

Considering evidence: The approach taken by the Hazardous Substances Advisory Committee in the UK

Chris D. Collins, Soil Research Centre, University of Reading, RG6 6AW, UK

Michelle Baddeley, Bartlett School of Construction and Project Management, University College London, WC1E 7HB, UK

Gill Clare,

Richard Murphy, Centre for Environmental Strategy, University of Surrey, GU2 7XH, UK

Susan Owens, Centre for Environment and Policy, University of Cambridge, UK

Sophie Rocks, Institute for Environment, Health, Risk and Futures, Cranfield University, MK43 0AL, UK

1. Introduction

The Hazardous Substances Advisory Committee (HSAC) provides expert advice to UK officials, Ministers and other relevant bodies on the protection of the environment, and human health via the environment, from potentially hazardous substances and articles¹. The Committee's membership is multi-disciplinary and independent, enabling it to approach the evidence from a range of different perspectives. The Code of Practice requires members to observe the highest standards of impartiality, integrity and objectivity in relation to the advice they provide; the Code also includes clear provisions for handling conflicts of interest.

Hazardous substances are often the subject of controversy, on which individuals, and different groups in society, hold divergent views. In formulating its advice, the Committee needs to analyse, interpret and assess the available evidence, often in situations where the uncertainties may be considerable. This paper documents the different kinds of evidence that might be available to the Committee; the criteria that HSAC adopts in its assessments; and the wider perspectives and concerns that have a bearing on the issues at hand. It also proposes a process through which the Committee's judgements about the quality of the available evidence could be communicated in an accessible form.

2. Types of evidence

HSAC recognises that evidence varies in its source, robustness and defensibility, and that these factors will influence the degree of confidence that assessors can assign to any given 'piece' of evidence, or to a body of evidence as a whole. While most of the scientific evidence assessed by HSAC derives from experimental or epidemiological studies, or is based on modelling of some kind, observational and anecdotal evidence may also be considered (Table 1). Evidence in the last two categories sometimes provides a first indication that a phenomenon is worthy of further investigation, and can lead to more systematic studies. It is likely that the availability of less systematic evidence will increase with the evolution of social media. Statistical evidence is often grounded on hypotheses which have been tested to a certain degree. However, it takes time and resources to collect statistically robust data, so that such studies may not reflect rapidly changing circumstances and emerging problems.

¹ The HSAC Terms of Reference and its Code of practice can be downloaded from <https://www.gov.uk/government/groups/hazardous-substances-advisory-committee>

Table 1. Categories of evidence

		<u>Type of evidence</u>				
		Experimental	Model-based	Epidemiological	Observational	Anecdotal
		<i>Obtained through a methodological approach to experimental design and data collection.</i>	<i>Computer modelling of effects or exposures to provide a measurement of impact.</i>	<i>Data based on studies of populations under real-world conditions.</i>	<i>Based on observations and experience.</i>	<i>Based on personal accounts of effects.</i>
		<u>Possible to show causality/association</u>	<u>Infers causality</u>	<u>Infers association</u>	<u>Infers association</u>	<u>Hypothetical association - potentially identifies issues of concern, not yet addressed in scientific research</u>
↑ ----- Increasing quality ----- ↓		Repeated experiments with a high degree of replication and controls following internationally accepted standards (e.g. OECD Test Guidance Documents)	Approach informed by empirical evidence, all processes and parameters revealed to allow repetition by others	Follows published guidance (e.g. WHO) with clear methods and rationale for data inclusion or exclusion	Field observations made in a systematic way, but without a specific experimental design	Relatively high incidence of specific effects; consistency between unconnected accounts; different accounts carefully collated.
		Not meeting widely-accepted experimental protocols; untested method, poorly reported	Model without antecedents, parameters from assumptions not measurements, processes a black box, i.e. cannot be repeated by others	Un-tested method, inadequately reported, using non-standard measurements of impact	Circumstantial evidence random or 'one off' events or phenomena	Uncorroborated, unconfirmed anecdotes: 'a friend of a friend...'

There can be significant variations of quality within each type of evidence. Examples (not exhaustive) are given of what might be considered 'high' or 'low' quality evidence within each column; in practice, there will be a gradation. No simple (horizontal) quality continuum between different types of evidence is implied; see sections 3 and 4 below.

3. Judging quality: considerations to take into account

In reviewing the scientific evidence, HSAC considers the extent to which any given study meets the following, widely-accepted criteria. HSAC may attach particular weight to evidence that conforms to these criteria, though 'weaker' evidence (in these terms) should not be dismissed: it can be part of the bigger picture when different sources of evidence are combined.

- **Transparency of aims.** A study should have a clearly stated purpose, in terms of the problem to which it relates and the research questions to be addressed. Conventionally, this is achieved through the

statement of a hypothesis. The hypothesis to be tested should preferably link to previous work, and the study should be clear about the ways in which it builds upon, or challenges, the evidence base. The nature of HSAC's work is such that the Committee is often focusing on substances that have not been subject to exhaustive scientific studies (nanomaterials would be one example). In this case the hypothesis may be that a suspected causal agent is responsible for harm and it is important to recognize that this is essentially an arbitrary formulation. In assessing the stated hypothesis, it has to be clearly structured so that it is properly testable and falsifiable. HSAC recognises that findings based on statistical evidence are conditional on the structure of the hypotheses, and also on a potentially arbitrary decision about significance levels (e.g. a 10% or 5% probability of Type 1 error – i.e. incorrectly rejecting a true null hypothesis) and confidence intervals selected by the researcher (e.g. a 90 or 95% probability of the true value lying within the interval).

- **Methodology and results.** For experimental evidence, the Bradford Hill features or characteristics of causal associations [1] provide an excellent starting point for investigating causality. These include: *temporality, strength of the association, consistency of the observations, biological plausibility of the effect and evidence for recovery following diminution of the agent suspected of causing stress.*

Within each experimental study HSAC would also have regard to the following:

Methodology. For given data sources the methods used should have a sound scientific basis and should be fully described, capable of repetition and appropriate to the aims of the study. The reproducibility of the method should be tested by statistical examination of the replicates where the variability should ideally be low. The risks of bias in data collection should have been considered in the study design, and the efforts made to minimise any recognisable bias should be declared. There should be evidence of sound laboratory procedure, such as the use of controls and analytical blanks.

Results and interpretation. Results should be presented in a transparent way and should have appropriate statistical validity and power (for example, the data set should be of a suitable size, and appropriate confidence intervals and significance levels should be used). The caveat to any study is '*under the experimental conditions described*', so it is important that the conditions are relevant to the problem under investigation. The interpretation of the data should consider potential sources of error in the study, and the extent to which these affect the degree of uncertainty assigned to the findings and conclusions. The *null and alternative hypotheses* should be carefully constructed so that the study gives robust findings, allowing researchers to be confident about their result.

- **Completeness.** A study should be sufficiently complete to enable third parties to review it and arrive at an independent interpretation of its findings, which may or may not coincide with that of the original authors. As already noted, the authors should themselves attempt to identify uncertainties and

weaknesses in a given study, though it may not be easy (or even possible) to be comprehensive in this respect (see Sections 4 and 5).

- **Independent review.** The source of a study, and the likelihood of bias, are important considerations when assessing the quality of evidence (see also Section 5). Peer review (that is, review of a study by those regarded as having expertise in the field) is critical in this respect, even if it is an imperfect process. Greater confidence is also gained as other independent scientists replicate the original findings.
- **Accessibility.** Studies should be published or available in archival form in the public domain, so that the evidence can be readily examined. Even if they are not freely available, the costs of access to cited studies should not be prohibitive. Ideally, the raw data on which the study is based should also be available in a comprehensible form, so that its use can be assessed by others.

A number of formalised approaches, including, for example, the ‘Klimisch criteria’ and Harris *et al.* [2, 3], suggest further attributes of what might be considered ‘high quality evidence’, and these can be incorporated into a ‘weight of evidence’ approach [4] and systematic review. HSAC’s remit means that it may be called upon to assess and report on a wide range of areas and potential hazards, and the specific approach used needs to be selected on a case-by-case basis. HSAC seeks to ensure that its recommendations are fully and transparently described.

It is important to be aware of the limitations of criteria such as those described above. Even studies that ‘tick all the boxes’ may, for example, be the subject of unconscious bias or ‘groupthink’; research questions and design may be influenced by experimenters’ prior beliefs, or driven by particular fashions or pressures to publish. Researchers may form their conclusions relatively quickly, using heuristics or ‘rules of thumb’ (quick decision-making devices that can be useful but can also lead to biased assessments of evidence) and when these approaches are applied they may or may not have a reasonable empirical basis [5]. Further, it is in the nature of certain forms of bias that they seem normal and unbiased to those who hold them. HSAC considers it important, therefore, always to reflect on research questions and assumptions when considering scientific evidence, and to ask, for example, ‘how has the study been framed?’, and ‘what might have been missed?’ It is just as important for HSAC itself to reflect on assumptions and possible biases in its own evaluation of the evidence - a process that is facilitated by the diversity of the Committee.

4. The wider context

Criteria for assessing the quality of evidence can be surprisingly difficult to distinguish from those that individuals and groups deploy, consciously or not, in judging how much credence to attach to evidence in a particular case. The difficulty lies in identifying a benchmark from which to determine whether evidence is ‘sound’, in some wholly impartial, objective sense. In assessing the available evidence, and offering an opinion, HSAC is conscious that the science alone, while of fundamental importance, is unlikely to settle

issues of deep controversy. Rather, those presented with new evidence (scientists and advisors, as well as pressure groups, publics and decision makers) tend to be influenced by:

- **Trust in the source of the evidence.** Important considerations are whether the evidence comes from an individual or institution seen to have authority from the recipient's perspective and whether the source has a known, explicit or inferred bias. Obvious interests in the issue at hand can reduce the trustworthiness of evidence ('they would say that wouldn't they?'). Conversely, evidence that takes an unpredictable or unexpected line can sometimes be persuasive.
- **Defensibility.** Recipients of evidence form views about whether it has been arrived at in a defensible way, and there is clearly overlap here with the standard, scientific criteria for 'good' evidence (Section 3). The wider point is that in matters of controversy, the quality of the evidence is itself likely to be a matter for dispute, because this becomes part of the process of questioning unwelcome findings. Assumptions, judgements and biases (which are always, and necessarily, present, sometimes hidden behind claims about 'objectivity') are likely to be exposed and questioned.
- **Conformity to the recipient's 'worldview'.** Worldviews may include beliefs about nature (for example, whether natural systems are fragile or robust) and positive/negative feelings about particular 'risky' activities; they may be shared within groups, communities and cultures [6, 7]. Even for the most 'objective' of recipients, evidence is likely to be filtered through a worldview.
- **Framing.** Evidence may be more persuasive if it relates to a meaningful framing of the problem from the recipient's perspective [8]. Individuals may be unimpressed by evidence suggesting that substance X is 'safe' if the risk in question is not what really bothers them about that substance (for example, if reassurances are based on potential harms to human health, but concerns are with possible effects in the wider environment). Alternatively, they are likely to seize upon evidence exonerating X, if X is a substance whose use they want to promote. Even high quality evidence, according to criteria such as those in Section 3, will make little difference if it relates to an issue that is not, in fact, the primary issue of concern.

HSAC needs to be aware of these wider considerations, in addition to 'purely' scientific matters, when reviewing the available evidence and will aim to reflect them when presenting an opinion (Figure 1. and S.I.). HSAC's view is that awareness of context and different perspectives will enhance both the quality of its own deliberations and the utility of its opinions for decision-makers.

5. Towards a transparent assessment

HSAC addresses different kinds of questions, for which the evidence varies in terms of type, quality and amount. The Committee needs, therefore, to be flexible in its specific approach when weighing the evidence and arriving at an opinion (examples of published HSAC opinions are provided in S.I.). Criteria of the kind outlined in Section 3 can be applied as appropriate to the scientific evidence but the body of evidence *as a whole* needs also to be considered, and HSAC will take account of important, wider questions such as those of problem framing. As noted in Section 3, HSAC also reflects on its own perspectives (assisted by the

diversity of its membership) and on the wider context within which the problem has been framed. The Committee's intention, in adopting this rounded approach (Figure 1.), is to reach opinions that are robust, relevant and defensible. HSAC also considers how it might best reflect the overall strength of any given assessment and communicate the degree of confidence in its opinion. Quantitative measures, while attractive in some senses, can be open to misinterpretation, but a number of useful systems exist for indicating levels of confidence in a qualitative way [9-11]. HSAC will adopt (and adapt) one or more of these systems, as appropriate to the case in hand, when presenting its conclusions.

References

- [1] A.B. Hill, The Environment and Disease: Association or Causation?, *Proceedings of the Royal Society of Medicine* 58(1965) 295–300.
- [2] C.A. Harris, A.P. Scott, A.C. Johnson, G.H. Panter, D. Sheahan, M. Roberts, J.P. Sumpter, *Principles of Sound Ecotoxicology*, *Environ. Sci. Technol.* 48(2014) 3100-3111.
- [3] H.J. Klimisch, M. Andreae, U. Tillmann, A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data, *Regul. Toxicol. Pharmacol.* 25(1997) 1-5.
- [4] D.L. Weed, Weight of evidence: A review of concept and methods, *Risk Anal.* 25(2005) 1545-1557.
- [5] Baddeley, M. Herding, social influences and behavioural bias in scientific research. *EMBO reports.* 16(2015):902-905.
- [6] K. Dake, Myths of nature: culture and the social constructions of risk, *Journal of Social Issues* 48 (1992) 21-37;
- [7] M. Douglas and A. Wildavsky, *Risk and Culture: An Essay on the Selection of Technical and Environmental Dangers*, Univ. of California Press, Berkeley, 1982.
- [8] D. A. Schön and M. Rein, *Frame Reflection: Toward the Resolution of Intractable Policy Controversies*, Basic Books, New York, 1994.
- [9] Annual Report of the Government Chief Scientific Adviser. *Innovation Managing Risk and Not Avoiding It (Evidence and Case Studies)*, (2014).
- [10] T.R. Society, *Resilience to Extreme Weather*, Royal Society, London, 2014.
- [11] C.B.F. Michael D. Mastrandrea, Thomas F. Stocker,, K.L.E. Ottmar Edenhofer, David J. Frame, Hermann Held, Elmar Kriegler,, P.R.M. Katharine J. Mach, Gian-Kasper Plattner, Gary W. Yohe,, a.F.W. Zwiers, *Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties* <http://www.ipcc-wg2.gov/meetings/CGCs/index.html>, 2010.

Figure 1. HSAC workflow

