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