Environment International xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Environment International

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

Full Length Article

Systematic reviews: Separating fact from fiction

Neal R. Haddaway [a,*], Gary S. Bilotta [b]

[a] MISTRA EviEM, Royal Swedish Academy of Sciences, Stockholm, Sweden
[b] Aquatic Research Centre, University of Brighton, Brighton BN2 4GJ, United Kingdom

ARTICLE INFO

Article history:
Received 9 April 2015
Received in revised form 30 June 2015
Accepted 6 July 2015
Available online xxxx

Keywords:
Evidence-based practice
Evidence reviews
Evidence-based medicine
Evidence synthesis
Systematic review

ABSTRACT

The volume of scientific literature continues to expand and decision-makers are faced with increasingly unmanageable volumes of evidence to assess. Systematic reviews (SRs) are powerful tools that aim to provide comprehensive, transparent, reproducible and updateable summaries of evidence. SR methods were developed, and have been employed, in healthcare for more than two decades, and they are now widely used across a broad range of topics, including environmental management and social interventions in crime and justice, education, international development, and social welfare. Despite these successes and the increasing acceptance of SR methods as a 'gold standard' in evidence-informed policy and practice, misconceptions still remain regarding their applicability. The aim of this article is to separate fact from fiction, addressing twelve common misconceptions that can influence the decision as to whether a SR is the most appropriate method for evidence synthesis for a given topic. Through examples, we illustrate the flexibility of SR methods and demonstrate their suitability for addressing issues on environmental health and chemical risk assessment.

1. Background

Keeping up with information has never been easy, even before the 'Age of Information' (Bastian et al., 2010). For example, in 1753, when James Lind published his landmark review of what was then known about scurvy, he needed to point out that "…before the subject could be set in a clear and proper light, it was necessary to remove a great deal of rubbish" (Lind, 1753). The scientific evidence-base on many topics continues to grow, with a doubling of the number of cited references every 9 years over recent decades (Van Noorden, 2014). Systematic reviews (SRs) can be a powerful method for locating, appraising, and summarising evidence on a given topic. The methodology was originally developed for use in medicine, and its refinement in this field has been largely led by the Cochrane Collaboration (http://www.cochrane.org/), which was founded in 1993 after a renowned Scottish doctor, Archibald Cochrane (1979), reproached the medical profession for not having managed to organise a "critical summary, by speciality or subspeciality, adapted periodically, of all relevant randomised controlled trials" (Chalmers et al., 1992).

The Cochrane Collaboration is now 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. These experts aim to help healthcare practitioners, policy-makers, patients, their advocates and carers, make better-informed decisions about healthcare, by preparing, updating, and promoting the accessibility of SRs on the effectiveness of healthcare interventions. The Cochrane Collaboration have published more than 5000 SRs, 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). The SR practices of the Cochrane Collaboration have incited the development of other international initiatives including; the Campbell Collaboration (http://www.campbellcollaboration.org/), which was established in 2000 to prepare, maintain, and disseminate SRs on the effectiveness of social interventions in Crime & Justice, Education, International Development, and Social Welfare (Davies and Boruch, 2001); and the Collaboration for Environmental Evidence (http://www.environmentalevidence.org/), which was established in 2008 as an open community of scientists and managers who, from their initial centres in Australia, South Africa, Sweden, Canada, and the UK, prepare SRs on environmental topics (Pullin and Knight, 2013).

Across all disciplines, there are reportedly more than 4000 SRs being produced every year, and data show that the rate of production is increasing (Bastian et al., 2010). Nevertheless, SRs are still relatively new and unfamiliar to some disciplines, including environmental health and chemical risk assessment, for which there have only been a handful of SRs attempted so far (e.g. Adams et al., 2014; Alderman et al., 2012; Johnson et al., 2014; Liu et al., 2015; Schinasi and Leon, 2014; Shah and Balkhair, 2011). It is hoped that this Special Issue of *Environment International* will increase awareness of the potential value of SRs in this field. The aim of this article in particular, is to separate the facts

* Corresponding author.
E-mail address: neal_haddaway@hotmail.com (N.R. Haddaway).

http://dx.doi.org/10.1016/j.envint.2015.07.011
0160-4120/

from the fiction, addressing twelve common misconceptions that can influence the decision as to whether a SR is the most appropriate method for evidence synthesis for a given topic. The misconceptions covered in this article were identified at a workshop on SRs for Chemical Risk Assessment (Whaley et al., 2015), but they are also evident in the literature (where specified), and their prevalence has been confirmed by an online survey of SR experts (distributed through Twitter).

1.1. Misconception 1: a review is systematic if articles are identified through a systematic search, or a stepwise approach to inclusion

There is a widely held misconception that a literature review becomes a SR if the search and inclusion of articles is performed in a systematic way. This is a fallacy, which risks degrading the reputation of SRs as the 'gold standard' of evidence-informed policy and practice. This risk was actualized by a recent World Bank article by Evans and Popova (2015) which claimed to have conducted a SR of SRs on the effectiveness of methods to improve learning outcomes for children in low and middle income countries. Evans and Popova (2015) claimed to have identified six SRs, and in comparing the reviews' discovered divergent conclusions. However, as highlighted in a response to this article by Langer et al. (2015), neither Evans and Popova's (2015) own review, nor the majority of the reviews which they evaluated can be considered as SRs. A SR normally involves a number of purposeful formalised stages (formulating the question[s]; developing and publishing a protocol; conducting the searches; selecting the eligible studies; appraising the selected studies; extracting data for analysis and interpretation; disseminating and updating the review) (Bilotta et al., 2014a). Whilst the exact format of SRs may differ between the SR-coordinating bodies (including the Cochrane Collaboration, Campbell Collaboration, Collaboration for Environmental Evidence, European Food Safety Authority, the EPPI-Centre, and the Centre for Reviews and Dissemination.), three broad minimum standards are common to all of these organisations' SRs: (i) SR methods should be described in sufficient detail to allow full repeatability and traceability; (ii) they must include a systematic approach to identifying and screening relevant academic and grey literature, and (iii) they should include critical appraisal of the validity (quality and generalisability) of included studies to give greater weight to more reliable studies (Langer et al., 2015). Various resources exist that help readers to critique the quality of SRs (e.g. Scott et al. 2006).

1.2. Misconception 2: systematic reviews can only be used to answer questions that relate to the efficacy of interventions

Many of the early SRs in the healthcare field were initially limited to investigations of the efficacy of clinical interventions (e.g. Stjernswärd, 1974; Chalmers, 1975; Cochran et al., 1977; Smith and Glass, 1977), and whilst these types of questions lend themselves readily to SRs, they are not the only questions that can be, or have been, answered by SRs. This applies equally to medicine as it does to other disciplines (Petticrew, 2001), and it is a particularly salient point for consideration of the appropriateness of SRs to address questions from environmental health and chemical risk assessment.

Increasingly more common are SRs of the impacts of exposure to incidental factors, or indirect effects. An example of this sort of review includes SRs on the effect of maternal exposure to perfluorooctanoic acid – a chemical used in consumer products to impart fire resistance and oil, stain, grease, and water repellence – on human foetal growth (Koustas et al., 2014; Johnson et al., 2014; Lam et al., 2014). Another example of this sort of review includes the effect of occupational exposure to agricultural pesticide chemical groups on the incidence of non-Hodgkin's lymphoma (Schinasi and Leon, 2014). A further example, this time from the discipline of environmental science, includes the effect of climate change on Himalayan glacier mass (Miller et al., 2013). Other forms of SR can investigate the efficacy of different measurement methods, such as methods for measuring carbon in terrestrial carbon pools (Petrokofsky et al., 2012).

Systematic reviews can assess the effects or efficacy of any factor, not just the effectiveness of interventions.

1.3. Misconception 3: systematic reviews can only be used to answer narrow questions

Some have claimed that SRs focus on narrow questions that have limited practical utility, and that SRs are capable of investigating only single populations, interventions and outcomes (e.g. Doerr et al., 2014). Whilst many of the early healthcare SRs did have a relatively narrow, well-defined scope, the range of populations, interventions and outcomes now included in SRs, in healthcare and other fields, has expanded considerably. For example, a SR commissioned by the UK Department for Health on the effects of population-wide drinking water fluoridation strategies (McDonagh et al., 2000), considered multiple positive (e.g. reduction in incidence of tooth decay and cavities) and negative (e.g. dental fluorosis, cancer, bone fracture and bone development problems) outcomes. This SR also considered if any beneficial effects from water fluoridation were over and above that offered by the use of numerous alternative interventions and strategies (multiple interventions). It also examined how any beneficial effects from water fluoridation varied across different social groups and between geographical locations (multiple populations). Another example of a SR that considered multiple interventions and outcomes is provided by a recent Collaboration for Environmental Evidence SR of the human wellbeing impacts of a variety of terrestrial protected areas (Pullin et al., 2013). This SR iteratively included all measures of wellbeing identified in the evidence base. Similarly, a recent Campbell Collaboration SR examined multiple interventions (behavioural, psychological, educational and vocational) to facilitate multiple employment outcomes for cancer survivors (Fong et al., 2015).

Advances in SR methodology have seen the development of systematic maps (SMs) as a means of collating and cataloguing larger volumes of evidence following SR methodology as far as meta-data (information on study methods and context) extraction without fully synthesising the findings of included studies. SMs aim to produce a readily interrogable database of relevant studies on a subject and synthesis extends only to describing the evidence base rather than any findings of the included studies (CEE, 2013). SMs are highly valued by commissioners that wish to know how much evidence exists on a topic, and what form that evidence takes. SMs identify knowledge gluts (bodies of evidence that are sufficient in volume to permit full synthesis in SR) and knowledge gaps (areas of research that are conspicuous in their absence and warrant further primary research). SM methods have been used by a variety of different evidence synthesis coordinating body reviews, for example by: the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) to describe evidence on the relationship between obesity and sedentary behaviour in young people (Kalra and Newman, 2009); the Campbell Collaboration to describe evidence on the extent and impact of parental mental health problems on families and the acceptability, accessibility and effectiveness of interventions (Bates and Coren, 2006); the Collaboration for Environmental Evidence to describe evidence on the relationships between biodiversity and poverty (Roe et al., 2014).

Systematic reviews and SMs will always require focused, well-defined questions to ensure that projects remain manageable, that only evidence relevant to the review topic is included, and that the review conclusions are also focused and applicable in practice. This is not a disadvantage of the method, but rather a strength.

1.4. Misconception 4: systematic reviews can only include quantitative data from randomised controlled trials

There is a misconception that SRs are restrictive in the types of data that can be included; some believe SRs to be only capable of using

Please cite this article as: Haddaway, N.R., Bilotta, G.S., Systematic reviews: Separating fact from fiction, Environ Int (2015), http://dx.doi.org/10.1016/j.envint.2015.07.011

quantitative data from high-quality randomised controlled trials[1] (RCTs). This is a fallacy. Whilst early SRs in healthcare may have focussed on evidence from RCTs, there are now hundreds of healthcare SRs that do not use any evidence from RCTs at all. Some of these non-RCT containing SRs have been among the most influential of all the Cochrane Collaboration's reviews. For example, a Cochrane SR by Weller and Davis-Beaty (2002) investigated evidence on the effectiveness of condoms in reducing heterosexual HIV transmission. This SR used cohort studies of sexually-active couples where one of the partners was HIV-positive and one of them was HIV-negative at the start of the study. The sexually-active couples were grouped into those that always (100% of the time) or never (0% of the time) used condoms during sex. The follow-up of the initially HIV-negative partners who all had a known exposure to the disease, provided data on the effectiveness of condoms at reducing disease transmission and incidence.

In the above example, it was clearly not ethically-possible or desirable to set-up RCTs to study the effectiveness of the intervention of interest. In some topics it is not physically-possible to conduct RCTs of the intervention of interest, and thus any SR of the evidence will not contain RCT studies. A prime example of this is SRs on global climate change topics. This concept cannot be applied to world climate because there is only one Earth and we cannot randomly allocate replicates to different control and elevated carbon dioxide concentrations, thus in this case reviewers may have to make use of observational evidence (e.g. Miller et al., 2013). The majority of SRs by the Collaboration for Environmental Evidence use evidence from observational studies[2] and do not use any evidence from RCTs. This is not always because it may not be physically-possible to conduct RCTs on environmental topics, but simply because it is not a widely-used research method in this discipline.

There is a drive for more widespread use of RCTs across many disciplines, from education to international development, when the purpose of the studies is to assess the effectiveness of interventions (Haynes et al., 2012). Sometimes, the purpose of primary studies and SRs is to assess something other than the effectiveness of an intervention, or to use quantitative data as a means of doing so. Qualitative data are study findings that are not expressed in terms of numbers but rather textual descriptions, such as responses to open-ended interview questions. Qualitative syntheses are techniques to combine findings in a textual way, for example by identifying themes across studies. Primary research in environmental sciences most often takes the form of quantitative data, and as such, the majority of SRs in environmental sciences and public health have used quantitative syntheses. However, qualitative SR methods are common in subjects that deal with qualitative data, such as social research. The Cochrane Collaboration published its first SRs of qualitative evidence in November 2013 (Glenton et al., 2013; Gülmezoglu et al., 2013), but SRs by the Campbell Collaboration and Collaboration for Environmental Evidence have included qualitative data for some time now, and the continued expansion of SR use in social sciences attests to their adaptability to qualitative data (Petticrew and Roberts, 2008).

1.5. Misconception 5: systematic reviews must include a meta-analysis

There is a wide-held belief that SRs that focus on quantitative data MUST include a quantitative synthesis, i.e. meta-analysis, or that narrative synthesis alone is insufficient (Doerr et al., 2014). This is another fallacy. Syntheses within SRs may be narrative, such as a structured summary and discussion of the studies' characteristics and findings, qualitative, involving a structure, non-numerical approach, such as a thematic analysis to produce a conceptual framework, or quantitative, involving statistical analysis such as meta-analysis or meta-regression (Deeks et al., 2011). Meta-analysis – the statistical combination of results from two or more separate studies – is the most commonly used statistical technique in evidence syntheses, although it is stressed that meta-analysis is not appropriate in all SRs (Deeks et al., 2011).

Meta-analysis and related techniques can be used if there is a consistent outcome measure to estimate the size of the effect and the uncertainty surrounding that effect size, and to investigate whether the effect is consistent across studies. This is often relatively easy for SRs that assemble studies making one particular comparison between two treatment options, and is further facilitated where consistent units are reported. But as mentioned above, some SRs include a broader range of study types, data types, and/or examine multiple interventions and outcomes; for example, the SR mentioned above, on the impacts of terrestrial protected areas on human wellbeing, performed no meta-analysis because of incomparable data (Pullin et al., 2013). In these reviews, meta-analysis may be unfeasible or inappropriate - like comparing apples and oranges - so a structured narrative synthesis may be more suitable.

1.6. Misconception 6: systematic reviews cannot capture socio-political-, economic-, or health- context

There is a misconception that SRs cannot capture socio-political-, economic-, or health- context, and that SR is not suitable for interdisciplinary topics (Dafforn et al., 2015; and response by Haddaway and Bayliss, 2015). As seen with other misconceptions detailed above, early SRs did tend to focus on single populations, interventions and outcomes without necessarily considering the wider socio-political-, economic-, or health- implications, but a brief survey of the libraries of the Campbell Collaboration, the Collaboration for Environmental Evidence, and the Cochrane Collaboration illustrates that this is definitely no longer the case. A plethora of multidisciplinary SRs exist that do consider the wider context. Campbell Collaboration SRs are particularly good at this since their remit is often a diverse range of healthcare, social science, environmental sciences, political science and international development. There is certainly a challenge in combining different disciplines in single SRs, since the disciplines involved often have divergent approaches to data collection and synthesis. Synthesis experts are committed to raising awareness of the suitability of SR for multidisciplinary SRs and their emphasis on placing findings in context. In fact, the Campbell and Cochrane Economics Methods Group (CCEMG), which was formally registered as a Cochrane Collaboration methods group in 1998 and has been jointly registered as a Campbell Collaboration Methods Group since 2004, have produced detailed guidance on how to incorporate context information such as economic evidence in SRs (Shemilt et al., 2008). Increasingly, the Collaborations will work together on SRs to ensure that the SRs incorporate socio-political-, economic-, and/or health- data, with the first multi-registered SRs recently published between the Collaboration for Environmental Evidence and the Campbell Collaboration: '*Effects of Decentralised Forest Management (DFM) on Deforestation and Poverty in Low and Middle Income Countries*' (Samii et al., 2014a) and '*Effects of Payment for Environmental Services (PES) on Deforestation and Poverty in Low and Middle Income Countries*' (Samii et al., 2014b).

1.7. Misconception 7: systematic reviews take long periods of time to complete and require considerable funding

There is a perception that all SRs take long periods of time to complete (~2 years) require considerable funding (hundreds of thousands

[1] RCTs are studies in which study units are randomly allocated to intervention or control groups and followed up over time to assess any differences in outcome rates. Randomisation with allocation concealment (double blinded) ensures that on average known and unknown determinants of outcome (including confounding factors) are evenly distributed between groups.

[2] Observational studies are studies in which natural variation in interventions (or exposure) among study units is investigated to explore the effect of the interventions (or exposure) on outcomes.

Please cite this article as: Haddaway, N.R., Bilotta, G.S., Systematic reviews: Separating fact from fiction, Environ Int (2015), http://dx.doi.org/10.1016/j.envint.2015.07.011

of US$). In reality, the time required to complete a SR, from submission of the protocol to submission of the full review, varies from 8 months to 24 months for SRs published by the Collaboration for Environmental Evidence (Andrew Pullin, Pers. Comm.). A recent survey of authors of evidence reviews suggests that different disciplines vary in their SR time requirements (Table 1). The CEE believes that if an up-to-date SM already exists, conversion to a full SR can be undertaken in roughly 4 months. The time periods often quoted for SRs which have led to this misconception, are at the upper end of the scale and are inclusive of all processes related to the review, including administrative and publication bottlenecks and times of inactivity (e.g. whilst waiting to hear from authors of the primary research for raw data or clarification of the methods used in some of the primary studies cited).

Three of the most influential factors affecting the time required to undertake a SR are (1) the complexity of the review question and the type and volume of evidence available, (2) the experience level of the lead reviewer(s) and (3) availability of the lead reviewer(s) and the review team. An experienced team with strong time commitments to the review stand a good chance of completing a SR in under a year (excluding publication time for the protocol and full review report), depending on the topic and volume of the evidence. There have been attempts to develop 'rapid review' methods such as Rapid Evidence Assessments and Quick Scoping Reviews (e.g. Collins et al., 2014). Often, however, the timescales quoted for these rapid methods overlap those of SRs. Some SRs have been undertaken in short timescales even in comparison to more rapid evidence review methods, such as Quick Scoping Reviews (Haddaway unpublished data).

The cost of completing SRs depends on the size of the review team and the quantity of evidence available to review, but normally costs vary between US $30,000 and US $300,000 for SRs published by the Collaboration for Environmental Evidence (Andrew Pullin, Pers. Comm.). SR coordinating bodies vary in the publication costs for endorsed reviews: for example, CEE's journal *Environmental Evidence* charges article processing fees of US $2255 for the protocol and US $1127.50 for the subsequent full review (authors from more than 90 low-income countries receive automatic waivers on submission and do not have to pay article-processing charges). Ultimately, the decision to fund a SR should take into account whether the resource requirements (both time and money) outweigh the costs and risks of using a less reliable form of evidence synthesis that may very quickly become redundant when new studies and data become available. Where the human cost of failing to produce robust syntheses is high, then the value of SRs is clear; illustrated by the life-saving potential of a SR conducted to examine infant sleeping position and sudden infant death syndrome. Although advice to place infants on their backs to sleep was widely available in the early 1990s, the authors of the sudden infant death syndrome SR showed that the mortality benefit of this sleep position would have been apparent if a SR had been performed any time after 1970 (Gilbert et al., 2005). Such a review potentially could have saved 60,000 infant lives in the United Kingdom, Europe, the United States, and Australia (Freeman et al., 2006).

1.8. Misconception 8: systematic reviews fail to provide decision-makers with easily digestible summaries of the evidence and its synthesis

There is a misconception that SRs are not designed to readily support decision-making (Doerr et al., 2014). SRs are often written with an expert audience in mind, particularly SRs in medicine. However, all medics are trained in evidence-based medicine and are provided with a broad understanding of SR methodology: this is not an issue for these SRs in medicine. However for SRs in other disciplines, no such universal training exists. It is indeed true that SRs are often difficult to interpret and are typically long documents featuring extensive methodological detail that may make them hard for the uninitiated to digest. This is not a fault of SRs, however, and cannot be a criticism of the methodology. The detail is included to explain transparently the methods by which the reviewers searched for and evaluated the evidence, such that any informed reader can examine the review in detail, and understand from the background, the reasons behind every decision (Goldacre, 2014). Doerr et al. (2014) are right to recommend "that reviewers should produce a simple and brief summary of recommendations that follow from the review's synthesis of available evidence", but this is not at odds with SR guidance. In fact, SR coordinating bodies, such as the Collaboration for Environmental Evidence (CEE), advocate the use of policy briefs (Collaboration for Environmental Evidence, 2013), fact sheets, research briefs and short summaries tailored to the needs of researchers, decision-makers and non-technical stakeholders alike: see for example, the website of MISTRA EviEM, a CEE Centre in Stockholm (http://www.eviem.se) and the summary media relating to a SR on biomanipulation (Bernes et al., 2015).

1.9. Misconception 9: systematic reviews take money away from primary research/SRs are a substitute for primary research

SR methods are sometimes advocated as a more cost-efficient means of arriving at an answer to an evidence need than commissioning primary research. Indeed, cases exist in medicine where primary research is shown to be no longer necessary, since combining studies in one synthesis shows significance where individual studies fail to find any. For example, a SR which conducted cumulative meta-analytical techniques on 64 trials investigating the effectiveness of the drug Aprotinin at controlling perioperative bleeding showed that the effectiveness was apparent after only 12 trials (Fergusson et al., 2005). Thus this SR identified 52 unnecessary trials that had a SR been performed after the twelfth study, the treatment effect would have been apparent, duplicate trials would have been avoided, and patients would have experienced the benefit of a useful drug ten years earlier (Freeman et al., 2006). However, SRs will not make primary research obsolete – SRs themselves rely on primary research – but they may encourage policy-makers and funding bodies to avoid funding further duplicative studies in areas where there is already clear evidence. This is a positive step, saving valuable resources and directing them to areas where further research will have more benefit. Indeed, several SRs have been funded as part of primary research projects to better assess current understanding and ensure that the primary research methodologies to be used are based on

Table 1
Reported time requirements for systematic reviews across four major systematic review coordinating bodies (categories are total time requirements in months followed by reported person-months in brackets). Results of unpublished survey by Haddaway 2014.

	Review coordinating body			
	Campbell SR	CEE SR	Cochrane SR	EPPI-Centre SR
	www.campbellcollaboration.org	www.environmentalevidence.org	www.cochrane.org	http://eppi.ioe.ac.uk
Median time category (months)	23–24 (3–4/11–12)	17–18 (15–16)	23–24 (5–6)	17–18 (5–6)
Minimum time (months)	11–12 (3–4)	7–8 (3–4)	3–4 (1–2)	5–6 (3–4)
Maximum time (months)	51–52 (51–52)	35–36 (35–36)	51–52 (37–38)	27–28 (51–52)[a]

[a] Discrepancy in responses between total time and person time due to inaccurate response from one respondent. Next most-likely value is 29–30 months.

Please cite this article as: Haddaway, N.R., Bilotta, G.S., Systematic reviews: Separating fact from fiction, Environ Int (2015), http://dx.doi.org/10.1016/j.envint.2015.07.011

a critical assessment of the existing literature (e.g. Haddaway et al., 2014; Evans et al., 2014). Such combination of primary and secondary research is mutually complimentary, not competitive.

1.10. Misconception 10: systematic reviews must be registered with a 'collaboration'

There is a misconception that all SRs must be pre-registered with a collaboration such as the Cochrane Collaboration, Campbell Collaboration, or the Collaboration for Environmental Evidence (see Table 1). This again is a fallacy - anyone can conduct a SR and there is no requirement to register the review through one of these collaborations. There are a number of other organisations conducting SRs around the globe that are producing reproducible SRs and there are also specific publication platforms available to facilitate this, such as the *Systematic Reviews* journal. However, registration and publication of a SR through a coordinating body carries with it a number of key benefits in excess of those related to good SR practice, which must be considered carefully before choosing to publish a SR elsewhere. In summary, these are: peer-review of the protocol, advice during review activities, peer-review of the final full review, and prevention of duplication of efforts.

Publication of a SR protocol of some sort (not necessarily with a collaboration) on a publically-available platform is good practice. The purpose of protocol registration is to establish and document in advance, prior to knowledge of the available studies, the question(s) to be answered or hypothesis to be tested (relating to specific populations, interventions, comparisons and outcomes), the search terms and databases to be searched, and the quality assessment and data analysis methods to be used. This aims to reduce 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 (Bilotta et al., 2014a). Publication of the protocol via a SR coordinating body allows for independent peer-review by subject and methodology experts to ensure that mistakes in methodology are corrected from the outset: this is particularly true for search strategies, which are often corrected during protocol peer-review. Such peer-review benefits are also useful for publication of the full review: often mistakes and lack of detail in reporting the activities of the review can be readily corrected to maximise the quality and usability of the review that might otherwise be missed by the non-methodology experts that are likely to peer-review manuscripts in conventional journals. Finally, registration of a SR topic with a coordinating body helps to prevent duplication of effort as all new review topics are screened against ongoing reviews before being given the go-ahead.

There are other benefits in registering a review with a collaboration, including access to expertise, resources, and use of an updateable open-access publishing platform. Furthermore, these collaborations are working together to drive methodological improvements and innovations to raise standards and consistency between reviews (Bilotta et al., 2014b).

1.11. Misconception 11: systematic reviews provide the definitive answer to questions

There is a misconception that SRs provide definitive answers to questions. This is not always true. Whilst the rigour of SR methodology does mean that the findings of a review are less likely to be influenced by potential biases of the review authors, it doesn't guarantee conclusive answers. Ultimately the strength of the conclusions of a SR is dependent on the availability and quality of primary studies addressing the question of interest. There are examples, from all disciplines, of SRs that have been inconclusive owing to the limited availability of good quality primary studies. This is not a problem to do with the SR process itself. In fact, the SR process can be extremely useful for identifying research gaps and methods to improve the evidence-base on a given topic. Moreover, the reproducible and updateable nature of SRs means that if early inconclusive reviews are used to inform research design and increase the availability of good quality evidence, subsequent primary studies can later be incorporated into the updated version of the same SR, so that it can provide a more definitive answer. A fascinating example of this update process in practice has been provided by the Cochrane Collaboration's review of the effectiveness of neuraminidase inhibitors, such as oseltamivir (*Tamiflu*), for the prevention and treatment of influenza (Jefferson et al., 2014). Governments around the world had spent billions of pounds stockpiling these drugs in case of a flu pandemic. In this case the original SR was updated to include data that had previously been withheld by Roche (the drug company behind *Tamiflu*). The updated SR found that this drug, didn't work as well as the first review had suggested (Goldacre, 2014; Jefferson et al., 2014); concluding that "*the balance between benefits and harms should be considered when making decisions about use of both neuraminidase inhibitors for either the prophylaxis or treatment of influenza. The influenza virus-specific mechanism of action proposed by the producers does not fit the clinical evidence*".

1.12. Misconception 12: SRs are undertaken by outside researchers detached from the realities of practitioners

SRs are sometimes criticised as being academic works that are performed by researchers with little understanding of the practicalities faced by practitioners, and there is a misconception that SRs can or should be performed without subject experts (Doerr et al., 2014). This is simply not the case. SR review teams should always involve relevant subject experts: these are often leaders in the field not only in terms of research, but also practice and policy. Many SRs, including those undertaken according to guidelines set out by the EPPI-Centre (http://eppi.ioe.ac.uk), involve a review team that undertakes the review activities and a steering group that advises on how the review should be undertaken. Steering groups provide another opportunity for input from policy and practice experts as well as methodology and research experts. Irrespective of the make-up of the review team or steering group, however, a SR should always involve stakeholder engagement in order to ensure that the review and its outputs are as relevant as possible to the practical subject tackled by the review. Indeed, many SRs are accompanied by dissemination media tailored to practitioners, such as factsheets and infographics.

2. Summary

The use of SRs in numerous disciplines has increased in recent years, and SR methods are more widely understood and accepted. However, misconceptions still remain and proliferate today. Whilst SR advocates welcome constructive criticism and recognise its value in challenging dogma and developing novel approaches, many of these criticisms are in fact misconceptions. We have aimed to dispel the 12 misconceptions and myths detailed herein and hope that these points will help to clarify and promote the utility of thoughtful SR, whilst recognising that challenges facing evidence-based practice, for example non-reporting of intervention efficacy, publication bias, and poor reporting of study data remain (Bilotta et al., 2015; Haddaway, 2015; Pullin and Knight, 2005). These challenges, however, are not faults with SRs, but rather faults that SR methods aim to highlight and assess. Solutions to these challenges are complex, but worthy of additional effort. Future developments in SR methods, such as those investigated and proposed by specialist groups like CEE Methods Groups (www.environmentalevidence.org/method-groups), may help to mitigate these problems, but a concerted effort is required to solve the underlying problems rather that treating the symptoms. Systematic review methods are as useful and reliable for synthesising evidence for better decision-making in chemical risk assessment as they are in healthcare, environmental management, and social, criminal and educational fields.

Please cite this article as: Haddaway, N.R., Bilotta, G.S., Systematic reviews: Separating fact from fiction, Environ Int (2015), http://dx.doi.org/10.1016/j.envint.2015.07.011

Acknowledgements

The authors wish to thank Helen Bayliss and Andrew Pullin for discussions of the topic and two anonymous reviewers for comments on an earlier draft of the manuscript.

References

Adams, J.A., Galloway, T.S., Mondal, D., Esteves, S.C., Mathews, F., 2014. Effect of mobile telephones on sperm quality: a systematic review and meta-analysis. Environ. Int. 70, 106–112.

Alderman, K., Turner, L.R., Tong, S., 2012. Floods and human health: a systematic review. Environ. Int. 47, 37–47.

Bastian, H., Glasziou, P., Chalmers, I., 2010. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? PLoS Med. 7.e1000326 http://dx.doi.org/10.1371/journal.pmed.1000326.

Bates, S., Coren, E., 2006. SCIE Systematic map report 1: The extent and impact of parental mental health problems on families and the acceptability, accessibility and effectiveness of interventions Accessed 10th February 2015 http://www.scie.org.uk/publications/map/map01.asp.

Bernes, Claes, Carpenter, Stephen R., Gårdmark, Anna, Larsson, Per, Persson, Lennart, Skov, Christian, Speed, James DM, Van Donk, Ellen, 2015. What is the influence of a reduction of planktivorous and benthivorous fish on water quality in temperate eutrophic lakes? A systematic review. Environmental Evidence 4 (1), 7.

Bilotta, G.S., Milner, A.M., Boyd, I., 2014a. On the use of systematic reviews to inform environmental policies. Environ. Sci. Pol. 42, 67–77.

Bilotta, G.S., Milner, A.M., Boyd, I., 2014b. Quality assessment tools for evidence from environmental science. Environ. Evid. 3, 14. http://dx.doi.org/10.1186/2047-2382-3-14.

Bilotta, G.S., Boyd, I.L., Milner, A.M., 2015. How to increase the potential policy impact of environmental science research. Environ. Sci. Eur. 27, 9. http://dx.doi.org/10.1186/s12302-015-0041-x.

Collaboration for Environmental Evidence, 2013. Guidelines for Systematic Review and Evidence Synthesis in Environmental Management. Version 4.2. Collaboration for Environmental Evidence: www.environmentalevidence.org/Documents/Guidelines/Guidelines4.2.pdf.

Collins, A., Miller, J., Coughlin, D., Kirk, S., 2014. The Production of Quick Scoping Reviews and Rapid Evidence Assessments: A How to Guide. Beta Version 2. UK Defra, Joint Water Evidence Group.

Chalmers, T.C., 1975. Effects of ascorbic acid on the common cold. An evaluation of the evidence. Am. J. Med. 58, 532–536.

Chalmers, I., Dickersin, K., Chalmers, T.C., 1992. Getting to grips with Archie Cochrane's agenda. BMJ 305, 786–788.

Cochran, W.G., Diaconis, P., Donner, A.P., Hoaglin, D.C., O'Connor, N.E., Peterson, O.L., Rosenoer, V.M., 1977. Experiments in surgical treatments of duodenal ulcer. In: Bunker, J.P., Barnes, B.A., Mosteller, F. (Eds.), Costs, Risks and Benefits of Surgery. Oxford University Press, Oxford, pp. 176–197.

Cochrane, A.L., 1979. 1931–1971: a critical review, with particular reference to the medical profession. In: Cochrane, A.L., Teeling-Smith, G., Wells, N.E. (Eds.), Medicines for the Year 2000. Office of Health Economics, London, pp. 1–11.

Dafforn, K.A., Glasby, T.M., Airoldi, L., Rivero, N.K., Mayer-Pinot, M., Johnston, E.L., 2015. Marine urbanization: an ecological framework for designing multifunctional artificial structures. Front. Ecol. Environ. 13, 82–90. http://dx.doi.org/10.1890/140050.

Davies, P., Boruch, R., 2001. The Campbell Collaboration: does for public policy what Cochrane does for health. Br. Med. J. 323, 294–295.

Deeks, J.J., Higgins, J.P.T., Altman, D.G., 2011. Analysing Data and Undertaking Meta-Analyses. In: Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (Updated March 2011). The Cochrane Collaboration, London Chapter 9.

Doerr, E.D., Dorrough, J., Davies, M.J., Doerr, V.A.J., McIntyre, S., 2014. Maximizing the value of systematic reviews in ecology when data or resources are limited. Austral Ecol. 40, 1–11.

Evans, C., Burden, A., Morrison, R., Baird, A., Callaghan, N., Chapman, P., Cumming, A., Dixon, S., Evans, J., Gauci, V., Haddaway, N.R., Heppell, K., Healey, J., Holden, J., Hughes, S., Kaduk, J., Jones, D., Matthews, R., Menichino, N., Misselbrook, T., Page, S., Peacock, M., Pullin, A., Rayment, M., Ridley, L., Robinson, I., Stanley, K., Ward, H., Archer, E., Worrall, F., 2014. Interim Report on Project SP1210: lowland peatland systems in England and Wales — evaluating greenhouse gas fluxes and carbon balances. First Annual Data Report, March.

Evans, D., Popova, E., 2015. What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews. World Bank Policy Research Working Paper 7203.

Fergusson, D., Glass, K., Hutton, B., Shapiro, S., 2005. Randomized controlled trials of aprotinin in cardiac surgery: could clinical equipoise have stopped the bleeding? Clin. Trials 2, 218–232.

Fong, C.J., Murphy, K.M., Westbrook, J.D., Markle, M.M., 2015. Behavioral, Psychological, Educational, and Vocational Interventions to Facilitate Employment Outcomes for Cancer Survivors: A Systematic Review. Campbell Systematic Reviews 11, 5. http://dx.doi.org/10.4073/csr.2015.5.

Freeman, S.R., Williams, H.C., Dellavalle, R.P., 2006. The increasing importance of systematic reviews in clinical dermatology research and publication. J. Investig. Dermatol. 126, 2357–2360.

Gilbert, R., Salanti, G., Harden, M., See, S., 2005. Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002. Int. J. Epidemiol. 34, 874–887.

Glenton, C., Colvin, C.J., Carlsen, B., Swartz, A., Lewin, S., Noyes, J., Rashidian, A., 2013. Barriers and facilitators to the implementation of lay health worker programmes to improve access to maternal and child health: qualitative evidence synthesis. Cochrane Database Syst. Rev. 10, CD010414. http://dx.doi.org/10.1002/14651858.CD010414.pub2.

Goldacre, B., 2014. What the Tamiflu saga tells us about drug trials and big pharma. The Guardian 10th April http://www.theguardian.com/business/2014/apr/10/tamiflu-saga-drug-trials-big-pharma.

Gülmezoglu, A.M., Chandler, J., Shepperd, S., Pantoja, T., 2013. Reviews of qualitative evidence: a new milestone for Cochrane. Cochrane Database Syst. Rev. ED000073 http://dx.doi.org/10.1002/14651858.ED000073.

Haddaway, N., 2015. A call for better reporting of conservation research data for use in meta-analyses. Conserv. Biol. http://dx.doi.org/10.1111/cobi.12449.

Haddaway, N.R., Bayliss, H.R., 2015. Clarification on the applicability of systematic reviews. Front. Ecol. Environ. 13, 129. http://dx.doi.org/10.1890/15.WB.003.

Haddaway, N.R., Burden, A., Evans, C.D., Healey, J.R., Jones, D.L., Dalrymple, S.E., Pullin, A.S., 2014. Evaluating effects of land management on greenhouse gas fluxes and carbon balances in boreo-temperate lowland peatland systems. Environmental Evidence 3 (1), 5.

Haynes, L., Goldacre, B., Torgerson, D., 2012. Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials. Cabinet Office: Behavioural Insights Team.

Jefferson, T., Jones, M.A., Doshi, P., Del Mar, C.B., Hama, R., Thompson, M.J., Spencer, E.A., Onakpoya, I.J., Mahtani, K.R., Nunan, D., Howick, J., Heneghan, C.J., 2014. Neuraminidase inhibitors for preventing and treating influenza in adults and children. Cochrane Database Syst. Rev. CD008965 http://dx.doi.org/10.1002/14651858.CD008965.pub4.

Johnson, P.I., Sutton, P., Atchley, D.S., Koustas, E., Lam, J., Sen, S., Robinson, K.A., Axelrad, D.A., Woodruff, T.J., 2014. The navigation guide—evidence-based medicine meets environmental health: systematic review of human evidence for PFOA effects on fetal growth. Environ. Health Perspect. 122, 1028–1039. http://dx.doi.org/10.1289/ehp.1307893.

Kalra, N., Newman, M., 2009. A Systematic map of the Research on the Relationship Between Obesity and Sedentary Behaviour in Young People. Technical report. EPPI-Centre, Institute of Education, London.

Koustas, E., Lam, J., Sutton, P., Johnson, P.I., Atchley, D.S., Sen, S., Robinson, K.A., Axelrad, D.A., Woodruff, T.J., 2014. The navigation guide—evidence-based medicine meets environmental health: systematic review of nonhuman evidence for PFOA effects on fetal growth. Environ. Health Perspect. 122, 1015–1027.

Lam, J., Koustas, E., Sutton, P., Johnson, P.I., Atchley, D.S., Sen, S., Robinson, K.A., Axelrad, D.A., Woodruff, T.J., 2014. The navigation guide—evidence-based medicine meets environmental health: integration of animal and human evidence for PFOA effects on fetal growth. Environ. Health Perspect. 122, 1040–1051.

Langer, L., Haddaway, N., Land, M., 2015. Losing sight in the jungle of reviews: a comment on the system behind 'systematic reviews'. Africa Evidence Network. http://www.africaevidencenetwork.org/loosing-sight-in-the-jungle-of-reviews-a-comment-on-the-system-behind-systematic-reviews/ (Accessed 27/06/2015).

Lind, J., 1753. A treatise of the scurvy in three parts. Containing an inquiry into the nature, causes and cure of that disease, together with a critical and chronological view of what has been published on the subject. Miller, London.

Liu, J.C., Pereira, G., Uhl, S.A., Bravo, M.A., Bell, M.L., 2015. A systematic review of the physical health impacts from non-occupational exposure to wildfire smoke. Environ. Res. 136, 120–132.

McDonagh, M.S., Whiting, P.F., Wilson, P.M., Sutton, A.J., Chestnutt, I., Cooper, J., Misso, K., Bradley, M., Treasure, E., Kleijnen, J., 2000. Systematic review of water fluoridation. BMJ 321, 855–859.

Miller, J., Rees, G., Warnaars, T., Young, G., Collins, D., Shrestha, A., 2013. What is the evidence for glacial shrinkage across the Himalayas? CEE Library CEE review10-008. www.environmentalevidence.org/SR10008.html.

Petrokofsky, G., Kanamaru, H., Achard, F., Goetz, S.J., Joosten, H., Holmgren, P., Lehtonen, A., MCS, M., ASP, P., Wattenbach, M., 2012. Comparison of methods for measuring and assessing carbon stocks and carbon stock changes in terrestrial carbon pools. How do the accuracy and precision of current methods compare? A systematic review protocol. Environ. Evid. 1, 6.

Petticrew, M., 2001. Systematic reviews from astronomy to zoology: myths and misconceptions. BMJ 322, 98–101.

Petticrew, M., Roberts, H., 2008. Systematic Reviews in the Social Sciences: A Practical Guide. John Wiley & Sons, Oxford.

Pullin, A.S., Knight, T.M., 2005. Assessing conservation management's evidence base: a survey of management-plan compilers in the United Kingdom and Australia. Conserv. Biol. 19, 1986–1989.

Pullin, A.S., Knight, T.M., 2013. Time to build capacity for evidence synthesis in environmental management. Environ. Evid. 2, 21–22.

Pullin, A.S., Bangpan, M., Dalrymple, S.E., Dickson, K., Haddaway, N.R., Healey, J.R., Hauari, H., Hockley, N., Jones, J.P.G., Knight, T., Vigurs, C., Oliver, S., 2013. Human well-being impacts of terrestrial protected areas. Environ. Evid. 2, 19.

Roe, D., Fancourt, M., Sandbrook, C., Sibanda, M., Giuliani, A., Gordon-Maclean, A., 2014. Which components or attributes of biodiversity influence which dimensions of poverty? Environ. Evid. 3, 3.

Samii, C., Lisiecki, M., Kulkarni, P., Paler, L., Chavis, L., 2014a. Effects of Decentralized Forest Management (DFM) on deforestation and poverty in low and middle income countries: a systematic review. Campbell Syst. Rev. 10. http://dx.doi.org/10.4073/csr.2014.10.

Samii, C., Lisiecki, M., Kulkarni, P., Paler, L., Chavis, L., 2014b. Effects of Payment for Environmental Services (PES) on deforestation and poverty in low and middle income countries: a systematic review. CEE Library CEE 13-015b.

Schinasi, L., Leon, M.E., 2014. Non-Hodgkin lymphoma and occupational exposure to agricultural pesticide chemical groups and active ingredients: a systematic review and meta-analysis. Int. J. Environ. Res. Public Health 11, 4449–4527.

Please cite this article as: Haddaway, N.R., Bilotta, G.S., Systematic reviews: Separating fact from fiction, Environ Int (2015), http://dx.doi.org/10.1016/j.envint.2015.07.011

Scott, I., Greenberg, P., Poole, P., Campbell, D., 2006. Cautionary tales in the interpretation of systematic reviews of therapy trials. Intern. Med. J. 36, 587–599.
Shah, P.S., Balkhair, T., 2011. Air pollution and birth outcomes: a systematic review. Environ. Int. 37, 498–516.
Shemilt, I., Mugford, M., Byford, S., Drummond, M., Eisenstein, E., Knapp, M., Mallender, J., McDaid, D., Vale, L., Walker, D., 2008. Chapter 15: incorporating economics evidence. In: JPT, H., Green, S. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008].
Smith, M.L., Glass, G.V., 1977. Meta-analysis of psychotherapy outcome studies. Am. Psychoanal. 32, 752–760.
Stjernswärd, J., 1974. Decreased survival related to irradiation postoperatively in early breast cancer. Lancet 304, 1285–1286.
Weller, S.C., Davis-Beaty, K., 2002. Condom effectiveness in reducing heterosexual HIV transmission. Cochrane Database Syst. Rev. 1, CD003255. http://dx.doi.org/10.1002/14651858.CD003255.

Please cite this article as: Haddaway, N.R., Bilotta, G.S., Systematic reviews: Separating fact from fiction, Environ Int (2015), http://dx.doi.org/10.1016/j.envint.2015.07.011