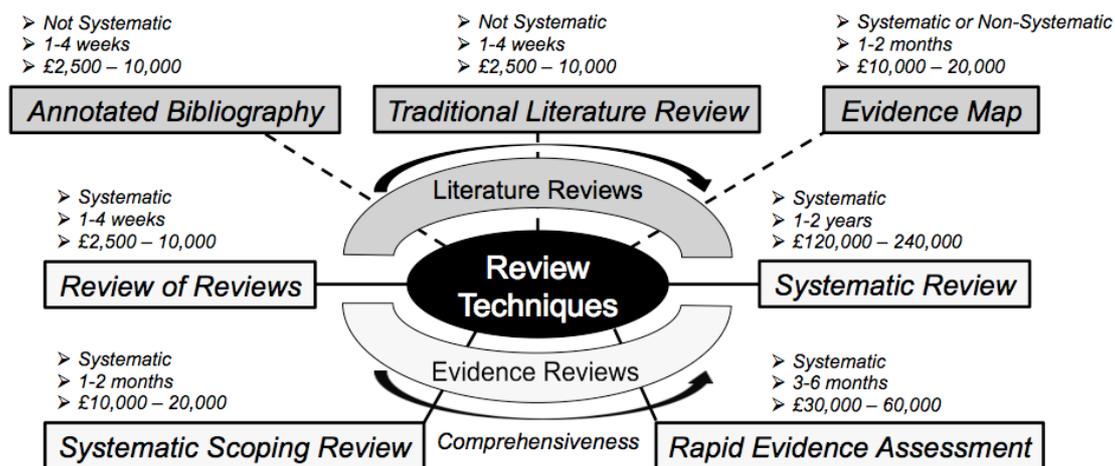


Figure 1: Evidence Reviews Framework

The evidence reviews framework is developed from both a (literature) review of evidence review practices in academia and from the practical application of new scales (such as Figure 2, which is discussed in the next section) to BEIS-commissioned projects conducted in 2016-2017. As Figure 1 shows, there are four main types of evidence review: reviews of reviews, systematic scoping reviews, rapid evidence assessments (REAs) and systematic reviews. All four types of evidence review use the systematic techniques discussed in the previous section. In Figure 1, as the reader moves from left to right of the diagram, the comprehensiveness (and resource requirements) of the review increases for both evidence reviews and literature reviews. Indicative costs (in UK pounds) and timescales are provided based on the experiences of BEIS. However, as discussed above, costs will vary based on the context of the project, government and country in question.

Review of Reviews

Reviews of reviews are the least comprehensive type of evidence review, but they are the most time-efficient. They use systematic techniques and reduce the scope of the review by just focusing on review studies (usually literature reviews, as few evidence reviews exist in the energy and climate policy field), which have already brought together the primary evidence on a particular topic. Data are then extracted from the review studies directly, and where required, the original studies may be retrieved in order to obtain the primary data. A drawback of this type of evidence review is that it relies on the existence of other reviews, thus potentially excluding primary evidence that has not yet been included in reviews. As such, the conclusions that can be drawn on the nature of the evidence base as a whole are limited. Useful references for this type of evidence review are: the UK Civil Service's Rapid Evidence Assessment Toolkit (2014) (for a summary), the Cochrane

Handbook for Systematic Reviews of Interventions (2017) (for detailed guidelines), Smith *et al.* (2011) (for a methodological approach) and Peters *et al.* (2006) (for a case study example).

From a government perspective, these reviews are useful in providing a starting point for understanding the evidence base quickly (within 1-2 weeks) and often are conducted internally, which reduces the timescales further by removing the time that would otherwise be required for procurement. Furthermore, they are more transparent and more reliable than non-systematic review techniques, and can be used to answer both broad and narrow research questions. This can be particularly advantageous in government where a broad overview of the evidence base is required within a short timescale in the early stages of policy development. However, there is a risk that without the required understanding of the strengths and weaknesses of different review techniques, policymakers may view the review as a comprehensive synthesis of the current state of evidence for a particular area. Nevertheless, by itself it is unlikely to inform the design of policies due to the complex nature of evidence collection, assessment and analysis within governments like the UK (for example, in-house secondary data analysis, in-house reviews, commissioned primary or secondary data collection and analysis, the use of submitted evidence from stakeholders, and evidence collected through call-for-evidence publications).

Systematic Scoping Reviews

Systematic scoping reviews are resource-efficient evidence reviews that can provide a reasonably comprehensive synthesis of the evidence base in a short period of time at low cost. They are usually 1-2 months in duration and are systematic. However, in certain circumstances, partially systematic scoping reviews may be conducted where the review is based on submitted evidence from stakeholders rather than through the active searching of databases and evidence sources using search strategies and inclusion criteria. However, in such cases, submitted evidence should still be subjected to quality assessment, as discussed in the next section.

From a government perspective, this paper advocates the use of systematic scoping reviews as a good compromise between comprehensiveness and government time and budget constraints. Despite this, systematic scoping reviews are suitable for answering a small number of narrow research questions rather than broader questions, as this is a necessary requirement for reducing the scope of the review but whilst maintaining a reasonably comprehensive synthesis of the evidence base. One practical challenge is where the government analyst in charge of delivering the review does not sit within an analytical team that has a research budget to commission research. In such situations, the analyst must make the case to those teams that do have research budgets based on policy needs and priorities. Alternatively, the analyst may conduct the review themselves, but this depends on their available capacity (as evidence reviews can be time-intensive) and their capability to conduct an evidence review (evidence review expertise is limited in the energy and climate policy field). Although reviews of reviews are more commonly undertaken internally, to date BEIS has primarily commissioned systematic scoping reviews. One exception to this is a systematic scoping review of the global evidence base on carbon capture and storage (CCS), which BEIS undertook internally in 2017.

In 2016-2017, BEIS commissioned three systematic scoping reviews to inform the development of policies for domestic heat and demand-side response (the response of consumers to price change or incentive payments – Albadi and El-Saadany, 2008). All three of the reviews adopted the framework and methodological technique shown in

Figure 1 and Figure 2 respectively. Two of the reviews focussed on the UK evidence base and the international evidence base on heating controls, in order to examine the energy savings, cost-effectiveness and usability of different types of controls. The results and details of the reviews are provided in BEIS (2016b; 2017a). The varied strength of the evidence, broken down by heating control type, was useful in informing the development of a domestic heat policy within the correct timescales for the policy (during the pre-call-for-evidence stage of policy development when the impact assessment was being produced). The design of two comprehensive reviews that used the same review protocol and which could be completed within 1-2 months to address an important policy need, contributed to increasing the policy impact of the review.

The third systematic scoping review examined the evidence base on household occupancy patterns, and the results are provided in BEIS (2016a). The review was not only used to inform domestic heat policies, but to contribute to the design of a large government-commissioned data collection project, as robust, representative real-world data are crucial for policy design and implementation. The review highlighted that the evidence base for occupancy patterns is currently not generalisable to the UK population. This review was the first evidence review that BEIS had commissioned since the department was formed in July 2016 and it aimed to test the use of the method to inform the new department's work. The review developed and tested the scale presented in Figure 2 (in the next section) and established a methodological approach for undertaking systematic scoping reviews in the field, which was then subsequently built upon in the two systematic scoping reviews discussed above (BEIS, 2016b; 2017a).

Rapid Evidence Assessments

Rapid Evidence Assessments (REAs) are the most commonly commissioned type of evidence review by the UK Government and their popularity is increasing within BEIS. Despite being more resource-intensive than systematic scoping reviews (in terms of time and budget), they comprehensively review the full evidence base on a particular topic and can be used to answer a much broader array of research questions than systematic scoping reviews or reviews of reviews. REAs are a condensed form of systematic review where the scope is narrowed primarily based on limiting the inclusion criteria rather than limiting the search strategy (such as the number of search terms or data sources). This is to ensure that the review remains comprehensive and can answer broader research questions. Examples of limiting the inclusion criteria include: only searching for English language documents, only including published studies and only accessing research that is freely available online. This is in contrast to systematic reviews, which have much broader inclusion criteria, as discussed below.

As with systematic reviews, a (non-systematic) literature review of the main databases and data sources that produce evidence on a given topic is first required. This is usually unnecessary for systematic scoping reviews and reviews of reviews due to their smaller scope, budget and timescales. Therefore, the choice of databases and data sources is usually determined based on the knowledge of the reviewers and the commissioning body. A challenge for commissioning REAs is the higher cost and resource required from a government perspective. As discussed above, there are important internal resource challenges that must be overcome before an REA gets to the stage of being commissioned externally. Firstly, a BEIS analyst must be identified with the capacity and expertise to project manage the REA. This is often the greatest challenge due to time constraints on government officials and limited internal expertise in the energy and climate policy field on how to conduct evidence reviews.

Unlike reviews of reviews and systematic scoping reviews, REAs are beyond the time capacity of most BEIS analysts, so to date they have been commissioned. This leads to the second challenge, which is that the REA project manager must identify those teams that have research budgets that might include the REA research topic within their funding remits. This involves the production of a business case and obtaining internal stakeholder support based on policy needs and priorities. If the REA addresses a priority policy area and has a strong rationale, funding might be authorised and the REA project manager can move through to the procurement stage.

A third challenge is external. Despite numerous academic and industry experts in the relevant topic area that an REA might focus on, there are few evidence review experts within the energy and climate policy field. Consequently, a common mistake made by potential bidders is that they present an approach to undertake a literature review rather than an evidence review that details a robust and detailed methodology that utilises systematic techniques and follows the eight stages listed at the start of the paper. In all four of the BEIS-commissioned evidence reviews discussed in this section, the BEIS project manager needed to have strong oversight of the contractors in order to ensure that they carried out the reviews correctly from a methodological perspective.

In 2016-2017, BEIS commissioned an REA to examine the evidence base on demand-side response in small energy users (households and small-to-medium-sized enterprises) to inform the development of demand-side response policies to 2025 (BEIS, 2017b). In addition to the five-month review, the project included five in-depth country case studies (Finland, Germany, Texas, Norway and Illinois), and developed a mixed-model analytical procedure for integrating the results of both methods (Saunders *et al.*, 2009). The level of depth of the case studies complemented the findings of the REA, as evidence reviews of this nature tend to focus across contexts, methods and evidence sources to look at the evidence base as a whole, rather than focussing on specific countries or contexts. The results of the REA were timed to feed in at the post-call-for-evidence stage of policy development in the design of policy proposals for demand-side response.

In addition to the challenges highlighted above, there were also challenges in developing a robust approach for integrating the systematically obtained documents from the REA and the non-systematically obtained documents from the case studies. After careful consideration, BEIS decided to keep the analyses separate, in order to ensure that all of the documents in the REA had followed the same systematic process. Despite this, the documents included in the case study were still subjected to the same quality assessment scale (shown in Figure 2 in the next section), in order to maintain consistency in the assessment of evidence quality. Integrating evidence reviews and other methods has been challenging beyond government research in the energy and climate policy field. In academia, some attempts have been made to do this (such as Warren, 2017; 2015, which integrated systematic reviews with Multi-Criteria Decision-Making (MCDM) analysis), but further methodological development is required.

Systematic Reviews

In fields such as the medical sciences, education, public health, social care and international development, systematic reviews form the top of the *Hierarchy of Evidence* pyramid (developed by the University of Illinois, 2014, and critiqued in Warren, 2014). They are resource-intensive but are fully comprehensive of the evidence base and usually include a large number of search terms, a large number of databases and data sources, a wide ranging list of inclusion and exclusion criteria (such as including snowballing, hand-searching and non-English documents), and are conducted by a larger team of reviewers

than the other types of evidence review. As Figure 1 shows, systematic reviews take a minimum of one year to complete, which is usually beyond the timescales of most governments from a commissioning perspective, as they require evidence to inform the development of policies at the correct stage in the policy process. In the UK, this occurs at the call-for-evidence stage of policy development, though it is not uncommon for evidence to be submitted before this period. However, after this period, the impact of submitted evidence is more limited, as policies would have entered the policy implementation stage of the policy process.

From a government perspective, systematic reviews are usually beyond the resources of governments if commissioned and still remain a method that has had limited application in the energy policy field (however, Warren (2015) is an example of the development of a methodological approach to apply systematic review methodology to the energy policy field). Nevertheless, independently conducted systematic reviews that are submitted by stakeholders to government officials at the correct stage of the policy process are invaluable for informing government assessments of the evidence base for a particular area. This paper encourages the development of systematic review expertise in the energy and climate policy field.

Assessing Evidence Quality

High quality evidence is crucial for underlying robust, defensible government policies. In the energy field, this is particularly challenging due to the diverse range of methods used, which vary in how they are conducted and what disciplinary perspective they are approached from. Furthermore, evidence from academia and industry also varies in how skilled experts are in understanding and communicating the policy implications of their work. In order to obtain a holistic understanding of the quality of evidence, it is important that a range of criteria are used to assess evidence quality. For example, peer review is an important indicator of evidence quality, but by itself, it does not give a comprehensive enough picture of quality and thus there is a need to use a broad range of criteria.

After the relevant studies have been selected in stage four (inclusion and exclusion criteria), stage five (quality assessment scale) of conducting evidence reviews involves assessing the quality of the evidence that has been retrieved. This filters the high quality studies from the poor quality studies. The challenge for energy and climate policy is that, as a field, it involves evidence from a diverse range of actors using different methods applied to different (but related) areas of focus from different perspectives in various timescales and a range of geographical contexts. Many existing scales struggle to deal with such diversity as they rely on specific methodological approaches and as such, they can place too much emphasis on, for example, blinding (Berger, 2006; Clark, 1999; Jadad, 1998), research reporting at the expense of research conduct (Cochrane Collaboration 2017), or qualitative evidence, thus struggling to include relevant quantitative-based studies in a robust way (Pawson *et al.* 2003).

Table 2 summarises the common metrics used in the disciplines that have a more established use of evidence reviews than in the energy policy sub-field (notably the medical sciences and social policy, which primarily apply integrative and interpretive evidence reviews respectively). The review of quality assessment scales included: the Jadad Scale (1998 – medical sciences), the guidelines of the Equator Network (Enhancing the Quality and Transparency of Health Research) (2018) and PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (2017) (medical sciences), outputs from the RAMESES project (Realist And Meta-narrative Evidence Syntheses:

Evolving Standards) (2017), the Pawson *et al.* TAPUPAS (Transparency, Accuracy, Purposivity, Utility, Propriety, Accessibility, Specificity) Framework (2003) (social care) (developed from earlier works, such as Pawson and Tilley, 1997; Pawson, 2002a; Pawson, 2002b), the Warren Scale (2014; 2015 – energy policy), the guidelines of NICE (National Institute for Care and Health Excellence) (2018) (health and social care), outputs from the UCL EPPI-Centre (Evidence for Policy and Practice Information and Coordinating Centre (2018) (multi-disciplinary) and the guidelines of CASP (Critical Appraisal Skills Programme) (2018) (health and social care).

Group	Criteria
Reporting errors	Type 1 error(s) – effect/relationship stated to exist when it does not
	Type 2 error(s) – effect/relationship stated not to exist when it does
	Type 3 error(s) – the right answer is given to the wrong question
	Transparency of data (links are provided and key data presented in an Appendix)
	Totals are given for all percentages
	Study explicitly provides answers to the research question(s)
	Conclusions match the data presented
	<i>Specific to policy/programme evaluation:</i> Details on implementation process
	<i>Specific to policy/programme evaluation:</i> Details on evaluation process
Study outline	Clear and justified rationale and study aim(s)
	Clear and justified research question(s)
Peer review	Single blinded peer review
	Double blinded peer review
	Number of peer reviewers
	Appropriate expertise of peer reviewer(s)
	External independent peer reviewer(s)
	Internal independent peer reviewer(s)
	Non-independent peer reviewer(s) (internal to project team or external)
Resource contributions	Acknowledgement of funding source
	Acknowledgement of other resources
	Statement(s) of conflicts of interest
	Statement(s) of copyright
	Statement(s) of regulatory compliance
Expertise	Expertise of author(s)
	Track record of author(s)
	Track record and expertise of author(s)'s institution(s)
	Track record and expertise of publishing institution(s)
Research Design	Study is legal
	Study is ethical
	Study is accessible to targeted audience(s)
	Meets the quality standards already used for the type of knowledge
	Clearly outlined methodological approach that is appropriate for answering the research question(s)
	Method(s) undertaken correctly
	States the sampling approach and the statistical representativeness (for quantitative research) and generalisability (for qualitative research) of the findings
	<i>Specific to certain methods (e.g. trials):</i> Blinded study
	<i>Specific to certain methods (e.g. trials):</i> Appropriate method for blinding
	<i>Specific to certain methods (e.g. trials):</i> Randomised study
	<i>Specific to certain methods (e.g. trials):</i> Appropriate method for randomisation
	<i>Specific to certain methods (e.g. trials):</i> Withdrawal rates stated
	<i>Specific to syntheses and reviews:</i> Comprehensiveness of search strategy
	<i>Specific to syntheses and reviews:</i> Comprehensiveness of inclusion criteria
	<i>Specific to syntheses and reviews:</i> Appropriateness of quality assessment scale
	<i>Specific to syntheses and reviews:</i> Appropriateness of data extraction technique(s)
<i>Specific to syntheses and reviews:</i> Appropriateness of data analysis technique(s)	
<i>Specific to syntheses and reviews:</i> Pilot test of review protocol undertaken	
Analysis	Causal expectations in the data are demonstrated through clearly outlined and appropriate approaches to analysis

	Lessons are clearly presented and applicable for wider use
Conclusions	Clearly and concisely presented conclusions that answer the research question(s)
	Research limitations and areas for further research are clearly outlined

Table 2: criteria for assessing the quality of evidence

Assessing all studies against all of the criteria shown in Table 2 would be resource-intensive and would unnecessarily penalise studies that excluded ‘surface’ reporting features, such as statements of conflicts of interest, copyright or regulatory compliance. In practice, commonly used scales adopt between five and seven quality assessment criteria and have a threshold at which studies are deemed to be of high quality. This is generally two-thirds of the available points, so for a scale with six quality criteria, a study would need to score four or more points to be included in the final sample. The quality scale presented in Figure 2 was developed by BEIS for the four evidence reviews it commissioned in 2016 (discussed in the previous section), which focused on: a) including quality criteria that could adequately deal with a diverse range of evidence and b) quality criteria that were appropriate for the focus of the four evidence reviews on specific topics within demand-side management policy. It is important to note that Figure 2 is just one example; it is not intended to be a proposed standard nor is it a standard scale used in BEIS. Instead, it demonstrates the point that a scale for assessing evidence quality should be tailored to the research question(s) and the methodological approach in question, selecting from the long list of quality criteria set out in Table 2.

- 2 points: Are the rationale (1 point) and research questions (1 point) clear and justified?
- 2 points: Does the document acknowledge resource contributions (1 point) and possible conflicts of interest (1 point)?
- 2 points: Has the document been peer reviewed or independently verified by one (1 point) or more (2 points) reputable experts?
- 1 point: Are the methods used suitable for the aims of the study?
- 1 point: Do the conclusions match the data presented?
- 1 point: Does the author / publishing organisation have a track record in the area?

Figure 2: An example of a scale used in evidence reviews commissioned by the UK’s Department for Business, Energy and Industrial Strategy (BEIS)

The scale in Figure 2 allowed all relevant evidence in the commissioned evidence reviews to be assessed equally and studies that scored two-thirds of the total number of points (six or more points out of a total of nine points) were deemed to be high quality (and were consequently included in the reviews).

There are three important considerations in using this scale. Firstly, reporting quality is considered as important as research quality as it ensures that studies are transparent, reliable and can be replicated. As Gough *et al.* (2012) argue:

“Although how well a piece of research is reported may be a proxy for the quality of the research itself, there is a danger of missing a good piece of research due to poor quality reporting; or wrongly judging a polished report as being indicative of a well-conducted piece of research.” (Gough *et al.*, 2012, p. 157)

Secondly, using a range of quality criteria offsets weaknesses in any single criterion. For example, the track record of the author(s) can be an indicator of the consistency of evidence quality. However, by itself, the criterion is limiting, as it excludes new authors in

the field that might have produced high quality work. Using a range of criteria across both research and reporting quality increases the overall reliability of the scale is increased. Note that if available information on particular criteria is lacking in the study document, it is possible to contact the author(s) to obtain further details and to ensure that potentially high quality studies are not excluded unnecessarily.

Thirdly, it may be desirable to weight some of the quality criteria (for example, in the Jadad scale (1998), an additional point is given for the criterion: “Was the study described as randomised?” if the method of randomisation was appropriate, but a point is deducted if the method was inappropriate). The use of weightings for specific quality criteria will depend on the focus of the evidence review in question, the characteristics of the relevant evidence base and the judgements of the reviewers as to what they consider to be the more important quality criteria for the research context.

Whilst using scales to assess evidence quality is an important part of evidence reviews, they could also be used to assess the quality of submitted evidence (evidence provided directly by stakeholders or experts), or to inform the design of primary evidence collection and reporting to ensure that new evidence is high quality.

Conclusion

Evidence reviews, commonly confused with literature reviews, have an important role to play in informing the development of energy and climate policy. The UK Department for Business, Energy & Industrial Strategy (BEIS)’s experiences with evidence reviews since the department was formed in July 2016 have been positive; they have been invaluable for informing the development of energy policies at an early stage of development. Furthermore, there have been methodological developments within BEIS to look at the integration of evidence reviews, particularly systematic scoping reviews and Rapid Evidence Assessments (REAs), with other methods, such as interviews, case studies (BEIS, 2017b) and non-systematic evidence searching (BEIS, 2017a). These developments allow new primary data to be combined with a comprehensive review of existing data to ensure that policies are developed based on up-to-date, robust evidence. This paper provides a framework through which stakeholders can understand the practicalities of different types of review. The paper recommends systematic scoping reviews and REAs as resource-efficient and comprehensive methods for informing policy development within the timescales and resources available to government departments.

It is important to ensure that the evidence reviewed is of high quality, so that policies maximise positive social impact, minimise negative impacts, are defensible from an expert perspective and learn from past experiences (both domestic and international experiences). BEIS’s experiences with using quality criteria and quality assessment scales have helped to determine the quality of a diverse evidence base in the energy policy field. Designing a flexible quality assessment framework with several criteria ensures that the review process covers both research quality and reporting quality. It also ensures that it is tailored to the style of the review, which will be determined by practical considerations such as timing, internal processes for sourcing funding, and limited external expertise for evidence reviews in the field of energy and climate policy.

Acknowledgements

This research was made possible by UK EPSRC support, grant number EP/H009612/1, as it is based on methodological techniques developed whilst at University College London

(UCL) and applied in the Department for Business, Energy & Industrial Strategy (BEIS). I would like to thank the anonymous peer reviewers for their helpful comments on the paper.

Bibliography

1. Albadi, MH and El-Saadany, EF, 2008, A summary of demand response in electricity markets, *Electric Power Systems Research* 78, 1989-1996
2. Berger, VW, 2006, Is the Jadad score the proper evaluation of trials? *The Journal of Rheumatology* 33, 8, 1710-1712
3. Campbell Collaboration, 2018: <http://www.campbellcollaboration.org>
4. Clark, HD, Wells, GA, Huët, C, McAlister, FA, Salmi, LR, Fergusson, D and Laupacis, A, 1999, Assessing the quality of randomised trials: reliability of the Jadad scale, *Controlled Clinical Trials* 20, 5, 448-452
5. Cochrane Collaboration, 2017, *Cochrane Handbook for Systematic Reviews of Interventions*, Version 5
6. Cochrane Collaboration, 2018: <http://www.cochrane.org>
7. Collaboration for Environmental Evidence (CEE), 2018: <http://www.environmentalevidence.org>
8. Critical Appraisal Skills Programme (CASP), 2018, *CASP Checklists*, <http://www.casp-uk.net/checklists>
9. Dixon-Woods, M, Agarwal, S, Jones, D, Young, B and Sutton, A, 2005, Synthesising qualitative and quantitative evidence – a review of possible methods, *Journal of Health Services Research Policy* 10, 1, 45-53
10. Equator Network (Enhancing the Quality and Transparency of Health Research) guidelines, 2018: [http://www.equator-network.org/?post_type=eq_guidelines&eq_guidelines_study_design=systematic-reviews-and-meta-analyses&eq_guidelines_clinical_specialty=0&eq_guidelines_report_section=0&s=+](http://www.equator-network.org/?post_type=eq_guidelines&eq_guidelines_study_design=systematic-reviews-and-meta-analyses&eq_guidelines_clinical_specialty=0&eq_guidelines_report_section=0&s=)
11. Gough, D, Oliver, S and Thomas, J, 2012, *An introduction to Systematic Reviews*, SAGE Publications Ltd., London, UK
12. Gurwick, NP, Moore, LA, Kelly, C and Elias, P, 2013, A systematic review of biochar research, with a focus on its stability in situ and its promise as a climate mitigation strategy, *PLoS ONE* 8, 9
13. Harden, A and Thomas, J, 2005, Methodological issues in combining diverse study types in Systematic Reviews, *International Journal of Social Research Methodology* 8, 3, 257-271
14. Jadad, A, 1998, *Randomised controlled trials: a users guide*, London: BMJ Books, UK

15. Muench, S and Guenther, E, 2013, A systematic review of bioenergy life cycle assessments, *Applied Energy* 112, 257-273
16. National Institute for Care and Health Excellence (NICE) Reviewing the Evidence, <https://www.nice.org.uk/process/pmg10/chapter/reviewing-the-evidence>
17. Oxford Dictionary, 2017: <https://en.oxforddictionaries.com/definition/evidence>
18. Papathanasopoulou, E, Queirós, AM, Beaumont, N, Hooper, T and Nunes, J, 2016, What are the local impacts of energy systems on marine ecosystem services, Collaboration for Environmental Evidence (CEE)
19. Pawson, R, 2002a, Evidence-based policy – in search of a method”, *Evaluation* 8, 2, 157-181
20. Pawson, R, 2002b, Evidence-based policy: the promise of ‘Realist Synthesis’”, *Evaluation* 8, 3, 340-358
21. Pawson, R, 2006, *Evidence-based Policy: A Realist Perspective*, London: SAGE, UK
22. Pawson, R and Tilley, N, 1997, *Realistic Evaluation*, London: SAGE, UK
23. Pawson, R, Boaz, A, Grayson, L, Long, A and Barnes, C, 2003, SCIE Knowledge review 03: Types and quality of knowledge in social care, *Social Care Institute for Excellence (SCIE)*, November 2003
24. Peters, JL, Sutton, AJ, Jones, DR, Rushton, L, Abrams, KR, 2006, A systematic review of systematic reviews and meta-analyses of animal experiments with guidelines for reporting, *Journal of Environmental Science Health B* 41, 7, 1245–1258
25. Petticrew, M and Roberts, H, 2006, *Systematic Reviews in the Social Sciences*, Blackwell Publishing, Oxford, UK
26. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, 2017: <http://www.prisma-statement.org>
27. RAMESES (Realist And Meta-narrative Evidence Syntheses: Evolving Standards) project, 2017: <http://www.ramesesproject.org>
28. Rehfuss, EA, Puzzolo, E, Stanistreet, D, Pope, D and Bruce, NG, 2014, Enablers and barriers to large-scale uptake of improved solid fuel stoves: a systematic review, *Environmental Health Perspectives* 122, 2, 120-130
29. Research Councils UK Centre for Energy Epidemiology, 2018: <http://cee.ac.uk>
30. Saunders, M, Lewis, P and Thornhill, A, 2009, *Research Methods for Business Students*, 5th Edition, Harlow: Pearson Education Limited
31. Smith, V, Devane, D, Begley, CM and Clarke, M, 2011, Methodology in conducting a systematic review of systematic reviews of healthcare interventions, *BMC Medical Research Methodology* 11, 15, 1-6

32. Snilstveit, B, Oliver, S and Vojtkova, M, 2012, Narrative approaches to systematic review and synthesis of evidence for international development policy and practice, *Journal of Development Effectiveness* 4, 3, 409-429
33. Sorrell, S, 2007, Improving the evidence base for energy policy: the role of systematic reviews, *Energy Policy* 35, 1858-1871
34. UCL Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre), 2018, *Knowledge Library*, University College London, <https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=60>
35. UK Civil Service, 2014, Rapid Evidence Assessment Toolkit, <http://webarchive.nationalarchives.gov.uk/20140305122816/http://www.civilservice.gov.uk/networks/gsr/resources-and-guidance/rapid-evidence-assessment>
36. UK Department for Business, Energy and Industrial Strategy (BEIS), 2018: <https://www.gov.uk/government/organisations/department-for-business-energy-and-industrial-strategy>
37. UK Department for Business, Energy and Industrial Strategy (BEIS), 2016b, *Scoping Review of Heating Controls Project*, London: HMSO, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/573888/Final_Report_-_Heating_Controls_Scoping_Review_Project.pdf
38. UK Department for Business, Energy and Industrial Strategy (BEIS), 2017a, *Heating Controls: International Evidence Base and Policy Experiences*, London: HMSO, <https://www.gov.uk/government/policies/household-energy>
39. UK Department for Business, Energy and Industrial Strategy (BEIS), 2016a, *Occupancy Patterns Scoping Review Project*, London: HMSO, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/565281/Final_Report_-_Occupancy_Patterns_Scoping_Review_Project.pdf
40. UK Department for Business, Energy and Industrial Strategy (BEIS), 2017b, *Realising the Potential of Demand-Side Response to 2025*, London: HMSO, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/657140/DSR_research_Executive_Summary.pdf
41. UK Energy Research Centre (UKERC)'s Technology and Policy Assessment group, 2018: <http://www.ukerc.ac.uk/programmes/technology-and-policy-assessment.html>
42. University of Illinois, 2014, *Levels of evidence*, <http://ebp.lib.uic.edu/nursing/node/12>
43. Warm Homes and Energy Conservation Act of 2000, London: HMSO, UK
44. Warren, P, 2014, The use of systematic reviews to analyse demand-side management policy, *Energy Efficiency* 7, 417-427
45. Warren, P, 2015, *Demand-Side Management Policy: Mechanisms for Success and Failure*, PhD Thesis, University College London (UCL), UK

46. Warren, P, 2017, Transferability of demand-side policies between countries, *Energy Policy* 109, 757-766
47. Watson, J, Byrne, R, Morgan Jones, M, Tsang, F, Opazo, J, Fry, C and Castle-Clarke, S, 2012, What are the major barriers to increased use of modern energy services among the world's poorest people, and are interventions to overcome these effective? Project Report, Collaboration for Environmental Evidence (CEE)