Science of the Total Environment 595 (2017) 537–546

Contents lists available at ScienceDirect

Science of the Total Environment

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

Strategic risk appraisal. Comparing expert- and literature-informed consequence assessments for environmental policy risks receiving national attention

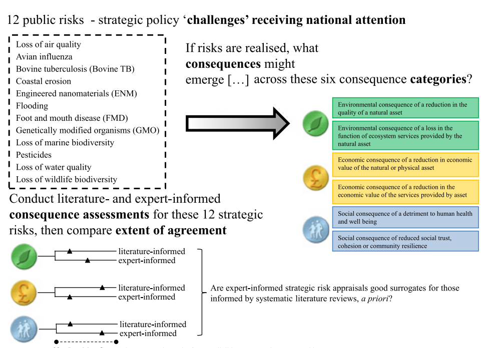
Jérôme Dagonneau [a], Sophie A. Rocks [a,*], George Prpich [a], Kenisha Garnett [a], Edgar Black [b], Simon J.T. Pollard [a]

[a] Cranfield University, Cranfield Institute for Resilient Futures, School of Water, Energy and Environment, Cranfield, Bedfordshire, MK43 0AL, United Kingdom
[b] Department for Environment, Food and Rural Affairs, Nobel House, 17 Smith Square, London, SW1P 3JR, United Kingdom

HIGHLIGHTS

- Strategic risk appraisals frequently rely on the opinions of technical policy staff and researchers in workshop settings
- For the first time, we compare expert- and literature-informed consequence assessments for 12 strategic environmental risks
- Of 36 literature- and expert-informed assessments compared, only 8 couples were statistically distinct
- Expert-informed consequence assessment appears a robust surrogate for *a priori* literature informed assessment
- When designing workshops, full representation of the risks is required, especially for socioeconomic assessments

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 22 February 2017
Received in revised form 31 March 2017
Accepted 31 March 2017
Available online xxxx

Editor: D. Barcelo

Keywords:
Risk
Strategic
Environment
Policy
Expert
Appraisal

ABSTRACT

Strategic risk appraisal (SRA) has been applied to compare diverse policy level risks to and from the environment in England and Wales. Its application has relied on expert-informed assessments of the potential consequences from residual risks that attract policy attention at the national scale. Here we compare consequence assessments, across environmental, economic and social impact categories that draw on 'expert'- and 'literature-based' analyses of the evidence for 12 public risks appraised by Government. For environmental consequences there is reasonable agreement between the two sources of assessment, with expert-informed assessments providing a narrower dispersion of impact severity and with median values similar in scale to those produced by an analysis of the literature. The situation is more complex for economic consequences, with a greater spread in the median values, less consistency between the two assessment types and a shift toward higher severity values across the risk portfolio. For social consequences, the spread of severity values is greater still, with no consistent trend between the severities of impact expressed by the two types of assessment. For the latter, the findings suggest the need for a fuller representation of socioeconomic expertise in SRA and the workshops that inform SRA output.

1. Introduction

Government departments managing public risks (Cabinet Office, 2012; Beddington, 2013) operate within budgetary limits established

* Corresponding author.
E-mail address: s.rocks@cranfield.ac.uk (S.A. Rocks).

http://dx.doi.org/10.1016/j.scitotenv.2017.03.293

within a political cycle and must prioritise the risks they manage across diverse policy portfolios. Some ministries have turned to policy-level (strategic) risk appraisals (SRA), among other inputs, to inform decisions on public expenditure so the highest residual risks can be considered for additional funding; assuming onward investment would reduce them further. At this level of analysis, it is the likelihood of a public national scale risk being realised that is being assessed (e.g. a substantive regional flood event; a national scale foot and mouth disease incursion), along with the associated consequences (environmental, economic and social) that might ensue. In SRAs, risks of varying character (Department for Environment, Transport and the Regions, 1998; Klinke and Renn, 2002; Pollard et al., 2004; Prpich et al., 2011) are appraised alongside one another and presented in a comparative analysis, often in a single schematic or 'heat map' (Prpich et al., 2013). This differs markedly from conventional environmental risk assessments where the analyst is concerned with estimating the likelihood of an adverse outcome (usually) in a spatiotemporal context; say the inhalation risk associated with emissions from a hazardous waste incinerator; or of hydrocarbon exposures to workers remediating a parcel of petroleum-contaminated land (see Defra, 2011 for examples).

For strategic risk appraisals, policy experts and their advisors assimilate expert knowledge at the policy level; drawing on the expertise of technical policy staff, their evidence programmes and the academic communities that research the risks being appraised; and then make well-reasoned judgements by interpreting the science base to compare risks alongside one another (Environment Agency, 2005; Mauelshagen et al., 2014). In SRA, expert-informed assessments of policy risk have become a practical, rapid surrogate for rigorous literature-informed assessments. But are we right to assume this surrogacy is valid? Do expert- and literature-informed assessments correlate?

Since the late 1980s, a substantive literature has grown around strategic environmental risk appraisal; to be specific, on the comparative analysis of multiple policy-level risks by environment ministries and their regulatory agencies (US Environmental Protection Agency, 1987; Morgan et al., 1996; German Advisory Council on Global Change, 1998; Feldman et al., 1999; Long and Fischhoff, 2000; Morgenstern et al., 2000; DeKay et al., 2001; Florig et al., 2001; Klinke and Renn, 2002; Morgan et al., 2001; Environment Agency, 2002, 2005; New Jersey Department of Environmental Protection, 2003; Pollard et al., 2004; Andrews et al., 2004; Linkov and Ramadan, 2005; Fischhoff and Morgan, 2009; International Risk Governance Council, 2011; Vlek, 2013). A compendium of techniques has been compared, their communication challenges described, the traction SRAs get with publics analysed and reviewed, and the visualisation of SRA outputs experimented with (Perhac, 1998; van Asselt, 1999; Klinke and Renn, 2002; Haimes et al., 2002; Environment Agency, 2002; Willis et al., 2004, 2010; Prpich et al., 2011, 2013; Soane et al., 2016; Rocks et al., 2017).

Researchers and advisors assisting policy officials with SRA have facilitated numerous workshops in which policy-level risks have been appraised, typically over one or two days, to inform outputs similar to those in Fig. 1. Fig. 1 style outputs are then used to stimulate discussions about the efficacy of existing risk management measures and the suitability of onward investments in risk management within available budgetary limits. Notwithstanding efforts to secure a representative make-up of technical expertise in SRA workshops, there has been little a priori analysis of whether the assessments garnered by them correlate with those gained from a more considered, albeit lengthy, analysis of the evidence for the same risks from the published literature. This paper attempts such an analysis and seeks to validate, or otherwise, the use of expert-informed SRA; our null hypothesis being '*there is no significant difference between literature- and expert-informed assessments of the environmental risks that attract national attention in SRA*'.

In SRA, evidence must be synthesised, simplified and made available for comparison, even though the risks appraised may seem incommensurate. The risks previously appraised by the authors (Prpich et al., 2011, 2013; Fig. 1) were associated with environmental hazards such as regional flooding, coastal erosion, pesticide impacts and engineered nanomaterials, and differed widely in their potential for harm, in how they were perceived and in the costs required to mitigate them (Science Advisory Council, 2012).

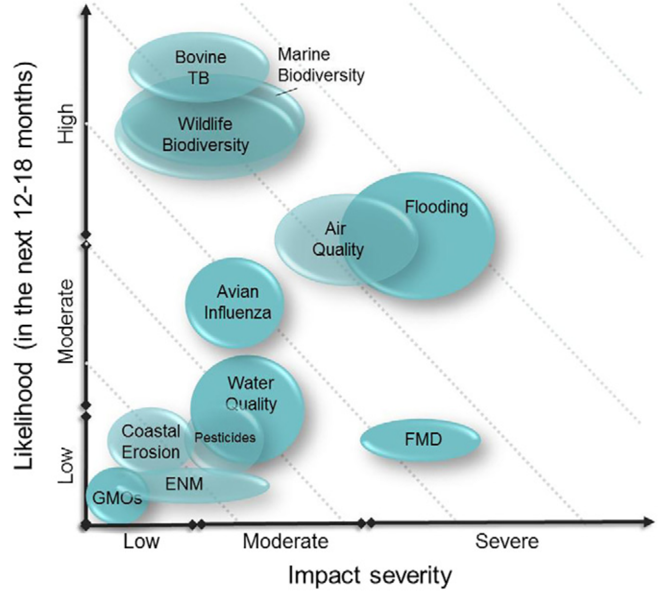


Fig. 1. Illustrative appraisal of 12 strategic risks for Defra (Science Advisory Council, 2012; Prpich et al., 2011, 2013). Ellipses reflect the relative magnitude and 2-dimensional uncertainty in likelihood and consequence (impact severity) for residual risks, assessed over a 12–18 month horizon, assuming existing risk management measures are in place. Their positions are informed through a flow of supporting evidence, independent analysis and deliberative process. Key for ellipses: GMOs genetically modified organisms; Bovine TB tuberculosis; ENM engineered nanomaterials; FMD foot and mouth disease.

In SRA workshops, policy experts appraise environmental risks that receive national attention, within a specific timeframe. Risks are expressed, usually logarithmically, in terms of a magnitude of likelihood and a severity of combined impact (Fig. 1); the latter being a mean of the impacts aggregated across environmental, economic and social categories of consequence. SRA methods are contentious among research and practitioner communities (Fischhoff and Morgan, 2009; House of Commons Science and Technology Committee, 2011; Cox, 2008; Vlek, 2013) because of the pragmatism necessary to deploy them and what are claimed as deviations from theoretical robustness in doing so (Cox, 2008). The authors, with others, have discussed the challenges and shortcomings of the methods elsewhere (Hofstetter et al., 2002; Pollard et al., 2004; Cox, 2008; Prpich et al., 2011, 2013; Vlek, 2013). This said, SRA tools see increasing use within Government and business circles (Beddington, 2013; Ernst and Young, 2010; World Economic Forum, 2011, 2017; Deloitte, 2013) and our research interest has been, therefore, to hone the tools used for strategic environmental risk appraisal (Duarte-Davidson et al., 1999; Environment Agency, 1999, 2004; Pollard et al., 2001, 2004; Prpich et al., 2011, 2013; Rocks et al., 2017). Particularly challenging for SRA workshops has been the assessment of consequences from residual risks, the aggregation of impacts and the overall presentation of their combined severity – the consequence assessment - which is the focus of this paper. A combined assessment of consequences and probabilities is necessary in SRA so to present the total risk for discussion. The probability assessment in SRA is not addressed here.

In this paper, we explore two sources of consequence assessment, experts and the literature, for 12 public risks. We have published a means of visualising differences between literature- and expert-informed assessments of consequence in SRAs (Fig. 2; Prpich et al., 2013; Dagonneau, 2013). In brief, the approach adopts six attributes of

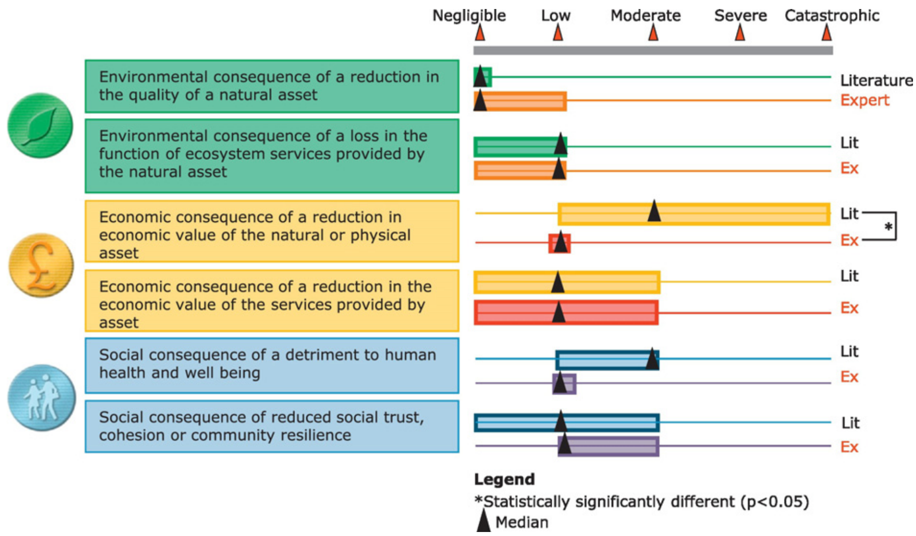


Fig. 2. Example expert- and literature-informed evaluations of consequences (n = 6) across environmental (leaf), economic (£) and social (people) consequence categories (2 attributes in each consequence category). Median severity values (denoted by black triangle) that differ with statistical significance in a Mann-Whitney *U* test as denoted by an asterisk (here, the consequences of a reduction in economic value of the natural or physical asset). This is a fictitious example for illustration only; see graphical abstract for engineered nanomaterials.

potential harm across three sustainability domains (environment, economic and social; represented by the leaf; £-sign and people infographics in Fig. 2; Prpich et al., 2011). For each strategic risk of concern, uncertainty in these consequence assessments is represented by the length of the wide bar along a severity axis, marked by ordinal levels from negligible to catastrophic harm; with literature- and expert-informed analyses presented together as a couple. A visual scan of the two bars in each of the six categories allows a ready assessment of alignment (or not) between these two sources of assessment, with significant differences between median assessments marked by an asterisk for the assessment couples (Fig. 2; economic consequence of a reduction […]).

In comparing the two types of assessment, we assess the validity of SRA workshops that often rely on expert input alone; frequently take place under a certain time pressure; and are occasionally assembled with sub-optimal representation of expertise across the gamut of risks appraised. A background comment on expert opinion is fitting. The quality of the opinion of an expert is difficult to verify. It is difficult to determine how 'expert' an expert is (Knol et al., 2010) and to assess the reliability and weight of opinion offered by an expert's interpretation of the evidence they have access to within a workshop setting, especially if their opinion is not supported by auditable data nor sources of verifiable information. In contrast, evidence from the peer reviewed literature appears more auditable, though is never completely transparent and, of course, open to its own biases. 'Rapid evidence appraisal', a tool used by Governments to appraise the scientific literature efficiently has emerged as one means of securing a representative and systematic review of the prior art as it relates to technical policy development (HM Treasury, 2011). It is hoped our analysis will be of value to risk practitioners in research, business and Government.

2. Materials and methods

The study was designed in conjunction with a Government department and examined 12 strategic risks attracting national attention within their policy portfolio (2011–12): poor air quality; poor water quality; foot and mouth disease (FMD); bovine tuberculosis (Bovine TB); avian influenza (AI); coastal erosion; regional flooding; genetically modified organisms (GMOs); engineered nanomaterials (ENM); reductions in wildlife biodiversity; reductions in marine biodiversity; and the use of pesticides. We deployed the methods in Prpich et al. (2011) using data collected through qualitative methods: systematic review and semi-structured interviews. The quantification, where possible, of impacts associated with these risks is detailed in Dagonneau (2013).

2.1. Literature-informed analysis

Evidence from the published literature was collected using a systematic process (Tranfield et al., 2003; Petticrew and Roberts, 2006) that included: (i) research questions; (ii) search strategy; (iii) appraisal of literature; (iv) data extraction; and (v) synthesis. Comprehensive searches were performed for each risk using Scopus™, ScienceDirect™ and Web of Knowledge™ and search terms in Boolean combinations (Fig. 3).

The outputs of these searches were reviewed and filtered to exclude documents not relevant to the six consequence categories of interest (boxes; Fig. 2). Additional sources were identified from relevant document bibliographies using a snowballing technique (Wohlin, 2014). Thus peer-reviewed papers and grey literature documentation informed the literature-based assessment of environmental, economic and social consequences for the 12 risks. The quality of relevant outputs was assessed using an evaluation matrix (Table 1; after Bowden, 2004), the evidence in each document being scored between 5 and 25 by reference to the quality indicators in Table 1. Evidence scored below 20 was excluded from the onward assessment to maintain presumed data quality. Table 2 summarises the outcome of this analysis.

Next, for documents supported by high quality evidence, key information related to impacts was extracted by reference to the 6 consequence categories (Fig. 2) and an impact severity from 'negligible' to 'catastrophic' assigned (Prpich et al., 2011; Appendix 1, Supplementary Information), with triangulation of this analysis completed by researcher SAR. For example, Evans et al. (2004) estimated flooding in the UK causes £1.04bn damage annually, rated here as 'severe' because economic damage between £1bn and £10bn was considered by policy officials, during construction of the matrix (Appendix 1), as a severe impact. The aggregated impact severity (i.e. environment, economy, social) was determined by the range of the values for the assessed evidence, offering a comprehensive analysis of the extent of impact across the three consequence categories.

2.2. Expert-informed analysis

For the expert-informed analysis, semi-structured interviews of technical policy experts from English government organisations and academia

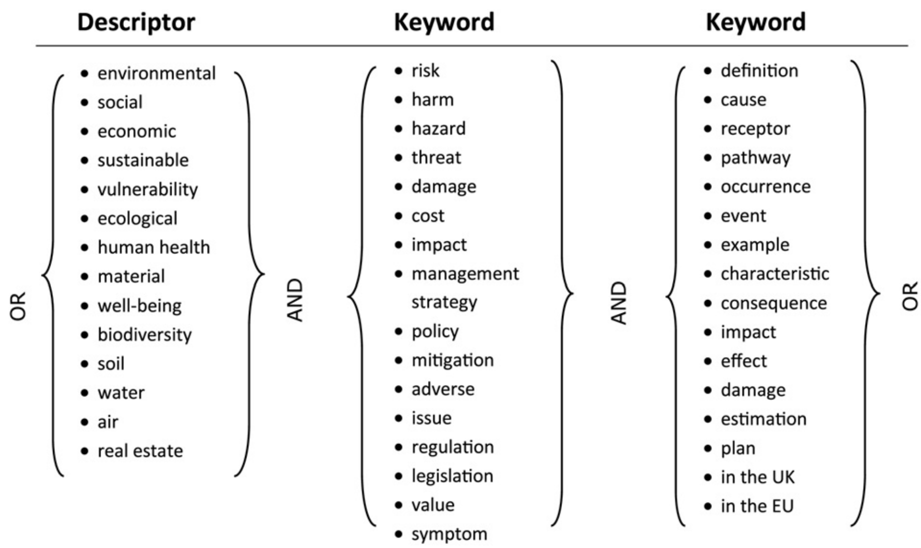


Fig. 3. Keywords and Boolean combinations used in systematic review.

were conducted. Assessments (2011) were elicited from policy professionals in Government organisations including Defra, the Environment Agency, the Centre for Environment, Fisheries and Aquaculture Science, the Joint Nature Conservation Committee, the Chemicals Regulation Directorate and also from domain-specific academics. Experts approached were also able to suggest participation from other experts they deemed to be skilled and reliable. At least three experts were identified for each of the risks studied. Generalist and normative experts were preferred, as the experts needed to be flexible so to consider evidence that differed from their held view, or that was at the margins of their specialism. Further, given the lexicon of risk terminology, to reduce any linguistic ambiguity, the principal researcher (JD) provided supporting material characterising the risk and the matrix used for assessing the policy risk. Interviews were designed by reference to the University's ethics policy and all experts were informed of their right to withdraw from the study at any point. The expert data were recorded anonymously.

Interviews followed a hybrid elicitation method, combining group and individual interviews (Keeney et al., 2001; Knol et al., 2010). For the consequences associated with each risk, at least three experts were selected to conduct the assessment. First, groups were offered a facilitated introduction where the researchers (JD, GP) presented the aims of the analysis and what was expected of the experts (Dagonneau, 2013 for detail). Individual open-ended interviews were then conducted (Robson, 2002) on the consequence categories. After the interviews, experts completed the matrix individually (Appendix 1; Supplementary Information) and supported their assessments with short narratives describing their rationale for judgement.

2.3. Comparing assessment types

This study was concerned with the agreement/disagreement between the two sources of data (p-statistic); and the dispersion/spread of impact severity scores between consequence categories. For the latter, the median, minimum and maximum values, lower and upper quartiles were determined for both the expert- and literature-informed assessment (Fig. 4, loss of air quality, literature-informed assessment, as an example).

Descriptive statistics and statistical tests (Mann-Whitney *U* test) were performed using Statsoft software, Statistica v11 (http://www.statsoft.co.uk/). For the expert- and literature-informed assessments, the central tendency was represented by the median, as the data were ordinal and the median preferred for illustrating the middle of the impact severity range rather than the most frequent. The non-parametric Mann-Whitney *U* test was used to identify statistically significant differences between the two types of assessment (denoted by * where this was the case) at a significance level of $p < 0.05$ for data with a small sample size ($n < 30$). The principal visuals are three schematics for environmental, economic and social consequences, across the risk portfolio (Figs. 6 to 8), where literature- and expert-based consequence assessments can be reviewed in tandem.

Fig. 2 style schematics were first constructed for each of the 12 risks evaluated (Appendix 2; Supplementary information; Dagonneau, 2013). Then, the two attributes of environmental (leaf), economic (£) and social (people) consequences for each risk were aggregated and presented in 3 comparator schematics (one each for environment, economic and social) across the whole risk portfolio to inform a discussion on the comparability

Table 1
Valuation of source quality (adapted from Bowden, 2004).

		Quality indicator					
		Theoritical basis	Scientifc method	Auditability	Validation	Objectivity	Score
Quality score	Very high	Well established theory	Best available practice; large sample; direct measure	Well documented trace to data	Independent measure of same variable	No discernible bias	Very high 20–25
	High	Accepted theory; high consensus	Accepted reliable method; small sample; direct measure	Poorly documented but traceable to data	Independent measurement of high correlation variable	Weak bias	High 15–20
	Moderate	Accepted theory; low consensus	Accepted method; derived data; analogue; limited reliability	Traceable to data with difficulty	validation measure not truly independent	Moderate bias	Moderate 10–15
	Low	Preliminary theory	Preliminary method; unknown reliability	Weak, obscure link to data	Weak, indirect validation	Strong bias	Low 5–10
	Very low	Crude speculation	No discernible rigour	No link back to data	No validation presented	Obvious bias	Very low 0–5
						Total Score	

Table 2
Summary of systematic review and literature appraisal.

Risk/activity	Screening and filtering the prior art				Number of lines of evidence relevant to the six consequence categories of interest					
	Body of literature	1st filter (abstract and title review)	2nd filter (full text review)	3rd filter (quality assessment of source)	Environmental consequence of a reduction in the quality of a natural asset	Environmental consequence of a loss in the function of ecosystem services provided by the natural asset	Economic consequence of a reduction in economic value of the natural or physical asset	Economic consequence of a reduction in the economic value of the services provided by asset	Social consequence of a detriment to human health and well being	Social consequence of reduced social trust, cohesion or community resilience
Loss of air quality	3678	369	26	14	8	8	5	1	11	2
Avian influenza	9722	265	24	14	3	3	3	0	4	5
Bovine tuberculosis (Bovine TB)	3154	154	47	24	2	12	13	2	6	4
Coastal erosion	1637	105	31	18	4	7	8	3	2	2
Engineered nanomaterials (ENM)	670	136	29	21	10	8	2	0	11	1
Flooding	4131	453	25	28	5	5	13	6	16	11
Foot and mouth disease (FMD)	9891	428	32	14	5	9	7	6	7	7
Genetically modified organisms (GMO)	1443	110	28	14	3	7	2	0	3	6
Loss of marine biodiversity	847	84	20	12	3	5	2	3	2	2
Pesticides	9842	236	28	17	6	14	2	3	8	2
Loss of water quality	1279	72	16	7	5	5	2	3	3	0
Loss of wildlife biodiversity	490	96	15	11	5	7	3	3	1	1

of the assessments and the SRA workshops that informed them (Figs. 6 to 8 below). The horizontal bars of impact severity were built using the full range (minimum and maximum) of values for each consequence assessment, in preference to the interquartile range that would have excluded the extremes (Cramer, 2004; Howitt and Cramer, 2011).

3. Results and discussion

3.1. Systematic review of consequences from public risks attracting national attention

Table 2 above reflects the limited extent to which high quality evidence is available on the consequence categories of interest for the risks studied. This is important and should inform the design and analysis of '*State of the Environment*' reports (e.g. European Environment Agency, 2015) that provide analyses of national or supra-regional environmental threats. Greater research effort is required to assimilate national level information on these risks. Note also that in practice, the analysis of consequences using quality-filtered literature sources secured a greater number of lines of evidence across the six categories than sources filtered (Fig. 5). This is because most of the filtered sources included statements relevant to more than one consequence category (Table 2).

Flooding and foot and mouth disease (FMD) were well-supported by evidence at the national scale (Fig. 5) including social impact evidence, reflecting the attention these risks have had in recent years in the UK (2000 onwards) in light of national events. Across the portfolio, the potential harm to the function of ecosystem services (Fig. 5; environmental; dotted stacked bar) was also well represented. Emergent risks

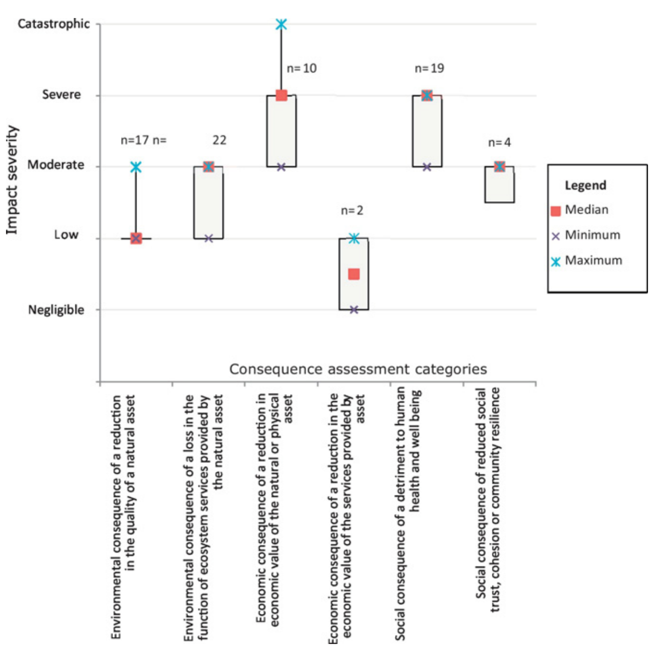


Fig. 4. Illustrative summary statistics representing the distribution of impact severity across 6 consequence categories from a literature-based assessment of evidence for the loss of air quality; showing median, interquartile range and range of values.

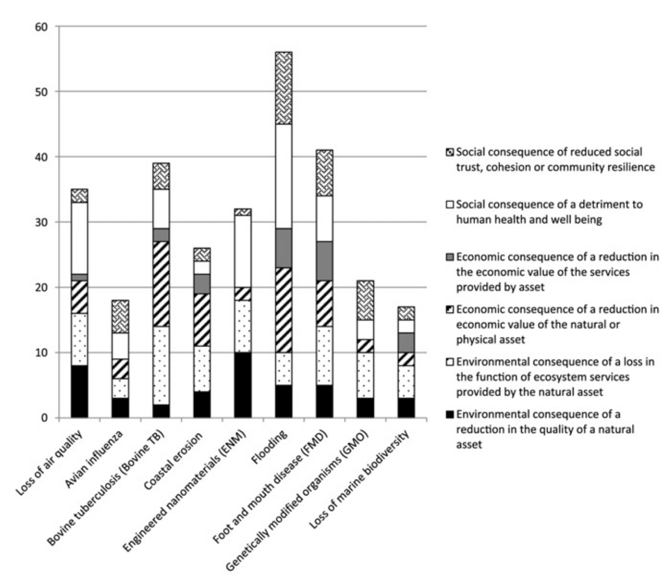


Fig. 5. The number of lines of evidence (n) from literature-informed assessment across 12 public risks to inform 6 categories of consequence assessment.

(avian influenza, GMO's and the potential loss of marine biodiversity) were less well represented and are clear candidates for evidence review and synthesis at the national scale. Appendix 2 (Supplementary information) provides the summary assessment for each risk. We comment on these elsewhere (Prpich et al., 2011). Of interest here is the comparison of assessment types for the consequence assessment that feeds into Fig. 1.

3.2. Comparing literature- and expert-informed consequence assessments

Our key results compare literature- and expert-informed assessments of the potential consequences from 12 environmental risks managed at a national level. These are presented in Figs. 6–8 allowing a summary comparison of alignment (or not) for the two types of assessment for each of three consequence categories. Each bar represents an aggregate of the two attributes of harm assessed for each category of consequence (as per Fig. 2; Appendix 2). The Mann Whitney *U* significance test identified assessment couples where medians were significantly different at $p < 0.05$ (Table 3).

The results offer insight on the appropriateness of relying on expert analysis alone in SRA and inform the development of SRA tools and workshops. The authors' experience of these tools and workshops is that consequence assessment becomes progressively challenging in the order: environment to economic to social. We discuss each consequence category in turn below. An important aside is the observation that virtually no published data exists on the relative likelihood (probability) of public environmental risks that attract national attention, other than that garnered by experts and their advisors within Government through the types of elicitation exercises that inform the national risk register (Cabinet Office, 2012; Beddington, 2013). Whilst scholars discuss individual consequences and localised aspects of specific risks, they rarely comment, if at all, on the likelihood of residual harms emanating from national risks; an observation that has attracted Government attention (Government Office for Science, 2011).

For environmental consequences, there is reasonable agreement between the two sources, with expert-informed assessments providing a narrower dispersion of severity and median values often similar in scale, with the exception of environmental consequences for avian influenza, bovine TB and the potential loss of wildlife biodiversity. Notably, flooding shows a considerable dispersion in severity values, reflected by the horizontal dimension of the flooding ellipse in Fig. 1.

The situation is somewhat more complex for the comparison of economic consequences. First we note greater spread in the values and less consistency between the two assessments with a general shift toward higher severity values across the risk portfolio reflecting, perhaps the substantive economic impacts of national scale events.

For the social consequences of harm across the 12 risks evaluated, the spread of severity values is greater still with no consistent trend between the severities expressed by the two assessments. We have observed that social concerns are often inadequately represented in SRAs, with facilitators reliant on policy colleagues' interpretations of non-Government organisation and citizen agendas often as a surrogate for deep social science expertise. Informed assessments are therefore critical.

Generally across the consequence categories we note the extremes of the ordinal 5-point scale are rarely used, one of Cox's (2008) criticisms, and that consequences judged as 'low' or 'moderate' were generally the norm, which is perhaps understandable given these are all policy risks being actively managed and so their assessment represents the extent of the residual risk remaining. The consequential 'bunching' of strategic risks is a feature of many SRAs (Fig. 1) and that can pose difficulties for the risk analyst as s/he seeks to distinguish between risk priorities and develop management strategies accordingly (Cox, 2008).

Next, the literature-informed assessments generally expressed a wider range (length of bar) of impact severity values (in 72% of the

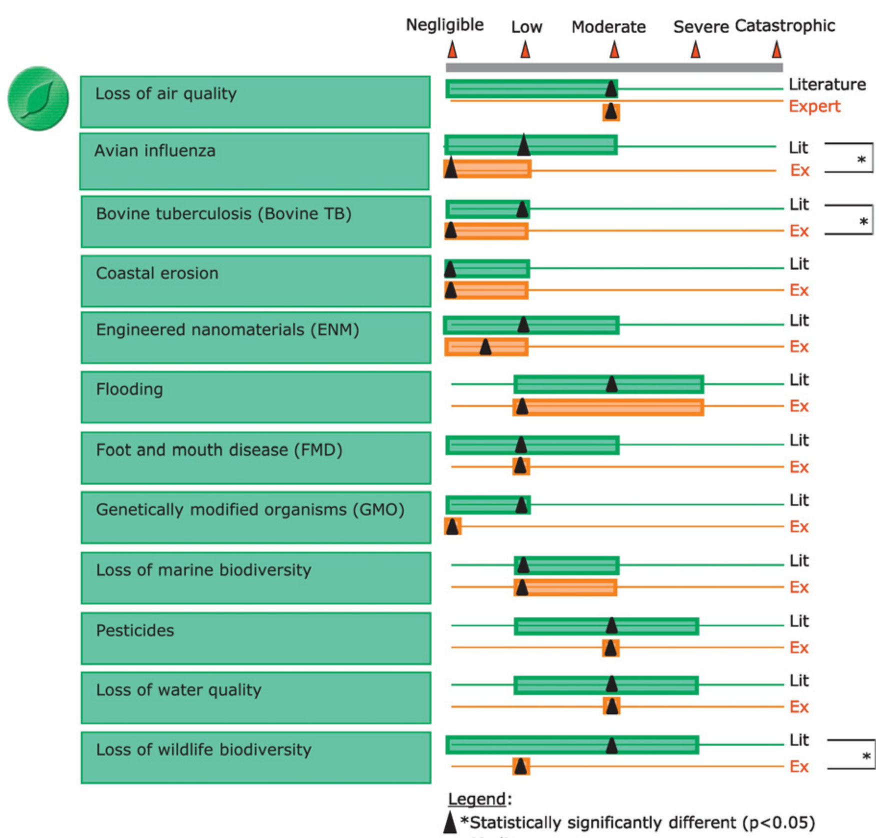


Fig. 6. Comparison of aggregate literature- and expert-informed assessments of potential environmental consequences from 12 environmental public risks.

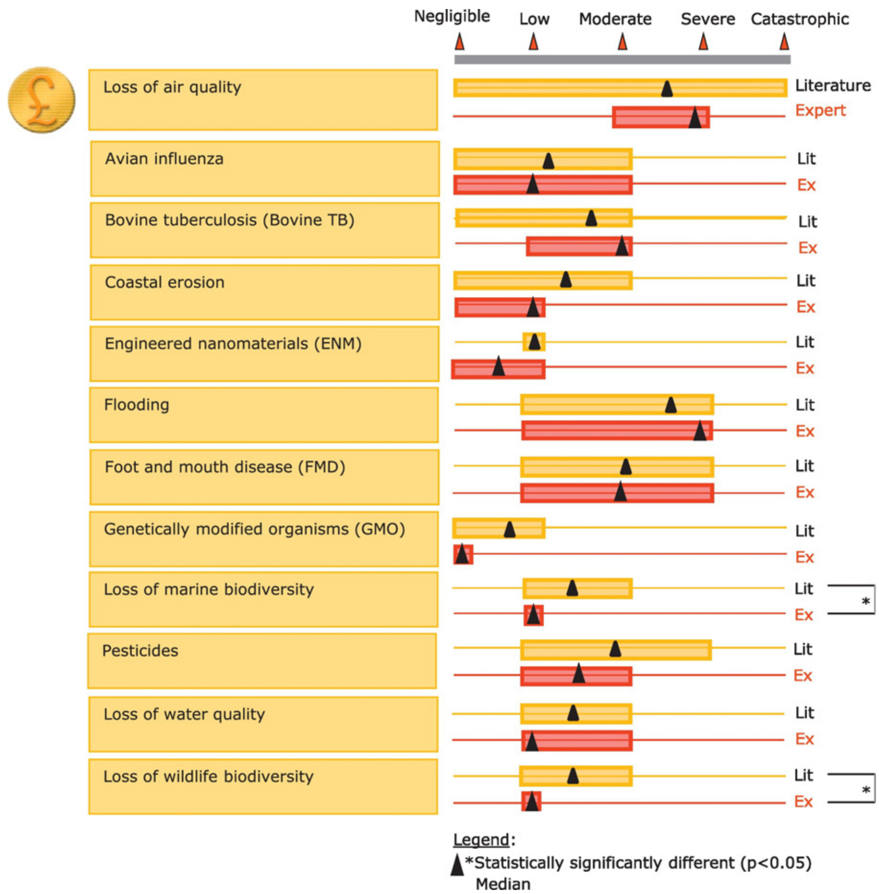


Fig. 7. Comparison of aggregate literature- and expert-informed assessments of potential economic consequences from 12 environmental public risks.

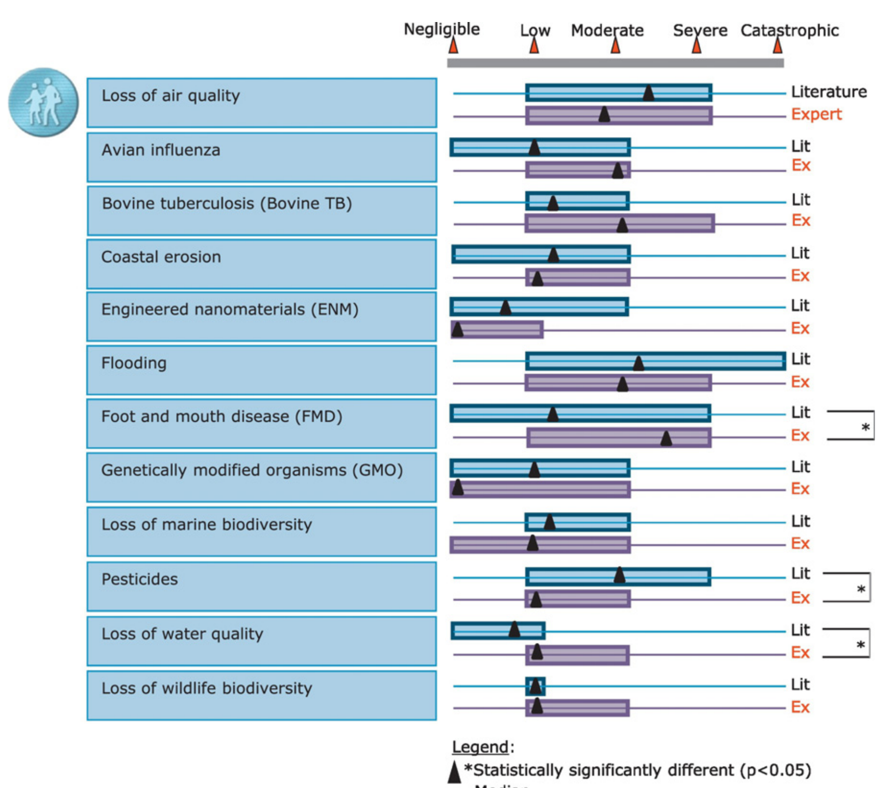


Fig. 8. Comparison of aggregate literature- and expert-informed assessments of potential social consequences from 12 environmental public risks.

Table 3
Mann Whitney *U* test for differences between literature- and expert-informed analyses across three consequence categories. Significance (shaded, starred) at values <0.05.

Risk	Consequence categories		
	Environment	Economy	Social
Loss of air quality	0.142	0.923	0.346
Avian influenza	0.035*	0.504	0.318
Bovine tuberculosis (Bovine TB)	0.011*	0.591	0.057
Coastal erosion	0.457	0.140	0.950
Engineered nanomaterials (ENM)	0.627	0.264	0.496
Flooding	0.347	0.338	0.417
Foot and mouth disease (FMD)	0.689	0.792	0.011*
Genetically modified organisms (GMO)	0.148	0.317	0.933
Loss of marine biodiversity	0.386	0.002*	0.953
Pesticides	0.153	0.162	0.002*
Loss of water quality	0.195	0.954	0.030*
Loss of wildlife biodiversity	0.007*	0.037*	0.447

cases; Figs. 6–8) compared to the expert-informed assessments; as demonstrated for the environmental consequence assessment (Fig. 6) for pesticides, loss of water quality, loss of wildlife biodiversity and derogations of air quality. This might be expected, as no consistent framing between independent research studies is expected in the open literature, unlike that for the expert assessments in this study that benefited from the context setting described; so comparison, moderation and consensus were all possible contributors to a tighter range of severity values (Fig. 6; air quality, Foot and mouth, pesticides).

We note a general closeness in median assessments between literature- and expert-informed assessments in the order environmental (most similar) > economic ≫ social. There are individual occasions when the differences are significant (Table 3), but the general observation suggests a deeper consensus between the sources of evidence and assessment for environmental consequences above economic and social consequences. Evaluating the economic and social consequences of harm is contentious in SRA workshops in the authors' experience and tends to drive polarised perspectives, not least because of differing views on the inclusion, qualitative nature and scale of externalities. Having asserted this, for the environmental and economic consequences, the literature-informed assessments are typically higher in their severity than those informed by expert-assessment; not so for the social consequences. Neither can be declared as accurate or not – they are just different - but this might suggest policy experts are perhaps more sanguine in their lesser assessments of consequence severity than the research community. We have to look hard to find comparator studies, but Nicollier-Fahrni et al. (2003) demonstrated in their experiments on the appropriateness of clinical indications for colonoscopy, that 68% of expert and literature-based evidence offered similar evaluations. In some cases, more evidence and more detailed narratives were used in the literature-informed assessment and there is no doubt that most expert-informed SRAs are completed under conditions of time pressure.

Why do literature-informed assessments appear to offer more severe assessments of consequence? Of the 8 significantly different comparisons (Table 3), only foot and mouth disease (social consequence assessment) and loss of water quality (social consequence assessment) has expert-informed assessment significantly higher in impact severity (Fig. 1); the remaining six were dominated by literature-informed assessment as determined by the Mann-Whitney *U* test. We suggest three plausible reasons: (i) 'sanguine experts' - an underestimation of impact severity by experts, consciously or subconsciously; (ii) 'evidence bias' – material differences in the lines of evidence utilised by each source; and (iii) 'assessor bias' – those assessing the evidence in the two types of assessment perceived the consequences both differently and systematically.

3.3. Implications for policy risk analysts and SRA workshops

Readers will recognise our comments on the contentious nature of SRA and the challenge of comparing seemingly incommensurate risks and their consequences across environmental, economic and social categories of harm. We are thereby cautious about our claims and note that SRA is best used to inform discussions about risk management strategies and relative priorities for onward investment in full view of the methodological limitations inherent to risk comparisons (Prpich et al., 2011). This said, SRA does provide a basis for comparing disparate risks and, in formalising expert judgement, it provides a traceability to decisions such as changes to investment in public risk management, the attention given to risk issues by 'state-of-the-environment reports' and strategic plans. However, SRA output cannot be used simply to prioritise issues 'at a glance'. Outputs are intended to allow technically aware users to analyse different issues, attach priorities, and communicate information to non-expert groups, having first understood the practical limitations. Furthermore, SRA is one among many strategic appraisal tools available to decision-makers (Hammond and Winnett, 2006; HM Treasury, 2011) that are necessarily imperfect, with the analyses they offer only ever as good as the evidence they draw on and the decisions they support - risk management should reflect sensible and meaningful conclusions rather than theoretical perspectives that run counter to sound judgement.

This study has its limitations. One might argue that no true comparison is valid unless it is completed by the same group of individuals, fully representative of the risks being studied, that perform the literature- and expert-informed assessments within a close time period of one another. Even then, these exercises would be susceptible to a host of inherent biases and so we are cautious about our claims. In this study, consistent application of the matrix in Appendix 1 was used as a modifier of some of this bias and triangulation was performed within the research team.

4. Conclusions

We address the hypothesis '*there is no significant difference between literature- and expert-informed assessments of the environmental risks that attract national attention in SRA*' and explored reasons for comparison and significant difference across environmental, economic and social categories of residual impact for 12 public risks. Accepting the limitations of such a study including the English policy perspective on risks that have global significance and risk character, we conclude with caution:

- Limited scholarly evidence is available on the extent of the environmental, social and economic consequences associated with the residual risks posed by national environmental threats.
- This said, for high quality research studies utilised by the research team, the literature- and expert-informed assessments showed good levels of agreement across 12 public risks, with only 8 of 36 assessment couples statistically distinct from one another as determined by a Mann-Whitney *U* test of their median consequence severities.
- The closeness in the median assessment of consequences, between literature- and expert-informed assessments, declines in the order environmental > (closer than) economic > social.
- The far extremes of consequence severity are not widely utilised for these residual risks, arguable reflecting the extent to which existing risk management measures are believed to be effective. This continues to pose a challenge for SRA practitioners.
- Expert-informed SRA appears a robust surrogate for a priori literature informed assessment, providing efforts are maintained to ensure full representation of the risks discussed in SRA workshops and that economic and social consequence categories can be informed by the deep expertise of those qualified to make these inputs.

Acknowledgements

We thank policy colleagues in the Department for Environment, Food and Rural Affairs (Defra) and their executive agencies for providing expert input. SAR, SJTP, JD and GP were funded through a Defra and a cross-Research Council (EPSRC, NERC, ESRC) research grant (EP/G022682/1) awarded to Cranfield University.\

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.scitotenv.2017.03.293.

References

Andrews, C.J., Hassenzahl, D.M., Johnson, B.B., 2004. Accommodating uncertainty in comparative risk. Risk Anal. 24 (5), 1323–1335.
Beddington, J., 2013. Addressing the challenges of the 21st century. J. Foundation for Science and Technol. 21 (1), 4–7.
Bowden, R.A., 2004. Building confidence in geological models. Geol. Soc. Lond., Spec. Publ. 239 (1), 157–173.
Cabinet Office, 2012. National Risk Register of Civil Emergencies. Cabinet Office, London, UK Available at:. https://www.gov.uk/government/publications/national-risk-register-for-civil-emergencies-2012-update.
Cox Jr., L.A., 2008. What's wrong with risk matrices? Risk Anal. 28 (2), 497–512.
Cramer, D., 2004. Fundamental Statistics for Social Research: Step-by-step Calculations and Computer Techniques Using SPSS for Windows. Routledge.
Dagonneau, J., 2013. Validating the Strategic Risk Appraisals of Policy Experts. (PhD thesis). Cranfield University, Cranfield, UK.
Defra, 2011. Guidelines for Environmental Risk Assessment and Management. Green leaves III. Revised Departmental Guidance. Cranfield University and Department for Environment, Food and Rural Affairs, PB13670, London, UK (79 pp.).
DeKay, M.L., Florig, H.K., Fischbeck, P.S., Morgan, M.G., Morgan, K.M., Fischhoff, B., Jenni, K.E., 2001. The use of public risk ranking in regulatory development. In: Fischbeck, P.S., Farrow, R.S. (Eds.), Improving Regulation: Cases in Environment, Health, and Safety. Washington, DC, Resources for the Future, pp. 208–230.
Deloitte, 2013. Exploring Strategic Risk. 300 Executives Around the World Say Their View of Strategic Risk is Changing (London, UK).
Department for Environment, Transport and the Regions, 1998. Management of Harm to the Environment: Criteria for the Management of Unplanned Releases. Department of the Environment, Transport and the Regions, London, UK (80 pp.).
Duarte-Davidson, R., SJT, Pollard, Yearsley, R., Llewellyn, G., Steele, J., 10–13 October 1999. Considering 'environmental harm': qualitative and semi-quantitative treatments for strategic risk assessment. In: Goossens, H J]–>L.H.J. (Ed.), Facing the New Millennium. Proceedings of the 9th Annual Conference on Risk Analysis (Rotterdam).
Environment Agency, 1999. Strategic Risk Assessment: Further Development and Trials, Research and Development Technical Report E70. Bristol, UK, Environment Agency.
Environment Agency, 2002. Strategic Risk Assessment: Phase II, Environment Agency R&D Project E2-041/TR, Prepared for the Environment Agency by Galson Sciences Limited. Environment Agency, Bristol, UK (150 pp.).
Environment Agency, 2004. Agricultural Strategic Risk Workshop Report. Prepared for the Environment Agency by Cranfield University Integrated Waste Management Centre and Galson's Sciences Limited (Unpublished).
Environment Agency, 2005. Informing strategic decisions with comparative risk assessment: a literature review. Science report RIS-05-01. Prepared for the Environment Agency by Cranfield University, Environment Agency, Bristol, UK (46 pp.).
Ernst and Young, 2010. The Ernst & Young business risk report 2010. The Top 10 Risks for Business. A Sector Wide View of the Risks Facing Business Across the Globe. Ernst & Young Global Limited, London, UK (45 pp.).
European Environment Agency, 2015. The European Environment. State and Outlook 2015. Executive Summary. European Environment Agency, Copenhagen, Denmark.
Evans, E., Ashely, R., Hall, J., Penning-Rowsell, E., Sayers, P., Thorne, C., Watkinson, A., 2004. Foresight Future Flooding. Vols. I and II. Office of Science and Technology, London.
Feldman, D.L., Hanahan, R.A., Perhac, R., 1999. Environmental priority-setting through comparative risk assessment. Environ. Manag. 23 (4), 483–493.
Fischhoff, B., Morgan, G., 2009. The science and practice of risk ranking. Horizons 10 (1), 40–47.
Florig, H.K., Morgan, M.G., Morgan, K.M., Jenni, K.E., Fischhoff, B., Fischbeck, P.S., DeKay, M., 2001. A deliberate method for ranking risks (I): overview and test bed development. Risk Anal. 21 (5), 913–921.
German Advisory Council on Global Change, 1998. World in transition - strategies for managing global environmental risks. Annual Report, German Advisory Council on Global Change (WGBU). Springer, Berlin, Germany.
Government Office for Science, 2011. Blackett review of high impact low probability risks. Government Office for Science, Report URN 12/519. Department for Business, Innovation and Skills, London, UK (45 pp.).
Haimes, Y.Y., Kaplan, S., Lambert, J.H., 2002. Risk filtering, ranking, and management framework using hierarchical holographic modeling. Risk Anal. 22 (2), 383–397.
Hammond, G.P., Winnett, A.B., 2006. Interdisciplinary perspectives on environmental appraisal and valuation techniques. Proc. Inst. Civ. Eng. Waste Resour. Manag. 159 (3), 117–130.
HM Treasury, 2011. The Magenta Book: Guidance for Evaluation. Her Majesty's Treasury, London, UK (accessed at http://www.hm-treasury.gov.uk/d/magenta_book_combined.pdf.
Hofstetter, P., Bare, J.C., Hammitt, J.K., Murphy, P.A., Rice, G.E., 2002. Tools for comparative analysis of alternatives: competing or complementary perspectives? Risk Anal. 22 (5), 833–851.
House of Commons. Science and Technology Committee, 2011. Scientific advice and evidence in emergencies. Third Report of Session 2010–2011, HC498. The Stationery Office, London, UK (88 pp.).
Howitt, D., Cramer, D., 2011. Introduction to SPSS Statistics in Psychology: For Version 19 and Earlier. fifth ed. Pearson Education Limited, Essex.
International Risk Governance Council, 2011. Concept note. Improving the management of emerging risks. Risks From New Technologies, System Interactions, and Unforeseen or Changing Circumstances. IRGC, Geneva, Switzerland (43 pp.).
Keeney, S., Hasson, F., McKenna, H.P., 2001. A critical review of the Delphi technique as a research methodology for nursing. Int. J. Nurs. Stud. 38 (2), 195–200.
Klinke, A., Renn, O., 2002. A new approach to risk evaluation and management: risk based, precaution based, and discourse based strategies. Risk Anal. 22 (6), 1071–1094.
Knol, A.B., Slottje, P., van der Sluijs, J.P., Lebret, E., 2010. The use of expert elicitation in environmental health impact assessment: a seven step procedure. Environ. Health 9 (1), 19.
Linkov, I., Ramadan, A.B., 2005. Comparative risk assessment and environmental decision making. Nato Science Series: IV: Earth and Environmental Sciences. NATO Science Series Vol. 38. Springer Link, Rotterdam.
Long, J., Fischhoff, B., 2000. Setting risk priorities: a formal model. Risk Anal. 20 (3), 339–351.
Mauelshagen, C., Denyer, D., Smith, M., Schiller, F., Rocks, S., Pollard, S.J.T., 2014. Effective risk governance for environmental policy making: a knowledge management perspective. Environ. Sci. Pol. 41, 23–32.
Morgan, M.G., Fischhoff, B., Lave, L., Fischbeck, P., 1996. A proposal for ranking risk within federal agencies. In: Davies, J.C. (Ed.), Comparing Environmental Risks: Tools for Setting Government Priorities. Resources for the Future, Washington, DC, pp. 111–147.
Morgan, K.M., DeKay, M.L., Fischbeck, P.S., Morgan, M.G., Fischhoff, B., Florig, H.K., 2001. A deliberative method for ranking risks (II): evaluation of validity and agreement among risk managers. Risk Anal. 21, 923–937.
Morgenstern, R.D., Shih, J.-S., Sessions, S.L., 2000. Comparative risk assessment: an international comparison of methodologies and results. J. Hazard. Mater. 78, 19–39.
New Jersey Department of Environmental Protection, 2003. Final Report of the New Jersey Comparative Risk Project.Available at:. http://www.state.nj.us/dep/dsr/njcrp/ (Trenton, NJ, Accessed on 20/12/2016).
Nicollier-Fahrni, A., Vader, J.P., Froehlich, F., Gonvers, J.J., Burnand, B., 2003. Development of appropriateness criteria for colonoscopy: comparison between a standardized expert panel and an evidence-based medicine approach. Int. J. Qual. Health Care 15 (1), 15–022.
Perhac, R.M., 1998. Comparative risk assessment: where does the public fit in? Sci. Technol. Hum. Values 23, 221–241.
Petticrew, M., Roberts, H., 2006. Systematic Reviews in the Social Sciences: A Practical Guide. Blackwell Publishing, Malden.
Pollard, S.J.T., Duarte-Davidson, R., Yearsley, R., Kemp, R., Crawford, M., 10–14 June 2001. Strategic risk assessment: a decision tool for complex decisions. In: Andersson, K. (Ed.), Proc. 2nd VALDOR Symposium Addressing Transparency in Risk Assessment and Decision-Making, Stockholm. Congrex, Stockholm, pp. 296–303.
Pollard, S.J.T., Kemp, R.V., Crawford, M., Duarte-Davidson, R., Irwin, J.G., Yearsley, R., 2004. Characterizing environmental harm: developments in an approach to strategic risk assessment and risk management. Risk Anal. 24, 1551–1560.
Prpich, G., Evans, J., Irving, P., Dagonneau, J., Hutchinson, J., Rocks, S., Black, E., Pollard, S.J.T., 2011. Character of environmental harms – overcoming implementation challenges with policy makers and regulators. Environ. Sci. Technol. 45, 9857–9865.
Prpich, G., Dagonneau, J., Rocks, S.A., Lickorish, F., Pollard, S.J.T., 2013. Scientific commentary: strategic analysis of environmental policy risks—heat maps, risk futures and the character of environmental harm. Sci. Total Environ. 463, 442–445.
Robson, C., 2002. Real World Research: A Resource for Social Scientists and Practitioner Researchers. second ed. Blackwell, Oxford, UK.
Rocks, S.A., Schubert, I., Soane, E., Black, E., Muckle, R., Petts, J., Prpich, G., Pollard, S.J.T., 2017. Engaging with comparative risk appraisals - public views on policy priorities for environmental risk governance. Risk Anal. http://dx.doi.org/10.1111/risa.12735 (in press).
Science Advisory Council, 2012. Paper SAC(12)04. Risk Issues. Agenda Paper for SAC Meeting 12th January, 2012. Defra, London, UK.
Soane, E., Schubert, I., Pollard, S., Rocks, S., Black, E., 2016. Confluence and contours: reflexive management of environmental risk. Risk Anal. 36, 1090–1107.
Tranfield, D., Denyer, D., Smart, P., 2003. Towards a methodology for developing evidence informed management knowledge by means of systematic review. Br. J. Manag. 14 (3), 207–222.
US Environmental Protection Agency, 1987. Unfinished Business: A Comparative Assessment of Environmental Problems. U.S. Environmental Protection Agency, Office of Policy Analysis, Office of Policy, Planning and Evaluation, Washington, DC, US.
van Asselt, M.B.A., 1999. Improving decision-making under uncertainty: an integrated approach to strategic risk analysis. In: Goosens, H J]–>L.H.J. (Ed.), Facing the New Millennium. Proc. 9th Annual Conference on Risk Analysis, Rotterdam, 10–13 October 1999. Delft University Press, Delft, Netherlands.
Vlek, C., 2013. How solid in the Dutch (and the British) national risk assessment? Overview and decision-theoretic evaluation. Risk Anal. 33, 948–971.
Willis, H.H., DeKay, M.L., Morgan, M.G., Florig, H.G., Fischbeck, P.S., 2004. Ecological risk ranking: development and evaluation of a method for improving public participation in environmental decision making. Risk Anal. 24, 363–378.

Willis, H.H., MacDonald Gibson, J., Shih, R.A., Geschwind, S., Olmstead, S., Hu, J., et al., 2010. Prioritizing environmental health risks in the UAE. Risk Anal. 30, 1842–1856.
Wohlin, C., May 13–14, 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering. Proc. EASE 2014 (London, UK).
World Economic Forum, 2011. Global risks 2011. An initiative of the risk response network. A World Economic Forum Report, sixth ed. (Switzerland, ISBN 92-95044-47-9).
World Economic Forum, 2017. The Global Risk Report 2017. 12th ed. (Geneva; Switzerland, ISBN 978-1-944835-07-1).