

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/envsci

Review

On the use of systematic reviews to inform environmental policies



Gary S. Bilotta^{a,b,*}, Alice M. Milner^{b,c}, Ian Boyd^{b,d}

^a School of Environment and Technology, University of Brighton, Brighton BN2 4GJ, UK

^b Department for Environment, Food and Rural Affairs, London SW1P 3JR, UK

^c Department of Geography, University College London, London WC1E 6BT, UK

^d College Gate, University of St Andrews, Fife KY16 9AJ, UK

ARTICLE INFO

Published on line 18 June 2014

Keywords:

Systematic review

Literature review

Environmental science

Environmental policy

ABSTRACT

Environmental research varies in its methodological quality, degree of bias, and relevance to policy questions. Using this heterogeneous, and sometimes polarised, research to inform environmental policies can be challenging. Policy-making in the healthcare field sometimes uses systematic reviews (SRs) to tackle these issues and present a comprehensive, policy-neutral, transparent and reproducible synthesis of the evidence. However, there is less familiarity with SRs in the environmental field. The aim of this article is to: (1) summarise the process of conducting SRs, using best practice methods from the healthcare field as an example, (2) explain the rationale behind each stage of conducting a SR, and (3) examine the prospects and challenges of using SRs to inform environmental policy. We conclude that existing SR protocols from healthcare can be, and have been, applied successfully to environmental research but some adaptations could improve the process. The literature search stage could be expedited by standardising the reporting and indexing of environmental studies, equivalent to that in the healthcare field. The consistency of the study appraisal stage of SRs could be augmented by refining the existing quality assessment tools used in the healthcare field, enhancing their ability to discriminate quality and risk of bias in non-randomised studies. Ultimately, the strength of evidence within SRs on environmental topics could be improved through more widespread use of randomised controlled trials as a research method, owing to their inherently lower risk of bias when conducted according to best practice.

© 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/3.0/>).

* Corresponding author at: School of Environment and Technology, University of Brighton, Brighton BN2 4GJ, UK. Tel.: +44 01273 643318. E-mail address: G.S.Bilotta@Brighton.ac.uk (G.S. Bilotta).

<http://dx.doi.org/10.1016/j.envsci.2014.05.010>

1462-9011/© 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/3.0/>).

1. Introduction

Environmental research varies in its methodological quality,¹ degree of bias,² and relevance to policy. Using this heterogeneous, and sometimes polarised, research to inform environmental policies can be a challenging task, which at present is often first approached through the use of narrative literature reviews (Boyd, 2013). It is recognised that these types of literature reviews are vulnerable to author bias, which can occur when the review authors intentionally or unintentionally select or emphasise research according to their own opinions, prejudices or commercial interests (Higgins and Green, 2011). Furthermore, narrative literature reviews rarely consider, in a reproducible and meaningful manner, the methodological quality, degree of bias, and therefore reliability of the primary studies that are cited. These features of narrative literature reviews could lead to ill-informed environmental policies.

In evidence-based policy-making in the healthcare field, systematic review (SR) processes are used in order to tackle these issues, helping to present a comprehensive, policy-neutral, transparent and reproducible synthesis of the evidence. These SR processes are exemplified by the activities of the Cochrane Collaboration (<http://www.cochrane.org>); an international network of more than 31,000 researchers and practitioners (a mix of volunteers and paid staff who are affiliated to the organisation), from over 120 countries, who work to help healthcare practitioners, policy-makers, patients, their advocates and carers, make well-informed decisions about healthcare, by preparing, updating, and promoting the accessibility of SRs on the effectiveness of healthcare interventions. The Cochrane Collaboration have published over 5000 SRs so far, all of which are freely available online in the *Cochrane Database of Systematic Reviews*, which is part of *The Cochrane Library* (<http://www.cochrane.org/cochrane-reviews/about-cochrane-library>).

There is a common belief outside of healthcare, however, that SRs intrinsically adopt a biomedical model that is of relevance only to medicine, for example only capable of using randomised controlled trials (RCTs) and only capable of answering certain types of questions (Petticrew, 2001). As demonstrated in this article, this belief is unjustified. The practices of the Cochrane Collaboration have spurred the development of another international initiative; the Campbell Collaboration (<http://campbell.gse.upenn.edu>), who prepare, maintain, and disseminate SRs on the effectiveness of social and behavioural interventions in education, social welfare, and crime and justice (Davies and Boruch, 2001). More recently, these practices have spurred the founding of the Collaboration for Environmental Evidence – CEE

(<http://www.environmentalevidence.org/>); an open community of scientists and managers who, from their initial centres in Australia, South Africa, Sweden and the UK, have started to prepare SRs on environmental topics. Nevertheless, at present many environmental researchers, practitioners and policy-makers are typically less familiar with exactly what a SR involves, and often have major misconceptions about their history and purpose (Petticrew, 2001). The aim of this article is to: (1) summarise the process of conducting a SR, using the Cochrane Collaboration's exemplary methodology as an example (<http://handbook.cochrane.org/>), (2) explain the rationale behind each stage of the process, and (3) examine the prospects and challenges of using SRs to inform environmental policies.

2. The process of conducting a Cochrane systematic review

The key stages of producing a Cochrane systematic review (CSR), as described in the Cochrane Handbook (<http://handbook.cochrane.org/>), are illustrated in Fig. 1 and are summarised and compared to traditional literature reviews in Table 1:

2.1. The rationale behind each stage of a Cochrane systematic review

2.1.1. Formulating a question

As with any research, the first and most important decision in preparing a CSR is to determine its focus (O'Connor et al., 2011). This is best done by clearly framing the questions the review seeks to answer. Well-formulated questions will guide many aspects of the review process, including determining eligibility criteria, searching for studies, collecting data from included studies, and presenting findings (Jackson, 1980; Cooper, 1984; Hedges, 1994). In CSRs, questions are stated broadly as review 'Objectives', and specified in detail as 'Criteria for considering studies for this review' (O'Connor et al., 2011). A statement of the objectives typically begins with a precise statement of the primary objective, normally in the format of a single sentence. For example, for CSRs this may take the form: 'To assess the effects of [treatment, intervention or comparison] for [health problem] in [types of people, disease or problem and setting if specified]'. This might be followed by one or more secondary objectives, relating to different participant groups, different comparisons of interventions or different outcome measures (O'Connor et al., 2011). As this example suggests, the detailed specification of the review question requires consideration of several key components (Richardson et al., 1995; Counsell, 1997), including the types of populations (or participants), types of interventions and comparisons, and the types of outcomes that are of interest (PICO – Participants, Interventions, Comparisons and Outcomes) (O'Connor et al., 2011). As well as focussing review conduct, the contents of these sections are used by readers in their initial assessments of whether the review is likely to be directly relevant to the issues they face (O'Connor et al., 2011).

Systematic reviews are likely to be more relevant to the end-user and of higher quality if the initial questions and the

¹ Methodological quality is the term used to describe the extent to which a study's design, conduct and analysis have minimised selection, measurement and confounding bias (West et al., 2002, p. 2). Some authors argue that a more complete definition should also include external validity, appropriateness of statistical analyses, and use of ethical procedures (Berlin and Rennie, 1999).

² Bias is the term used to describe a systematic error or deviation in results or inferences from the truth.

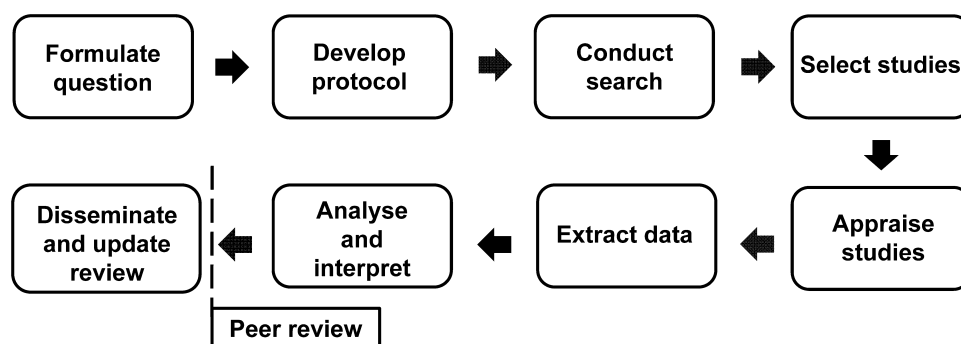


Fig. 1 – Key stages of conducting a CSR.

protocols (see Section 2.1.3) are informed by advice from stakeholders with a range of experiences, in terms of both the topic and the methodology (Khan et al., 2001; Rees et al., 2004; Thomas et al., 2004). The contribution of consulted stakeholders during the development of a review question and

protocol, should be documented in the Acknowledgements section of the protocol or review (Green and Higgins, 2011). Titles for CSRs are agreed by and registered with Cochrane Review Groups, who then oversee the process from publishing the protocol to publishing the final review.

Table 1 – A comparison of the features of CSRs (<http://handbook.cochrane.org/>) and traditional narrative literature reviews.

Stage	Cochrane systematic reviews	Narrative reviews
Formulating a question	Start with clear question to be answered or hypothesis to be tested (relating to specific populations, interventions, comparisons and outcomes)	May start with clear question to be answered, but they more often involve general discussion of the subject with no stated hypothesis
Developing a protocol	Protocol to be used is established and documented in advance, prior to knowledge of the available studies. This reduces the impact of review authors' bias, promotes transparency of methods and processes, reduces the potential for duplication, allows peer review of the planned methods, and enables easy maintenance of reviews in the light of new findings	Do not normally follow a pre-published protocol
Conducting the search	Strive to locate all relevant published and unpublished studies to limit the impact of publication and other biases	Do not usually attempt to locate all relevant literature. Often focus on published studies only
Selecting studies	Involve explicit description of what types of studies are to be included to limit selection bias on behalf of author(s)	Usually do not describe why certain studies are included and others are excluded
Appraising studies	Examine in a systematic and unbiased manner, the methods used in the primary studies, and investigate potential biases in those studies and sources of heterogeneity between study results	Often do not consider differences in study methods or study quality
Extracting data for analysis	Data analyses may be narrative, such as a structured summary and discussion of the studies' characteristics and findings, or quantitative, that is involving statistical meta-analysis, supported by Cochrane's review writing software	Typically limited to narrative analyses
Interpreting the synthesis	Base their conclusions on those studies which are most methodologically sound, and present a policy-neutral summary of the body of evidence, supported by Cochrane's review writing software	Often do not differentiate between methodologically sound and unsound studies. Sometimes present a policy-aligned summary of the body of evidence.
Disseminating and maintaining the review	Written by more than one author and peer-reviewed by a number of experts. They are then published online and are sometimes also co-published in peer-reviewed healthcare journals. They must remain free for dissemination in any and all media. Authors are committed to maintaining and updating these reviews, at least every two years, on the <i>Cochrane Library</i>	Written by one or more authors, and are sometimes peer-reviewed by experts, but are not always available as open access articles, and are not updated on a central database in the light of new findings

Adapted from Petticrew (2001).

2.1.2. Application to environmental research

It is not difficult to see how the specified PICO components of CSR questions could be adapted for SRs on environmental topics. Human participants may be replaced with specified animal or plant populations, habitats, ecosystems or members of society; healthcare interventions may be replaced with environmental management options such as the use of different agricultural techniques or different plant/animal disease control measures; comparisons between healthcare interventions (e.g. comparing the effectiveness of a pharmaceutical drug against a placebo), may be replaced with environmental studies using before-and-after approaches or interventions versus control experiments; human health outcomes may be replaced with metrics of animal, plant, or ecosystem health or productivity, or even social outcomes. Indeed, the CEE (2013) guidelines for conducting SRs on environmental topics, recommend adopting the PICO approach to question formulation. Policy-based questions may be a starting point to guide SRs on environmental topics, though those carrying out reviews need to be satisfied that the policy development process that has been the source of questions is robust. There may be reasons for challenging policy-makers to understand the rationale for certain questions before embarking upon a particular SR.

2.1.3. Developing a protocol

Preparing a CSR is complex and involves many judgements; in order to minimize the potential for bias in the review process, these judgements should be made in ways that do not depend on the findings of the studies included in the review (Green and Higgins, 2011). Since CSRs are by their nature retrospective, it is important that the methods to be used should be established and documented in advance (Green and Higgins, 2011). Publication of a protocol for a review prior to knowledge of the available studies reduces the impact of review authors' bias, promotes transparency of methods and processes, reduces the potential for duplication, and allows for stakeholder engagement/review of the planned methods (Light and Pillemer, 1984).

While the intention is that reviews will adhere to the published protocol, changes in a review protocol are sometimes necessary (Green and Higgins, 2011). It is important, however, that these changes should not be made on the basis of how they affect the outcome of the research study (Green and Higgins, 2011). Post hoc decisions made when the impact on the results of the research is known, such as excluding selected studies from a SR, are highly susceptible to bias and should be avoided (Green and Higgins, 2011).

In the case of CSRs, protocols are published before the completed SR in the *Cochrane Database of Systematic Reviews* (CDSR) (Green and Higgins, 2011). Changes in the protocol are documented and reported in the 'Differences between protocol and review' section of the completed review, and sensitivity analyses exploring the impact of deviations from the protocol are undertaken when possible (Green and Higgins, 2011).

2.1.4. Application to environmental research

It is not difficult to imagine how authors might prepare a Cochrane-style protocol for a SR on an environmental topic.

Indeed, the CEE (2013) provide a template for developing, registering and publishing a SR protocol. Their latest guidance can be found at www.environmentalevidence.org/Instructionsforauthors.html. For examples of almost 100 protocols produced at the time of writing this article, visit the CEE Library (www.environmentalevidence.org/Library.htm). For examples of reviews that are in progress, visit: <http://www.environmentalevidence.org/Reviewsinprogress.html>. The protocols produced so far by the CEE, vary in their detail and level of specification, partly reflecting the fact that environmental researchers are, at present, less familiar with the SR process. It would be advantageous for the CEE and wider environmental community to build expertise in protocol registration. Ideally, those who oversee the process should be specialists in this skill, but not have a detailed knowledge of the evidence available to answer the question. This will help to eliminate potential bias resulting from experts directing the review in a certain direction, based on their existing knowledge of the topic.

2.1.5. Conducting the search

Literature searches for CSRs aim to be as extensive as possible to ensure that as many as possible of the relevant studies are included in the review (Lefebvre et al., 2011). It is, however, necessary to strike a balance between striving for comprehensiveness and maintaining relevance when developing a search strategy (Lefebvre et al., 2011). Increasing the comprehensiveness or sensitivity of a search,³ will reduce its precision,⁴ and will retrieve more non-relevant articles (Lefebvre et al., 2011).

The Cochrane Handbook provides detailed guidance on developing a search strategy for a CSR. The search strategy for a CSR (search terms to be used, databases to be searched, etc.) is described in its review protocol, though searching can be an iterative process in which the terms that are used are modified, based on what has already been retrieved (Lefebvre et al., 2011). There are diminishing returns for search efforts; after a certain stage, each additional unit of time invested in searching returns fewer references that are relevant to the review (Lefebvre et al., 2011). Consequently there comes a point where the rewards of further searching may not be worth the effort required to identify the additional references (Lefebvre et al., 2011). The decision as to how much to invest in the search process depends on the question the review addresses and the resources that are available. Lefebvre et al. (2011) suggest that at a conservatively estimated reading rate of two abstracts per minute, the results of a database search can be 'scan-read' at the rate of 120 per hour; so the high yield and low precision associated with CSR searching is not as daunting as it might at first appear.

In CSRs the full final search strategies used for each database searched are included in an appendix of the CSR, so all search strategies should be saved, and notes taken of the

³ Comprehensiveness or sensitivity is the number of relevant reports identified divided by the total number of relevant reports in existence.

⁴ Search precision is the number of relevant reports identified divided by the total number of reports identified.

number of records retrieved for each database searched (Lefebvre et al., 2011). This ensures that the search procedure is transparent, auditable, and reproducible.

2.1.6. Application to environmental research

The literature search processes used in CSRs were originally designed with a focus on evidence derived from RCTs examining the effectiveness of healthcare treatments. At present, non-randomised studies (NRSs) are a more common study methodology in environmental science.⁵ These types of studies could be more time-consuming to search for during the literature search and selection stage of a SR on environmental topics. For example, when a CSR aims to include RCTs only, various approaches are available to restrict the search strategy to RCTs, including:

- I. search for previous reviews of the review question,
- II. use resources such as CENTRAL or Cochrane Review Group-specific registers that are 'rich' in RCTs,
- III. use methodological filters and indexing fields, such as publication type in MEDLINE, to limit searches to studies that are likely to be RCTs, and
- IV. search trial registers.

However, to restrict the search to particular NRS designs is more difficult; study design labels are not used consistently by authors and are not indexed reliably by journals or bibliographic databases. Search results thus often contain large numbers of irrelevant citations and abstracts often do not provide adequate detail about NRS design. Therefore, unlike the situation when reviewing studies from RCTs, it may be necessary to obtain and read many full reports in order to identify eligible studies. This challenge does not prevent NRSs from being included in SRs. Indeed, the Cochrane Collaboration have produced a number of reviews which include evidence from NRSs, including qualitative studies and economic data. The literature search and study eligibility assessments may, however, take longer to complete than they would with a review based solely on RCTs.

Authors, publishers and hosts of bibliographic databases could all contribute to improving the reporting and indexing of environmental NRSs so that they are easier to search and check against eligibility criteria in the future. In evidence-based medicine, this standardisation of reporting and indexing of research began in 1996 when an international group of epidemiologists, statisticians, clinical trialists, and medical editors, some of whom were involved with establishing the Cochrane Collaboration, published the CONSORT statement; a checklist of items to be addressed in a report of the findings of RCTs (Turner et al., 2013). CONSORT has twice been revised and updated over time, and the impact has been noted as one of the major milestones on health research methods over the

last century (Gabriel and Normand, 2012). Environmental science would benefit from developing an environmental equivalent of the CONSORT statement for NRSs.

In the meantime, the CEE (2013) provide detailed guidance on how best to conduct literature searches for environmental studies.

2.1.7. Selecting the eligible studies

The findings of a SR depend critically on decisions relating to which studies are included, and on decisions relating to which data from these studies are presented and analysed (Higgins and Deeks, 2011). The methods used for these decisions must be transparent, and they should be chosen to minimize biases and human error. A CSR is a review of studies that meet pre-specified criteria for inclusion in the review (Higgins and Deeks, 2011). Since each of the studies discovered from the literature search stage may have been reported in several articles, abstracts or other reports, a comprehensive search for studies for the review may identify many reports from potentially relevant studies (Higgins and Deeks, 2011). Two distinct processes are therefore used to determine which studies can be included in the review: one is to link together multiple reports of the same study; and the other is to use the information available in the various reports to determine which studies are eligible for inclusion (Higgins and Deeks, 2011). Although sometimes there is a single report for each study, it should never be assumed that this is the case as this could introduce substantial biases if studies are inadvertently included more than once in a meta-analysis (Tramèr et al., 1997).

A typical process for selecting studies for inclusion in a CSR is as follows: (I) merge search results using reference management software, and remove duplicate records of the same report. (II) Examine titles and abstracts to remove obviously irrelevant reports (authors should generally be over-inclusive at this stage). (III) Retrieve full text of the potentially relevant reports. (IV) Link together multiple reports of the same study. (V) Examine full-text reports for compliance of studies with eligibility criteria. (VI) Correspond with investigators, where appropriate, to clarify study eligibility (it may be appropriate to request further information, such as missing results, at the same time). (VII) Make final decisions on study inclusion and proceed to data collection (Higgins and Deeks, 2011).

Decisions about which studies to include in a review are among the most influential decisions that are made in the review process (Higgins and Deeks, 2011). However, they involve judgement – to help ensure that these judgements are reproducible in CSRs, more than one author repeats the process independently, and this is overseen by the Cochrane Review Group. Using at least two authors, searching independently, reduces the possibility that relevant reports will be discarded (Edwards et al., 2002).

Experts in a particular area frequently have pre-formed opinions that can bias their assessments of both the relevance and validity of articles (Cooper and Ribble, 1989; Oxman and Guyatt, 1993). Thus while it is important that at least one author is knowledgeable in the area under review, it is an

⁵ The Cochrane Collaboration define NRSs as quantitative and/or qualitative studies that do not use randomisation to allocate units to comparison groups. These include observational studies, cohort studies, case-control studies, controlled before-and-after studies, interrupted-time-series studies, and controlled trials that use quasi-randomised design.

advantage to have a second author who is not a content expert (Higgins and Deeks, 2011). Disagreements about whether a study should be included can generally be resolved by auditable discussion (Higgins and Deeks, 2011). Often the cause of disagreement is a simple oversight on the part of one of the review authors, but when the disagreement is due to a difference in interpretation, this may require auditable arbitration by another reviewer (Higgins and Deeks, 2011).

2.1.8. Application to environmental research

Study selection methods used in CSRs could be used, with little or no amendment, in SRs on environmental topics. The CEE (2013) suggest that it is good practice at the beginning of the abstract relevance assessment stage for two reviewers to undertake the same process on a random sub-sample of articles from the original list (the recommended sample is a minimum of 50 articles or 10% up to a maximum of 200 references). To check for consistency in the interpretation of the selection criteria, reviewer relevance decisions can be compared by performing a kappa analysis, which adjusts the proportion of records for which there was agreement, by the amount of agreement expected by chance alone. A kappa rating of 'substantial' (>0.5) is recommended to pass the assessment.

For the latest guidance on how to conduct the study selection stage for environmental SRs, visit www.environmentalevidence.org/Instructionsforauthors.html.

Inconsistencies among reviewers at this stage of the SR, could change the outcome of the review. Owing to the importance of decisions made at this stage, it would be sensible for the CEE to enforce the independent assessment of reproducibility of this stage, rather than just recommending it as good practice. Likewise, it would be worthwhile enforcing that at least one of the members responsible for judgements of study eligibility, is a non-expert on the review topic.

2.1.9. Appraising the selected studies

This stage of a CSR is designed to ensure that the review authors are cognisant of the potential biases within primary studies and of how such biases could impact review results and subsequent conclusions (Higgins et al., 2011a). Cochrane SRs assess the methodological quality of primary studies through use of an objective system developed by the Grades of Recommendation, Assessment, Development and Evaluation (GRADE) Working Group (GRADE Working Group, 2004; Schünemann et al., 2006; Guyatt et al., 2008a,b). This approach is now used by the World Health Organisation (WHO) and the UK National Institute for Health and Care Excellence (NICE) among 20 other bodies internationally. The GRADE approach specifies four levels of quality (High, Moderate, Low, and Very Low). The highest quality rating is for evidence from RCTs. Review authors can, however, downgrade evidence from RCTs to moderate, low, or even very low quality evidence, depending on the presence of the five factors, including limitations in the design and implementation of available studies suggesting high risk of bias; indirectness of evidence; unexplained heterogeneity or inconsistency of results; imprecision of results; and high

probability of publication bias (Higgins et al., 2011a). Review authors will generally grade evidence from sound observational studies as low quality. If however, such studies yield large effects and there is no obvious risk of bias explaining those effects, review authors may rate the evidence as moderate or, if the effect is large enough, even high quality (Higgins et al., 2011a). The very low quality level includes, but is not limited to, studies with critical problems and unsystematic clinical observations (e.g. case series or case reports) (Higgins et al., 2011a).

The risk of bias within studies is assessed through the use of the Cochrane Collaboration's Risk of Bias Tool (Higgins et al., 2011a), which helps review authors to identify potential biases and support judgements of the degree to which they may have influenced the findings of each study. As part of the SR process, authors record their judgments of the risk of bias and provide evidence for each potential source of bias. Through the combined use of GRADE and the Risk of Bias Tool, review authors are guided in their evaluation of each of the included studies. This increases the transparency and auditability of evidence appraisal stage, reducing the potential for authors' bias to influence the conclusions of the review, while helping the authors to discover the consistencies and account for the variability in similar appearing studies through accounting for potential biases in the primary research (Cooper and Hedges, 1994). At least two authors assess the risk of bias within primary studies, and these authors are double-blinded before agreeing a final assessment.

While the Cochrane Collaboration prefer to use evidence derived from RCTs, owing to the lower potential for bias (when conducted according to best practice), the group have produced many SRs based on evidence derived from NRSs. Some review authors have tried to develop, optimise and 'validate' search strategies for NRS (Wieland and Dickersin, 2005; Fraser et al., 2006; Furlan et al., 2006; Golder et al., 2006a,b), and there is a dedicated NRS Methods Group who provide guidance to support authors who are considering including NRS in CSRs (Higgins and Green, 2011). In principle, the assessment of the risk of bias in NRSs is exactly the same as it is for RCTs, but the Collaboration advise that review authors must pay extra attention to the weaknesses of the designs that have been used (such as noting their potential to ascertain causality); the execution of the studies through careful consideration of their risk of bias, especially the potential for selection bias and confounding factors to which all NRSs are suspect; and the potential for reporting biases, including selective reporting of outcomes owing to the lack of study registration systems (Higgins and Green, 2011).

Deeks et al. (2003) noted that there were at least 194 existing tools (scales which score the studies based on a number of weighted criteria; and checklists which assess studies against criteria without producing a score), that could be or have been used to assess methodological quality of NRSs. Until relatively recently, CSRs used a variety of these tools (Lundh and Gøtzsche, 2008). The Cochrane Collaboration's current recommended tool for assessing risk of bias, however is neither a scale nor a checklist. It is a domain-based evaluation, in which critical assessments are made separately

for different domains of bias (or aspects of research design).⁶ These domains were selected on the basis of empirical evidence linking them to biased findings (Wood et al., 2008; Gluud, 2006). The approach was developed between 2005 and 2007 by a working group of methodologists, editors and review authors. Because it is impossible to know the extent of bias (or even the true risk of bias) in a given study, the possibility of validating any proposed tool is limited (Higgins and Green, 2011). Experiences in application of the first (2008) version of the tool led to some criticisms concerning its ease of use and reliability (i.e. inter-rater consistency) (Hartling et al., 2009, 2012). In response to these criticisms, a revised version was published in 2011, and a working group has since been established to continue development of the Risk of Bias Tool, with version 2.0 due to be released in 2014 (Turner et al., 2013). It is also accepted that issues of study design arise when using the tool to assess risk of bias in NRSs, and therefore an ongoing Cochrane Methods Innovation Fund project will lead to the release of a new version of the tool for assessing NRSs (Turner et al., 2013). Nevertheless, the use of alternative scales for assessing quality or risk of bias is explicitly discouraged in CSRs. The Collaboration argue that while scales offer appealing simplicity; theoretical (Greenland and O'Rourke, 2001), and empirical evidence (Juni et al., 1999), suggests that their associations with intervention effect estimates are inconsistent and unpredictable (Balk et al., 2002; Emerson et al., 1990; Schulz et al., 1995). Furthermore, calculating a summary score inevitably involves assigning 'weights' to different items in the scale, in ways that are difficult to justify (Higgins et al., 2011b).

2.1.10. Application to environmental research

Practitioners conducting SRs on environmental issues should consider testing, modifying, and adopting the quality assessment tools and risk of bias tools developed by the Cochrane Collaboration; capitalising on more than twenty year's of theoretical and empirical research that has been invested in these tools, while attempting to enhance the ability of the tools to discriminate between different levels of quality within NRSs, and widening the possible sources of bias considered in the Risk of Bias Tool to make it more relevant to environmental research.

That said, while NRSs are the current *modus operandi* for environmental research, RCTs could and should be used more often as a means of addressing research hypotheses. The RCT is widely regarded as the design of choice for the assessment of the effectiveness of interventions in healthcare, and this is

such for a reason. The main benefit of the RCT is the use of a randomisation procedure that, when properly implemented, ensures that the allocation of any study unit to one intervention or another cannot be predicted. The randomisation process makes the comparison groups equal with respect to both known and unknown prognostic factors at baseline, apart from chance bias (D'Agostino and Kwan, 1995). RCTs also tend to benefit from so-called 'inherited properties', which generally mark them out as higher quality studies (Deeks et al., 2003). These properties include the fact that they are prospective studies, with written protocols specifying, and thus standardising, important aspects of study unit enrolment, interventions, observation and analysis (Abel and Koch, 1999). RCTs are also more likely to employ specific measures to reduce or remove bias, such as blinded outcome assessment (Deeks et al., 2003).

Ultimately the strength of the findings of a SR are determined by the quality and risk of bias in the primary studies cited. If environmental researchers were to use RCTs more widely as a research method (where appropriate and feasible), then it is likely that the environmental community would be able to increase the impact of their SRs. In the meantime, the environmental community will have to make use of the evidence currently available, which as mentioned above is often derived from NRSs. The choice of which quality appraisal tool to use to assess these studies is critical, as it has been empirically demonstrated that the use of different quality scales for the assessment of the same studies results in different estimates of quality (Moher et al., 1998; Juni et al., 1999).

At present, the CEE (2013) does not place restrictions on the use of existing checklists or critical appraisal tools as a basis for study appraisal, but requires that authors either explain why they used the chosen method as is (no modification, because not considered to be needed, and why), or adapted the method for their SR (in which case the decisions made must be stated and justified). They suggest that review-specific *a priori* assessment criteria for appraising the quality of methodology should be included in the SR protocol, and that two or more assessors should be used for study appraisal. The environmental community should attempt to determine the most suitable tool for quality appraisal in environmental SRs. This would simplify the SR process and reduce the potential for authors to select a quality appraisal tool that emphasises research that meets their own opinions, prejudices or commercial interests.

2.1.11. Extracting data for analysis and interpretation

Analyses within CSRs may be narrative, such as a structured summary and discussion of the studies' characteristics and findings, or quantitative, that is involving statistical analysis (Deeks et al., 2011). Meta-analysis – the statistical combination of results from two or more separate studies – is the most commonly used statistical technique. Cochrane review writing software (RevMan) can perform a variety of meta-analyses, although it is stressed that meta-analysis is not appropriate in all CSRs (Deeks et al., 2011).

In CSRs the analysis plan follows from the scientific aim of the review. Reviews have different types of aims, and may therefore contain different approaches to analysis (Deeks et al., 2011). The most straightforward CSR assembles studies that

⁶ This tool classifies potential biases into *selection bias* (in the case of clinical trials this refers to systematic differences between baseline characteristics of the groups that are to be compared), *performance bias* (in the case of clinical trials this refers to systematic differences between groups in the care that is provided, or in exposure to factors other than the interventions of interest), *attrition bias* (in the case of clinical trials this refers to systematic differences between groups in withdrawals from a study), *detection bias* (in the case of clinical trials this refers to systematic differences between groups in how outcomes are determined), and *reporting bias* (in the case of clinical trials this refers to systematic differences between reported and unreported findings), and *other biases*.

make one particular comparison between two treatment options. Meta-analysis and related techniques can be used if there is a consistent outcome measure to (I) establish whether there is evidence of an effect, (II) estimate the size of the effect and the uncertainty surrounding that size, and (III) investigate whether the effect is consistent across studies.

Some reviews may have a broader focus than a single comparison. The first is where the intention is to identify and collate studies of numerous interventions for the same disease or condition. The second, related aim is that of identifying a 'best' intervention. Such reviews may include multiple comparisons and meta-analyses between all possible pairs of treatments, and require care when it comes to planning analyses (Deeks et al., 2011).

Occasionally review comparisons have particularly wide scopes that make the use of meta-analysis problematic. When reviews contain very diverse studies a meta-analysis might be useful to answer the overall question of whether there is evidence that a particular intervention can work. But use of meta-analysis to describe the size of effect may not be meaningful if the implementations are so diverse that an effect estimate cannot be interpreted in any specific context (Deeks et al., 2011).

An aim of some CSRs is to investigate the relationship between the size of an effect and some characteristic(s) of the studies (Deeks et al., 2011). This is uncommon as a primary aim in CSRs, but may be a secondary aim.

2.1.12. Application to environmental research

Similar considerations influence decisions on whether the synthesis should be narrative or quantitative in SRs on environmental topics. For the latest CEE guidance on this stage of conducting a SR on environmental topics, visit: www.environmentalevidence.org/Instructionsfor-authors.html. An example of an environmental SR with an entirely narrative synthesis (Davies et al., 2006) and a narrative synthesis that complements a quantitative synthesis (Bowler et al., 2010) are available in the CEE Library (<http://www.environmentalevidence.org/Reviews.html>). Open access data is now becoming more prevalent in environmental research, and this will increase the possibilities for SRs with quantitative meta-analyses in the future.

2.1.13. Disseminating and updating the review

CSRs are written by more than one author and are peer-reviewed by a number of experts (Green and Higgins, 2011). They are then published online and are sometimes also co-published in peer-reviewed healthcare journals (Green and Higgins, 2011). For the Cochrane Collaboration, there is one essential condition of co-publication: CSRs must remain free for dissemination in any and all media, without restriction from any of them (Green and Higgins, 2011). Since evidence on a given subject is generally dynamic and continually evolving, incorporating additional studies as they become available can change the results of a SR (Chalmers and Haynes, 1994). Therefore, SRs that are not maintained run the risk of becoming out of date and even misleading. An important feature of CSRs is that review authors are committed not only to preparing SRs of evidence, but also to maintaining (and

updating) these reviews on a regular basis (at least every two years) on the *Cochrane Library* (Green and Higgins, 2011).

2.1.14. Application to environmental research

The environmental equivalent to the Cochrane Collaboration (i.e. the CEE) was founded in 2008 (Pullin and Knight, 2013), and an online open-access library of SRs was created in 2012 to enable widespread dissemination. This is currently a relatively small but growing library of SRs on environmental topics. Policies need to be established to ensure that these reviews are maintained and updated regularly. If society is to move towards better-informed environmental policies for a sustainable global environment and the conservation of biodiversity, these SR activities need to be accelerated urgently (Pullin and Knight, 2013).

3. Conclusions

Systematic reviews are powerful tools that aim to provide comprehensive and reproducible summaries of evidence to guide policy decisions. They employ a range of methods that are designed to reduce the influence of author bias while considering bias in the primary studies. This article summarised how to conduct SRs according to best practice in the healthcare field, and explained the rationale for each stage of conducting a SR. It has demonstrated that existing CSR methods can be, and already have been, used in SRs on environmental topics. For example, to date the CEE have published over 60 SRs, with a further 30 SRs in progress. These SRs, which are all available from the CEE Library, cover a range of topics including pure environmental science questions such as 'What is the evidence for glacial shrinkage across the Himalayas?' (Miller et al., 2013), applied environmental management topics such as 'Evaluating the biological effectiveness of fully and partially protected marine areas' (Sciberras et al., 2013), and human–environment interaction questions such as 'What is the evidence that scarcity and shocks in freshwater resources causes conflict instead of collaboration?' (Johnson et al., 2011). We suggest, however that the process of conducting a SR on an environmental topic could be improved through several adaptations of both the SR process, and the manner in which environmental research is conducted, reported and indexed in the future. The literature search stage could be expedited by producing a statement of guidelines to standardise the reporting and indexing of environmental studies, equivalent to the CONSORT statement in healthcare. The consistency of the study appraisal stage of SRs could be augmented by refining the existing quality assessment tools used in the healthcare field, enhancing their ability to discriminate quality and risk of bias in NRSs. Ultimately, the strength of evidence within SRs on environmental topics could be improved through more widespread use of RCTs as a research method, owing to their inherently lower risk of bias when conducted according to best practice.

Society may be on the cusp of an evidence revolution in environmental management but it will take new contributors and investment to ensure that this has impact. The CEE recently proposed a five-year programme to build capacity for the conduct and use of SRs in the environment sector (Pullin

and Knight, 2013). This programme aims to: (1) increase the commissioning and use of SR in evidence-based policy; (2) develop the capacity of the global environmental research community to conduct SRs; and (3) develop the capacity of CEE to co-ordinate and promote the conduct of SRs in the environmental sector. All of these plans require a big effort on the part of those already active in CEE, but also provide opportunities for others to join in and contribute to the growth of the network (Pullin and Knight, 2013).

Acknowledgements

This work arises from Natural Environment Research Council funding (Grant NE/L00836X/1) in collaboration with the United Kingdom's Department for Environment, Food and Rural Affairs.

REFERENCES

- Abel, U., Koch, A., 1999. The role of randomization in clinical studies: myths and beliefs. *J. Clin. Epidemiol.* 52, 487–497.
- Balk, E.M., Bonis, P.A.L., Moskowitz, H., Schmid, C.H., Ioannidis, J.P.A., Wang, C., Lau, J., 2002. Correlation of quality measures with estimates of treatment effect in meta-analyses of randomized controlled trials. *J. Am. Med. Assoc.* 287 (22) 2973–2982.
- Berlin, J.A., Rennie, D., 1999. Measuring the quality of trials: the quality of quality scales. *J. Am. Med. Assoc.* 282 (11) 1083–1085.
- Bowler, D., Buyung-Ali, L., Knight, T., Pullin, A.S., 2010. How effective is 'greening' of urban areas in reducing human exposure to ground level ozone concentrations, UV exposure and the 'urban heat island effect'? *Environ. Evid.*, www.environmentalevidence.org/SR41.html.
- Boyd, I., 2013. Research: a standard for policy-relevant science. *Nature* 501, 159–160.
- Collaboration for Environmental Evidence, 2013. Guidelines for Systematic Review and Evidence Synthesis in Environmental Management. Version 4.2. Environmental Evidence. , www.environmentalevidence.org/Documents/Guidelines/Guidelines4.2.pdf.
- Chalmers, I., Haynes, B., 1994. Reporting, updating, and correcting systematic reviews of the effects of health care. *Br. Med. J.* 309 (6958) 862–865.
- Cooper, H., Ribble, R.G., 1989. Influences on the outcome of literature searches for integrative research reviews. *Knowl.: Creation Diffus. Util.* 10, 179–201.
- Cooper, H.M., 1984. The problem formulation stage. In: Cooper, H.M. (Ed.), *Integrating Research: A Guide for Literature Reviews*. Sage Publications, Newbury Park, CA.
- Cooper, H., Hedges, L. (Eds.), 1994. *The Handbook of Research Synthesis*. Russell Sage Foundation, New York.
- Counsell, C., 1997. Formulating questions and locating primary studies for inclusion in systematic reviews. *Ann. Intern. Med.* 127 (5) 380–387.
- D'Agostino, R.B., Kwan, H., 1995. Measuring effectiveness: what to expect without a randomized control group. *Med. Care* 33 (4) 95–105.
- Davies, P., Boruch, R., 2001. The Campbell Collaboration: does for public policy what Cochrane does for health. *Br. Med. J.* 323 (7308) 294–295.
- Davies, Z.G., Tyler, C., Stewart, G.B., Pullin, A.S., 2006. Are Current Management Recommendations for Conserving Saproxylic Invertebrates Effective? CEE review 05-011 (SR17). Collaboration for Environmental Evidence. , www.environmentalevidence.org/SR17.html.
- Deeks, J.J., Dinnes, J., D'amico, R., Sowden, A.J., Sakaravitch, C., Song, F., Petticrew, M., Altman, D.J., 2003. Evaluating non-randomised intervention studies. *Health Technol. Assess.* 7 (27) 1–179.
- Deeks, J.J., Higgins, J.P.T., Altman, D.G., 2011. Analysing data and undertaking meta-analyses. In: Higgins, J.P.T., Green, S. (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* (updated March 2011). The Cochrane Collaboration, 2011. (Chapter 9).
- Edwards, P., Clarke, M., DiGiuseppi, C., Pratap, S., Roberts, I., Wentz, R., 2002. Identification of randomized controlled trials in systematic reviews: accuracy and reliability of screening records. *Stat. Med.* 21 (11) 1635–1640.
- Emerson, J.D., Burdick, E., Hoaglin, D.C., Mosteller, F., Chalmers, T.C., 1990. An empirical study of the possible relation of treatment differences to quality scores in controlled randomized clinical trials. *Control. Clin. Trials* 11 (5) 339–352.
- Fraser, C., Murray, A., Burr, J., 2006. Identifying observational studies of surgical interventions in MEDLINE and EMBASE. *Br. Med. Counc. Med. Res. Methodol.* 6, 41.
- Furlan, A.D., Irvin, E., Bombardier, C., 2006. Limited search strategies were effective in finding relevant nonrandomized studies. *J. Clin. Epidemiol.* 59 (12) 1303–1311.
- Gabriel, S.E., Normand, S.L., 2012. Getting the methods right – the foundation of patient-centered outcomes research. *N. Engl. J. Med.* 367 (9) 787–790.
- Gluud, L.L., 2006. Bias in clinical intervention research. *Am. J. Epidemiol.* 163 (6) 493–501.
- Golder, S., Loke, Y., McIntosh, H.M., 2006a. Room for improvement? A survey of the methods used in systematic reviews of adverse effects. *BMC Med. Res. Methodol.* 6, 3.
- Golder, S., McIntosh, H.M., Duffy, S., Glanville, J., 2006b. Centre for Reviews and Dissemination and UK Cochrane Centre Search Filters Design Group, Developing efficient search strategies to identify reports of adverse effects in MEDLINE and EMBASE. *Health Inform. Libr. J.* 23, 3–12.
- GRADE Working Group, 2004. Grading quality of evidence and strength of recommendations. *Br. Med. J.* 328, 1490–1494.
- Green, S., Higgins, J.P.T., 2011. Preparing a Cochrane review. In: Higgins, J.P.T., Green, S. (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* (updated March 2011). The Cochrane Collaboration, 2011. (Chapter 2).
- Greenland, S., O'Rourke, K., 2001. On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. *Biostatistics* 2, 463–471.
- Guyatt, G.H., Oxman, A.D., Kunz, R., Vist, G.E., Falck-Ytter, Y., Schünemann, H.J., 2008a. What is 'quality of evidence' and why is it important to clinicians? *Br. Med. J.* 336 (7651) 995–998.
- Guyatt, G.H., Oxman, A.D., Vist, G.E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., Schünemann, H.J., 2008b. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Br. Med. J.* 336 (7650) 924–926.
- Hartling, L., Ospina, M., Liang, Y., Dryden, D.M., Hooton, N., Seida, J.K., Klassen, T.P., 2009. Risk of bias versus quality assessment of randomised controlled trials: cross sectional study. *Br. Med. J.* 339, b4012.
- Hartling, L., Hamm, M.P., Milne, A., Vandermeer, B., Santaguida, P.L., Ansari, M., Tsertsvadze, A., Hempel, S., Shekelle, P., Dryden, D.M., 2012. Testing the Risk of Bias tool showed low reliability between individual reviewers and across consensus assessments of reviewer pairs. *J. Clin. Epidemiol.* 66 (9) 973–981.

- Hedges, L.V., 1994. Statistical considerations. In: Cooper, H., Hedges, L.V. (Eds.), *The Handbook of Research Synthesis*. Russell Sage Foundation, New York, NY.
- Higgins, J.P.T., Altman, D.G., Sterne, J.A.C., 2011a. Assessing risk of bias in included studies. In: Higgins, J.P.T., Green, S. (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* (updated March 2011). The Cochrane Collaboration, 2011. (Chapter 8).
- Higgins, J.P.T., Altman, D.G., Gøtzsche, P.C., Juni, P., Moher, D., Oxman, A.D., Savović, J., Schulz, K.F., Weeks, L., Sterne, J.A.C., 2011b. *The Cochrane Collaboration's tool for assessing risk of bias in randomised trials*. *Br. Med. J.* 343, d5928.
- Higgins, J.P.T., Deeks, J.J., 2011. Selecting studies and collecting data. In: Higgins, J.P.T., Green, S. (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* (updated March 2011). The Cochrane Collaboration, 2011. (Chapter 7).
- Higgins, J.P.T., Green, S., 2011. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
- Jackson, G.B., 1980. *Methods for integrative reviews*. *Rev. Educ. Res.* 50, 438–460.
- Johnson, V., Fitzpatrick, I., Floyd, R., Simms, A., 2011. What is the Evidence that Scarcity and Shocks in Freshwater Resources Cause Conflict Instead of Promoting Collaboration? CEE Review 10-010. Collaboration for Environmental Evidence. , In: www.environmentalevidence.org/SR10010.html.
- Juni, P., Witschi, A., Bloch, R., Egger, M., 1999. The hazards of scoring the quality of clinical trials for meta-analysis. *J. Am. Med. Assoc.* 282 (11) 1054–1060.
- Khan, K.S., ter Riet, G., Glanville, J., Sowden, A.J., Kleijnen, J., 2001. *Undertaking Systematic Reviews of Research on Effectiveness: CRD's Guidance for those Carrying Out or Commissioning Reviews (CRD Report Number 4)*, 2nd edition. NHS Centre for Reviews and Dissemination, University of York, York, UK.
- Lefebvre, C., Manheimer, E., Glanville, J., 2011. Searching for studies. In: Higgins, J.P.T., Green, S. (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* (updated March 2011). The Cochrane Collaboration, 2011. (Chapter 6).
- Light, R.J., Pillemer, D.B., 1984. *Summing Up: The Science of Reviewing Research*. Harvard University Press, Cambridge, MA.
- Lundh, A., Gøtzsche, P.C., 2008. Recommendations by Cochrane Review Groups for assessment of the risk of bias in studies. *Br. Med. Counc. Med. Res. Methodol.* 8, 22.
- Miller, J., Rees, G., Warnaars, T., Young, G., Shrestha, A., Collins, D., 2013. What is the Evidence About Glacier Melt Across the Himalayas. CEE Review 10-008. Collaboration for Environmental Evidence. , In: www.environmentalevidence.org/SR10008.html.
- Moher, D., Pham, B., Jones, A., Cook, D.J., Jadad, A.R., 1998. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 352 (9128) 609–613.
- O'Connor, D., Green, S., Higgins, J.P.T., 2011. Defining the review question and developing criteria for including studies. In: Higgins, J.P.T., Green, S. (Eds.), *Cochrane Handbook of Systematic Reviews of Intervention Version 5.1.0* (updated March 2011). The Cochrane Collaboration, 2011. (Chapter 5).
- Oxman, A.D., Guyatt, G.H., 1993. The science of reviewing research. *Ann. N. Y. Acad. Sci.* 703, 125–133.
- Petticrew, M., 2001. Systematic reviews from astronomy to zoology: myths and misconceptions. *Br. Med. J.* 322 (7278) 98–101.
- Pullin, A.S., Knight, T., 2013. Time to build capacity for evidence synthesis in environmental management. *Environ. Evid.* 2 (1) 21–22.
- Rees, R., Kavanagh, J., Burchett, H., Shepherd, J., Brunton, G., Harden, A., Thomas, S., Oakley, A., 2004. *HIV Health Promotion and Men who have Sex with Men (MSM): A Systematic Review of Research Relevant to the Development and Implementation of Effective and Appropriate Interventions*. EPPi-Centre, Social Science Research Unit, Institute of Education, University of London, London, UK.
- Richardson, W.S., Wilson, M.S., Nishikawa, J., Hayward, R.S.A., 1995. The well-built clinical question: a key to evidence based decisions. *ACP J. Club* A12–A13.
- Schulz, K.F., Chalmers, I., Hayes, R.J., Altman, D.G., 1995. Empirical evidence of bias, dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *J. Am. Med. Assoc.* 273 (5) 408–412.
- Schünemann, H.J., Jaeschke, R., Cook, D.J., Bria, W.F., El-Solh, A.A., Ernst, A., Fahy, B.F., Gould, M.K., Horan, K.L., Krishnan, J.A., Manthous, C.A., Maurer, J.R., McNicholas, W.T., Oxman, A.D., Rubenfeld, G., Turino, G.M., Guyatt, G., 2006. An official ATS statement: grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. *Am. J. Respir. Crit. Care Med.* 174 (5) 605–614.
- Sciberras, M., Jenkins, S., Kaiser, M., Hawkins, S., Pullin, A., 2013. Evaluating the biological effectiveness of fully and partially protected marine areas. *Environ. Evid.* 2 (4) , In: <http://www.environmentalevidencejournal.org/content/2/1/4>.
- Thomas, B.H., Ciliska, D., Dobbins, M., Micucci, S., 2004. A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews Evid.-Based Nurs.* 1, 165–184.
- Tramèr, M.R., Reynolds, D.J., Moore, R.A., McQuay, H.J., 1997. Impact of covert duplicate publication on meta-analysis: a case study. *Br. Med. J.* 315 (7109) 635–640.
- Turner, L., Boutron, I., Hróbjartsson, A., Altman, D.G., Moher, D., 2013. The evolution of assessing bias in Cochrane systematic reviews of interventions: celebrating methodological contributions of the Cochrane Collaboration. *Syst. Rev.* 2 (1) 79.
- West, S., King, V., Carey, T.S., Lohr, K.N., McKoy, M.D.N., Sutton, S.F., Lux, L., 2002. *Systems to Rate the Strength of Scientific Evidence*. In: *Evidence Report/Technology Assessment No. 47* (Prepared by the Research Triangle Institute-University of North Carolina Evidence-based Practice Center under Contract No. 290-97-0011). AHRQ Publication No. 02-E016. Agency for Healthcare Research and Quality, Rockville, MD.
- Wieland, S., Dickersin, K., 2005. Selective exposure reporting and Medline indexing limited the search sensitivity for observational studies of the adverse effects of oral contraceptives. *J. Clin. Epidemiol.* 58, 560–567.
- Wood, L., Egger, M., Gluud, L.L., Schulz, K.F., Juni, P., Altman, D.G., Gluud, C., Martin, R.M., Wood, A.J.G., Sterne, J.A.C., 2008. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *Br. Med. J.* 336 (7644) 601–605.

Gary S. Bilotta Dr Bilotta is currently a Senior Lecturer in Physical Geography at the University of Brighton. His research links the disciplines of hydrology, geomorphology, biogeochemistry and freshwater ecology. He has particular expertise in developing water quality monitoring and assessment tools for international water resource legislation. At the time of writing this article Dr Bilotta was seconded to the UK Department for Environment, Food and Rural Affairs (Defra), as part of the Natural Environment Research Council (NERC) Knowledge Exchange Scheme. The aim

of this secondment was to drive improvements in the policy evidence-base and its use.

Alice M. Milner Dr Milner is currently a Research Associate at University College London specialising in climatic and environmental change on a range of time scales, past, present and future. Her research is focussed on understanding how ecosystem respond to abrupt climate changes during previous warm intervals, using terrestrial and marine sediments as natural archives of environmental change. At the time of writing, Dr Milner was seconded to the UK Defra as part of the NERC Knowledge Exchange Scheme to improve the use and management of evidence in policy.

Ian Boyd Professor Boyd is currently a Professor in Biology at the University of St Andrews and Chief Scientific Adviser to the UK Defra. His career has evolved from Physiological ecologist with the NERC's Institute of Terrestrial Ecology, to a Science Programme Director with the British Antarctic Survey, Director at the NERC's Sea Mammal Research Unit, Chief Scientist to the Behavioural Response Study for the US-Navy, Director for the Scottish Oceans Institute and acting Director and Chairman with the Marine Alliance for Science and Technology for Scotland. In parallel to his formal positions he has chaired, co-chaired or directed international scientific assessments; his activities focusing upon the management of human impacts on the marine environment.