Implementing systematic review techniques in chemical risk assessment: Challenges, opportunities and recommendations

Paul Whaley, Crispin Halsall, Marlene Ågerstrand, Elisa Aiassa, Diane Benford, Gary Bilotta, David Coggon, Chris Collins, Ciara Dempsey, Raquel Duarte-Davison, Rex FitzGerald, Malyka Galay-Burgos, David Gee, Sebastian Hoffmann, Juleen Lam, Toby Lasserson, Len Levy, Steven Lipworth, Sarah Mackenzie Ross, Olwenn Martini, Catherine Meads, Monika Meyer-Baron, James Miller, Camilla Pease, Andrew Rooney, Alison Sapietsu, Gavin Stewart, David Taylor

Systematic review (SR) is a rigorous, protocol-driven approach designed to minimise error and bias when summarising the body of research evidence relevant to a specific scientific question. Taking as a comparator the use of SR in synthesising research in healthcare, we argue that SR methods could also pave the way for a “step change” in the transparency, objectivity and communication of chemical risk assessments (CRA) in Europe and elsewhere. We suggest that current controversies around the safety of certain chemicals are partly due to limitations in current CRA procedures which have contributed to ambiguity about the health risks posed by these substances. We present an overview of how SR methods can be applied to the assessment of risks from chemicals, and indicate how challenges in adapting SR methods from healthcare research to the CRA context might be overcome. Regarding the latter, we report the outcomes from a workshop exploring how to increase uptake of SR methods, attended by experts representing a wide range of fields related to chemical toxicology, risk analysis and SR. Priorities which were established include: the conduct of CRA-focused prototype SRs; the development of a recognised standard of reporting and conduct for SRs in toxicology and CRA; and establishing a network to facilitate research, communication and training in SR methods. We see this paper as a milestone in the creation of a research climate that fosters communication between experts in CRA and SR and facilitates wider uptake of SR methods into CRA.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
1. Introduction

Systematic review (SR) is a rigorous, protocol-driven approach to minimising error and bias in the aggregation and appraisal of evidence relevant to answering a research question. SR techniques were initially developed in the fields of psychology, social science and health care and have, since the 1980s, provided a valuable tool for evidence-informed decision-making across many domains (Lau et al., 2013). In medicine, SRs have provided a valuable response to the need for consistent, transparent and scientifically-robust interpretations of the results of increasing numbers of often conflicting studies of the efficacy of healthcare interventions. SRs have taken on an increasingly fundamental role both in supporting decision-making in healthcare and, by channelling resources towards questions for which the answers are not yet known, reducing waste in research (Chalmers and Glasziou, 2009; Salman et al., 2014). It is now accepted practice in healthcare to use SR methods to assess evidence not only for the efficacy of interventions, but also on diagnostic tests, prognostics and adverse outcomes.

The extension of SR techniques to other fields is based on a mutual need across disciplines to make the best use of existing evidence when making decisions, a move for which momentum has been growing for several decades. For example, the What Works Clearinghouse was established in 2002 to apply SR techniques in support of American educational policy (US Institute of Education Sciences, 2015), and in 2000 the international Campbell Collaboration research network was convened to undertake and disseminate systematic reviews on the effects of social interventions in diverse fields such as crime and justice, education, international development and social welfare (Campbell Collaboration, 2015). Meta-analysis and SR in ecology have contributed to evidence-based environmental policy since the mid-1990s (Stewart, 2010); more recently, the Collaboration for Environmental Evidence (CEE) has been established to encourage conduct of SRs on a wide range of environmental topics (Collaboration for Environmental Evidence, 2015).

The potential advantages of adapting SR methodology to the field of chemical risk assessment (CRA) have also been recognised, with multiple research groups and organisations either developing and adopting (Woodruff and Sutton, 2014; Birnbaum et al., 2013; European Food Safety Authority, 2010; Rooney et al., 2014; Aiassa et al., 2015) or recommending (US National Research Council, 2014a, 2014b; US Environmental Protection Agency, 2013; Silbergeld and Scherer, 2013; Hoffmann and Hartung, 2006; Zoeller et al., 2015) the use of SR methods for evaluating the association between health effects and chemical exposures to inform decision-making. There are, however, a number of recognised challenges in extending SR methods to CRA, many of which derive from key differences in the evidence base between the healthcare and toxicological sciences.

SRs in medicine often focus on direct evidence for benefits and adverse effects of healthcare interventions derived from randomised controlled trials (RCTs) in humans. The evidence base for CRA is generally more complex, with a need to extrapolate from investigations in animals, in vitro and in silico, and then to synthesise findings with those from human studies if available. Furthermore, the human data tend to come from observational studies with greater and more varied potential for bias and confounding than RCTs, and the range of outcomes to be considered is usually much wider than in the assessment of healthcare interventions. Thus, when the various types of toxicological research are combined into a single overall conclusion about the health risks posed by a chemical exposure, reviewers are challenged with integrating the results from a broad and heterogeneous evidence base.

In spite of these differences, there is reason for thinking that SR methods can be applied successfully to CRA. For example, techniques for aggregating the results of different study types are already addressed in various frameworks currently in use in toxicology. These include: International Agency of Research on Cancer (IARC) Monographs (International Agency for Research on Cancer, 2006); the Navigation Guide (Woodruff and Sutton, 2014); and the US Office for Health Assessment and Translation (OHAT) (Rooney et al., 2014; US National Toxicology Panel, 2015) – though it should be noted that none of these approaches have yet applied SR methods to the exposure assessment component of CRA. Heterogeneous sources of evidence are a familiar challenge in all domains including clinical medicine (Lau et al., 1998), and SR of observational studies has a crucial role in identifying complications and side-effects of healthcare interventions (Sterne et al., 2014; Higgins and Green, 2011). The need for SR of pre-clinical animal trials of healthcare interventions, in order to better anticipate benefits and harms to humans, is another area in which methods being developed and implemented by a number of groups including SYRCLE (Hooijmans et al., 2012; van Luijk et al., 2014) and CAMARADES (Macleod et al., 2005; Sena et al., 2014). (Stewart and Schmid, 2015) argue that research synthesis methods (including systematic review) are generic and applicable to any domain if appropriately contextualised.

Given the sometimes controversial outcomes of CRAs and the growing public and media profile of the risks that chemicals may pose to humans and the environment, SR is increasingly viewed as a potentially powerful technique in reassessing and communicating how likely it is that a chemical will cause harm. SR methods add transparency, rigour and objectivity to the process of collecting the most relevant scientific evidence with which to inform policy discussions and could provide a critical tool for organising and appraising the evidence on which chemical policy decisions are based.

Consequently, in November 2014 a group of 35 scientists and researchers from the fields of medicine, toxicology, epidemiology, environmental chemistry, ecology, risk assessment, risk management and SR participated in a one-day workshop to consider the application of SR in CRA. The purpose was three-fold:

1. Identify from expert practitioners in risk assessment and SR the obstacles, in terms of practical challenges and knowledge gaps, to implementing SR methods in CRA;
2. Develop a “roadmap” for overcoming those obstacles and expediting the implementation of SR methods, where appropriate, by the various stakeholders involved in CRA;
3. Establish the foundations of a network to co-ordinate research and activities relating to the implementation of SR methods in CRA. The aim would be to support best practise in the application of SR techniques and promote the wider adoption of SR in CRA, both in Europe and elsewhere.

Participants heard seven presentations about recent developments in SR methods, their application to the risk assessment process, and their potential value to policy-makers. There were two break-out sessions in which participants were divided into three facilitated groups, firstly to discuss challenges to implementing SR methods in CRA, and then to suggest ways in which the obstacles could be overcome. These ideas were discussed in plenary before being summarised, circulated for comment and then published in this paper. The Workshop was conducted under the “Chatham House Rule” such that participants were free to refer to the information presented and discussed, provided they did not attribute it to identifiable individuals or organisations.

---

1 It is worth drawing a distinction between three sources of bias in the review process. There is potential for bias in the conduct of a review (e.g. because of inappropriate methods for identifying and selecting evidence for inclusion in the review); bias because the material available for the review is not representative of the evidence base as a whole (due to selective publication); and bias arising from flaws in the design, conduct, analysis and reporting of individual studies included in the review that can cause the effect of an intervention or exposure to be systematically under- or over-estimated. One of the major functions of SRs is to minimise bias in the conduct of a review, and, as far as possible, to ensure that potential bias from selective publication and methodological flaws in the evidence are properly taken into account when drawing conclusions in response to a research question.
The purpose of this overview paper is to present the rationale for exploring the application of SR methods to CRA, the various experts’ views on the challenges to implementing SR methods in CRA, and their suggestions for overcoming them. The remaining goals of the meeting are ongoing work, including the development of the roadmap concept for publication and the establishment of a network for supporting the use of SR in CRA.

2. The appeal of SR methods in CRA

Chemical risk assessment is a multi-step process leading to a quantitative characterisation of risk, which can then be used to inform the management of chemical substances so as to ensure that any risks to human health or the environment are managed optimally. CRAs entail four fundamental steps: hazard identification; hazard characterisation (often a dose–response assessment); exposure assessment; and risk characterisation (see Fig. 1). These steps draw on various fields of scientific research including environmental chemistry, toxicology (encompassing in vivo, in vitro, ecotoxicological and in silico methods), ecotoxicology, human epidemiology, and mathematical modelling.

There are many ways in which errors can occur in the interpretation of evidence from these varied disciplines, including failure to consider all relevant data, failure to allow appropriately for the strengths and limitations of individual studies, and over- or underestimating the relevance of experimental models to real-world scenarios (to name a few). Whether the appraisal of evidence is based on objective processes, or on subjective expert judgement and opinion, may also be an important factor in accurate interpretation of evidence: the assessment process always requires input from technical experts, which inevitably brings an element of subjectivity to the interpretation of the scientific evidence. Different experts may have varying degrees of practical and cognitive access to relevant information, place differing weight on individual studies and/or strands of evidence that they review and, when working in committee, may be more or less influenced by dominant personalities. This can result in misleading conclusions in which the potential for health risks is overlooked, underestimated or overstated. Furthermore, if the factors determining their assessment of evidence are undocumented, when expert opinions are in conflict it can be very challenging to distinguish which opinion is likely to represent the most valid synthesis of the totality of available evidence.

A recent illustrative example (see Box 1) of when expert scientists and reputable organisations have come to apparently contradictory conclusions about the likelihood of a chemical causing harm is the case of bisphenol-A (BPA). BPA is a monomer used in the manufacture of the resinous linings of tin cans and other food contact materials such as polycarbonate drinks bottles. It has been banned from use in infant-feed bottles across the EU (European Commission, 1/28/2011) because of “uncertainties concerning the effect of the exposure of infants to Bisphenol A” (European Commission, 5/31/2011b).

The European Food Safety Authority (EFSA) considers that current levels of exposure to BPA present a low risk of harm to the public (European Food Safety Authority, 2015a). The French food regulator ANSES takes a seemingly different stance on the risks to health posed by BPA (French Agency for Food, Environmental and Occupational Health, and Safety, 4/7/2014), determining there to be a “potential risk to the unborn children of exposed pregnant women”. On this basis, ANSES has proposed classifying BPA as toxic to reproduction in humans (French Agency for Food, Environmental and Occupational Health, and Safety, 2013), a proposal which has contributed to the French authorities’ decision to implement an outright ban on BPA in all food packaging materials (France, 12/24/2012). While the ban has been challenged by some stakeholders as being disproportionate under EU law (Tošenovský, 2014, 2015; Plastics Europe, 2015), the Danish National Food Institute has argued that EFSA has overestimated the safe daily exposure to BPA and that some populations are exposed to BPA at levels higher than can be considered safe (National Food Institute, Denmark, 2015); a view reflected in the conclusions of some researchers, e.g. (Vandenberg et al., 2014) but not others, e.g. (US Food and Drug Administration, 2014).

The example of BPA illustrates the challenges in reaching consensus even when interpreting the same evidence base regarding the potential toxicity of chemical exposures, either in terms of what is known and what is uncertain about the risks to health posed by BPA, and/or what response is appropriate to managing those risks and uncertainties. It also shows how, in the absence of that consensus, there is a danger that policy on BPA may become disconnected from the evidence base, either risking harm to health through continued exposure or incurring...
unnecessary economic costs through restricting the use of a chemical which is in fact sufficiently safe. It also suggests that if the reasons for disagreement about health risks posed by a chemical are not accessible to various stakeholders in the debate, it then becomes much more difficult for regulators to credibly resolve controversies about chemical safety, potentially undermining their authority in the long term.

This example highlights the potential for differences in the interpretation of evidence when assessing chemical toxicity and the need for a process that is not only scientifically robust but also transparent, so that the reasons for any disagreement can be readily identified – including giving stakeholders greater opportunity to understand when differences in policy stem from divergent assessments of risk, and when they stem from divergent opinions as to how those risks are best managed. It also suggests the importance of the following characteristics in risk assessments that are used to inform risk management decisions:

1. **Transparency**, in that the basis for the conclusions of the risk assessment should be clear (otherwise they may not be trusted and errors may go undetected).
2. **Validity**, in that CRAs should be sufficiently (though not necessarily maximally) scientifically robust in their methodology and accurate in their estimation of risks and characterisation of attendant uncertainties as to optimise the decisions that must be made in risk management.
3. **Confidence**, providing the user with a clear statement as to the overall strength of evidence for the conclusions reached and a characterisation of the utility of the evidence for decision-making (e.g. "appropriate for hazard identification but inappropriate for identification of a reference dose").
4. **Utility**, in that the output of the risk assessment should be in a form that is convenient and intelligible to those who will use it (outputs that are too detailed and complex to validate and readily comprehend lead to inefficiency and possibly erroneous decisions).
5. **Efficiency**, providing a clear justification of the choice of research question in the context of efficiently solving a CRA problem. Resources for CRA are often limited and it is wasteful to expend unnecessary effort on aspects of an assessment that will not be critical to decision-making (although for the purposes of transparency and validity, the reasons for focusing on a particular outcome or otherwise restricting the evaluation should be explained).
6. **Reproducibility**, in that the conclusions of the SR process when applied to the same question and data should ideally produce the same answer even when undertaken by different individuals (also described as “consistency”). In practise, different experts may reach different conclusions because they will not all make the same value judgments about the scope, quality and interpretation of evidence. Therefore, the process should be sufficiently rigorous that it is highly likely that scientific judgement would result in the same conclusion independent of the experts involved, and as a minimum the SR process should render transparent the reasons for all conclusions.

It may be perceived that the value of SR methods lies in their provision of unequivocal assessments of whether or not a chemical will induce specific harm to humans and/or wildlife in given circumstances. In practise, however, this will happen only if the evidence base is sufficiently extensive, there is unanimity in identification of the problem and in assessment of the quality of the evidence base, and also how the evidence is to be interpreted in answering the review question (without this, SRs will also produce different results). Often, the consensus and/or information may be relatively limited; in such circumstances, a SR will instead clearly state the limitations of the available data and consequent uncertainties. The value here is in the provision of a comprehensive and transparent assessment of what is not known and insight into the drivers of divergent opinion. From a research perspective, this yields valuable information about how research limitations and knowledge gaps contribute to ongoing uncertainty about environmental and health risks, allowing the subsequent efforts of researchers to be more clearly focused. From a policy perspective, SRs offer a transparent explanation as to why there are differences in opinion which can then be communicated to stakeholders.

Overall, SR contributes to achieving consensus not by eliminating expert judgement, nor by eliminating conflicting opinions about whether a compound should be banned (for example), but by providing a robust, systematic and transparent framework for reviewing evidence of risks, such that when there is disagreement, the reasons for it are clearly visible and the relative merits of differing opinions can be appraised. In this way, it may help to resolve controversies in the interpretation of the science which informs the risk management process.
3. SR and its application to CRA

3.1. Traditional vs. SR methods

SR methods are often contrasted with “traditional”, non-systematic narrative approaches to describing what is and is not already known in relation to a research question. In reality, the distinction between systematic and narrative review is a crude one, with narrative reviews encompassing a number of different approaches to reviewing evidence, from the caricature of one researcher writing about “my field, from my standpoint […] using only my data and my ideas, and citing only my publications” (Caveman, 2000), to thorough narrative critiques of comprehensively identified evidence relevant to answering an explicitly articulated question, as conducted by organisations such as IARC (International Agency for Research on Cancer, 2006).

Nonetheless, it is worth noting that only relatively recently has it been recognised that traditional narrative reviews are, to varying degrees, vulnerable to a range of methodological shortcomings which are likely to bias their summarisation of the evidence base (Chalmers et al., 2002). These include selective rather than comprehensive retrieval of evidence relevant to the review topic, inconsistent interpretation of the impact of methodological shortcomings on the validity of included studies, and even an absence of clear review objectives or conclusions which are drawn directly from the strengths and limitations of the evidence base (Mulrow, 1987; Mignini and Khan, 2006).

The presence of these shortcomings seriously challenges the reader’s ability to determine the credibility of a review. When there exist multiple competing reviews, each using opaque methods, it becomes almost impossible to judge their relative merits and therefore to base decisions on current best available evidence. The consequence is a proliferation of conflicting opinions about best practice that fail to take proper account of the body of research evidence. In the healthcare sciences, this was initially shown by Antman and colleagues when they found that, in comparison to recommendations of clinical experts, systematic aggregation of data from existing clinical trials of streptokinase to treat myocardial infarction would have demonstrated benefit some years before recommendations for its use became commonplace (Antman et al., 1992). More recently, cumulative meta-analyses have been shown to be more accurate in summarising current understanding of the size of effect of a wide range of healthcare interventions than researchers planning new clinical trials who have not used these methods (Clarke et al., 2014).

A SR is an approach to reviewing evidence which specifically sets out to avoid these problems, by methodically attempting “to collate all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question,” using “explicit, systematic methods that are selected with a view to minimising bias” (Higgins and Green, 2011).

In detail, this amounts to the pre-specification of the objective and methods of the SR in a written protocol, in which the aim of conducting the review is clearly stated as a structured question (for a SR of the effects of an intervention or exposure, this can establish a testable hypothesis or quantitative parameter that is to be estimated), along with the articulation of appropriate methods. The methods specified should include the techniques for identifying literature of potential relevance to the research question, the criteria for inclusion of the studies of actual relevance to the research question, how the internal validity of the included studies will be appraised, and the analytical techniques used for combining the results of the included studies. The purposes of the protocol are to discourage ad-hoc changes to methodology during the review process which may introduce bias, to allow any justifiable methodological changes to be tracked, and also to allow peer-review of the work that it is proposed, to help ensure the utility and validity of its objectives and methods.

The final SR itself consists of a statement of the objective, the search method, the criteria for including relevant studies for analysis, and the results of the appraisal of internal validity of the included studies, e.g. implemented as a “risk of bias” assessment in Cochrane Reviews of randomised trials (Higgins et al., 2011). The evidence is then synthesised using statistical meta-analytical techniques, narrative methods or both (depending on the extent to which meta-analysis is possible) into an overall answer to the research question. An assessment is then made of the strength of the evidence supporting the answer; in Cochrane reviews, this typically follows the GRADE methodology (Atkins et al., 2004), taking into account overall features of the evidence base including risk of bias across the included studies, publication bias in the evidence base, external validity or applicability of the evidence to the population of interest, heterogeneity of the evidence, and the overall precision of the evidence. This is finally followed by a concluding interpretation of what the SR as a whole determines is and is not known in relation to its objective.

In this, we emphasise the distinction between a SR and a meta-analysis. A meta-analysis pools the results of a number of separate studies in a single statistical analysis and may be a component of a SR; however, it does not necessarily incorporate the full set of methodological features which define the SR process (e.g. a meta-analysis may or may not include an assessment of the internal validity of included studies). While we acknowledge that some researchers use the terms “systematic review” and “meta-analysis” interchangeably, we believe the two approaches should be disambiguated. It is also worth noting that many reviews employ a combination of narrative and systematic methods; there were differing opinions among workshop participants as to the extent to which it is reasonable to expect all reviews to fully incorporate SR methods.

3.2. The current status of SR in environmental health, toxicology and CRA

While the use of SR methodologies is well established in healthcare to determine the effect of interventions on health outcomes or the accuracy of a diagnostic test, application of SR is relatively novel in the fields of toxicology and environmental health. Workshop participants heard how methods for SR of medical interventions have in the United States been adapted in both academic and federal contexts to the gathering and appraising of evidence for the effects of chemical exposures on human health: researchers at the University of California have developed the Navigation Guide (Woodruff and Sutton, 2014), and the US Office of Health Assessment and Translation (OHAT) at the US National Toxicology Program has developed the OHAT Framework for systematically reviewing environmental health research for hazard identification (Rooney et al., 2014).

The two approaches adopt the key elements of SR methods to questions in environmental health (which is directly relevant to the CRA process but does not include assessment of dose–response). Features that the two approaches have in common include: conducting a SR according to a pre-specified protocol; the development of a specific research question and use of “PECO” statements (see Box 2) in systematising review objectives and the methods that will be used to answer that question; an approach to appraising the internal validity of included studies adapted from the risk of bias appraisal tool developed by the Cochrane Collaboration (Higgins et al., 2011); an adaptation of the GRADE methodology (Atkins et al., 2004) for describing the certainty or strength of a body of evidence, incorporating risk of bias elements with other criteria such as for the assessment of relevance or internal validity of included studies. It is a term used in Cochrane Collaboration guidance on conduct of SRs and recommendations, Environ Int (2015), http://dx.doi.org/10.1016/j.envint.2015.11.002

Please cite this article as: Whaley, P., et al., Implementing systematic review techniques in chemical risk assessment: Challenges, opportunities and recommendations, Environ Int (2015), http://dx.doi.org/10.1016/j.envint.2015.11.002
“PECO” is an acronym representing: Population (the exposure group of interest, e.g. people of a certain age or rats in laboratory studies); Exposure (the compounds or exposure scenarios of interest, e.g. respiratory exposure to fine particulate matter); Comparator (the group to which the exposure group is being compared, e.g. vehicle-exposed controls in laboratory experiments or less exposed groups in epidemiological studies); Outcome (a deleterious change or marker thereof hypothesised to be brought about by the exposure). The purpose of a PECO statement is to provide a framework for developing the key question which a SR will answer, and also to determine the rationale for the inclusion and exclusion criteria that explicitly define which studies are relevant for the review.

Box 2. The use of PECO statements in the SR process.

3.3. Overcoming the challenges in implementing SR methods in CRA

Risk assessment for a chemical or group of chemicals is a multi-faceted process that normally requires consideration of multiple endpoints in relation to a variety of exposure scenarios, integrating evidence from epidemiological studies, bioassays in animals, mechanistic studies and studies on the distribution and determinants of exposure by different pathways and routes. In addition to resolving methodological issues relating to underdeveloped methods (e.g. how SR methods can be used as part of dose–response assessment or how they can be applied to exposure assessment), it is important to consider how SR should fit into the CRA process. One challenge going forward is to explore the circumstances in which applying more rigorous SR methods to assess scientific evidence would be warranted, which would require insight into the practicality and cost-effectiveness of applying such methods in those situations.

In principle, it should be possible to conduct SRs in any aspect of a CRA. Given the success in employing SR methods to support evidence-based practice in healthcare, it is intuitive that SRs could address specific aspects of CRA. Given the success in employing SR methods to support evidence-based practice in healthcare, it is intuitive that SRs could address specific aspects of CRA. On the other hand, it is clear that certain aspects of CRA are more amenable to SR methodology than others. For example, topics such as biological effects, where there is more agreement on the type of evidence that should be included, may be more suitable for SR methodology than topics such as exposure assessment, where there is more variability in the types of evidence that should be included.

Other tools are being developed to contribute to the systematic assessment of in vivo and ecotoxicity studies which have not been directly derived from Cochrane Collaboration methods. Presented at the Workshop was Scirap (Science in Risk Assessment and Policy), a system developed to improve the consistency with which the relevance and reliability of studies are appraised in the context of conducting a chemical risk assessment for regulatory purposes. It is also intended to reduce the risk of selection bias in the risk assessment process by providing a mechanism for including non-standardised study methods yielding potentially valuable data (Beronius et al., 2014; Scirap, 2014).

There are a number of other initiatives promoting and developing the use of SR methodologies in environmental and chemical risk assessment. Participants heard about how the European Food Safety Authority is integrating SR methods into its assessments of food and feed safety (European Food Safety Authority, 2015b, 2015c), and about the UK Joint Water Evidence Group methods for rapid and systematic assessments of evidence (Collins et al., 2014). Other coordinated initiatives include the Evidence-Based Toxicology Collaboration (Hoffmann and Hartung, 2006); the Collaboration for Environmental Evidence (Bilotta et al., 2014a; Land et al., 2015); and the Systematic Review Centre for Laboratory Animal Experimentation (SYRCLE).

3.4. Priorities for expediting the adaptation of SR methods to CRA

The priorities for expediting the adaptation of SR methods to CRA identified at the Workshop are as follows:

1. The development of a number of prototype CRA-focused SRs to explore how readily SR procedures can be integrated into the CRA process, to:
   a. identify additional methodological challenges in adapting SR methods to the CRA context and develop techniques to address them;
   b. acquire practical experience in managing resources when conducting SRs in CRA, including the conduct of scoping exercises for identifying high-value review questions, the further development and/or application of novel “rapid evidence review” methods (UK Civil Service, 2015), and how SR methods can be integrated into existing regulatory structures such as REACH (see Box 3 (European Chemicals Agency (2/26/2015)));

2. Technical development of SR methodologies for CRA purposes, in particular the further advancement of techniques for appraising and synthesising mechanistic, toxicological and human epidemiological studies, to include:
   a. refining tools for more consistent and scientifically robust appraisal of the internal validity of individual studies included in a CRA and the implications for interpretation of their findings; see e.g. Bilotta et al. (2014b).

This might include further development and validation of tools such as the SYRCLE methodology for assessing the internal validity of animal studies (Hooijmans et al., 2014); for SR of observational studies see e.g. Sterne et al. (2014),

Please cite this article as: Whaley, P., et al., Implementing systematic review techniques in chemical risk assessment: Challenges, opportunities and recommendations, Environ Int (2015), http://dx.doi.org/10.1016/j.envint.2015.11.002
4. Conclusions

While systematic review methods have proven highly influential in healthcare, they have yet to make widespread impact on the process of chemical risk assessment. While there is much promise in the concept of adapting SR methods to CRA to give definitive answers to specified research questions, or to enable identification of the reasons for failure to resolve debate, a number of challenges to implementing SR methods in CRA have been identified. These include particular concerns about approaches to assessing bias and confounding in observational studies, the effort involved in conducting SRs, and the subsequent benefits of conforming to SR standards. Recent experience from both regulatory agencies and academics already yields some clear recommendations which would expedite the wider implementation of SR methods in CRA, potentially increasing the efficiency, transparency and scientific robustness of the CRA process.

Acknowledgements

The views expressed in this manuscript are those of the authors and do not necessarily represent the views or policies of their employers or otherwise affiliated organisations. EA is employed by the European Food Safety Authority (EFSA); however, the present article is published under her sole responsibility and may not be considered as an EFSA scientific output.

References


Caveman, 2000. The invited review — or, my field, from my standpoint, written by me using only my data and my ideas, and citing only my publications. J. Cell Sci. 113 (Pt 18), 3125–3126.
Chalmers, I., Glasziou, P. 2009. Avoidable waste in the production and reporting of re-
search evidence. Lancet 374 (9683), 86–89. http://dx.doi.org/10.1016/S0140-6736(09)
12555-8.

Chalmers, I., Hedges, L.V., Cooper, H. 2002. A brief history of research synthesis. Evalua-

Clarke, R., Hopewell, S., Flach, V., 2014. Accumulating research: a systematic account of how cumu-
lative meta-analyses would have provided important knowledge, improved health, reduced
downstream and reduced resources. PLoS ONE 9 (7), e102670. http://dx.doi.org/10.1371.
journal.pone.0102670.

Collaboration for Environmental Evidence (2015): The Collaboration for Environmental Evi-


European Food Safety Authority, 2015c. Principles and process for dealing with data and
evidence in scientific and non-scientific areas. Available online at http://www.anses.fr/en/
content/bisphenol-a-exposure.html, accessed 2/18/2015.

European Food Safety Authority (2015a): No consumer health risk from bisphenol A ex-
F6553AACC19D178279D8DF154EAC8558.tpdila17v_1?cidTexte=
19442 du 24 décembre 2012 visant à la suspension de
l'utilisation de certains fluoralkyl substances on the concentra-
tions of perfluoroalkyl acids and their precursors in the environment? A systematic re-


(3), 485–488.

food.dtu.dk/english/~media/Institutter/FoedevarerensInstitut/Publication; Pub-

Nordic Cochrane Centre (2014): Review Manager (RevMan). Version 5.3: Cochrane Col-

Plastics Europe (1/5/2015): French ban on the use of Bisphenol A (BPA) in food contact:
In conflict with European law and risk assessment—severe distortion of the market—
bisphenol-a-europe.org/uploads/Modules/Mediarioum/stm_re_french_bpa-ban-

1307972.

Royaj, N., Valdsteino, A. (2013): Health Assessment Workspace Collaborative (HAWC). Ver-
sion Solid Hammer: UNC-CH Software. Available online at https://hawcproject.org/

tegrating the value and reducing waste in biomedical research regulations and manage-
62297-7.

Schünemann, H.J., Moja, L. 2015. Reviews: Rapid! Rapid! Rapid! ... and systematic

mental Science at Stockholm University; Institute of Environmental Medicine at Karolinska Institutet in Stockholm; MistraPharma. Available online at http://www.

Sena, E.S., Currie, G.L., McCann, S.K., Macleod, M.R., Howells, D.W. 2014. Systematic re-
views and meta-analysis of preclinical studies: why perform them and how to ap-
2288-14-5.

Silbergeld, E., Scherer, R.W. 2013. Evidence-based toxicology: strait is the gate, but the
road is worth taking. ALTEx 30 (1), 67–73.


1098/selphy.2009.0546.

the need for trans-disciplinary evidence synthesis methodologies. Research Synthesis

European Parliament, Parliamentary questions, E-004315-15, Subject: measures
to prevent the severe distortion of the market caused by the French ban on the use of Bisphenol A. The National Academies Press, Washington, D.C.

US National Research Council, 2014a. A Framework to Guide Selection of Chemical Alter-
 natives. The National Academies Press, Washington, D.C.

Assessment Systematic Review of Per-and Pseudoalkylated Substances on the concentra-
tions of perfluoralkyl acids and their precursors in the environment? A systematic re-

(97)08468-7.