Policies and Procedures

JNCC Evidence Quality Assurance (EQA) Policy

Appendix 1. Bias, conflicting evidence and uncertainty

This appendix is an edited version of Evidence Quality Guidance Note 1 (EQGN1), written in 2013-14 by Helen Baker and edited by Richard Ferris and Matt Smith

https://jncc.gov.uk/about-jncc/corporate-information/evidence-quality-assurance/

To find out more about JNCC visit https://jncc.gov.uk/about-jncc/

Appendix 1. Bias, conflicting evidence and uncertainty [1]

1. Bias in interpretation of evidence

1.1. The interpretation of evidence can be biased by a number of factors, including for example:

- lack of evidence and/or poor transferability of evidence;
- selective choice of evidence to underpin conclusions and advice;
- dismissal of evidence that conflicts with other evidence;
- failure to account for the quality of evidence included and its associated uncertainty;
- poor choice of additional analytical (meta-analysis) methods;
- poor data management underpinning meta-analyses;
- poor data analysis (i.e. incorrect use of statistics)
- combining evidence and expert opinion without a suitable audit trail demonstrating how differing types of evidence have been combined;
- poorly designed methods for obtaining expert opinion.

The JNCC Evidence Quality Policy is designed to help reduce bias in the interpretation of evidence so that advice and response options are as robust as available good quality evidence allows. The following guidelines should help staff in meeting the principles set out in the policy.

1.2 Searching for, collating and reviewing third party evidence

Third party evidence includes all information that is gathered from sources outside JNCC.

Searching and sourcing:
JNCC staff have access to Scopus, which enables searches of published research literature using specific search terms. In addition, some internet search engines are specifically designed to find scientific literature, including 'grey' literature (reports) (e.g. Science Direct, Google Scholar). Scopus searches can be filtered to show only open access results. Most of the literature can be obtained through OpenAthens logins supplied by the Defra librarians, who also provide a service supplying electronic interlibrary loans for the standard fee using project budgets. Papers can also be obtained by email request to the main author and/or through Researchgate. Although there is increasing open access availability of important literature and more funders are demanding open access publication of work they fund, relying on open access journals and grey literature may result in bias and JNCC staff should remain aware of this when searching for evidence.

JNCC uses the free reference managing software Zotero, which should be available at team level to assist with managing information on relevant literature and generating reference lists. Decisions on which papers to acquire and review take much more time and this time needs to be estimated when planning a project and setting deadlines.

Conservation Evidence (https://www.conservationevidence.com/) provides a fully referenced summary of the evidence concerning a wide range of conservation actions based on a database of over 5,000 papers.

Collation and review:

[1] This appendix is an edited version of EQGN 1, written in 2013-14 by Helen Baker and edited by Richard Ferris and Matt Smith

It is important to ascertain what constitutes reasonable effort in collating and reviewing relevant evidence. There is a trade-off between reducing bias in our assessments of evidence and the use (or impact) of the evidence product. Therefore, the effort spent reducing bias needs to be proportionate to the expected use (or impact) of the evidence product. A basic risk assessment can help in making a decision on how much effort to invest in searching, collating and reviewing published evidence (see the Evidence Quality Policy section 5.2). All categories of knowledge synthesis methods have been analysed by the EU-funded Eklipse project (http://www.eklipse-mechanism.eu/expert_group_on_methods), with their costs and benefits, strengths and weaknesses, to help select the most appropriate method for each task. The 21 methods analysed include:

Systematic reviews
Meta-analysis
Multiple expert consultation with Delphi
Non-systematic literature review
Bayesian Belief Networks

http://www.eklipse-mechanism.eu/apps/Eklipse_data/website/EKLIPSE_D3-1-Report_FINAL_WithCovers_V6.pdf [2]

1.3 What is good or high quality evidence?

Adopting good scientific practice in evaluating evidence will help in judging the quality of evidence. Key questions to consider include:

- Are the data gathering methods fit for purpose and scientifically sound?
- Given the type and amount of data, are the analytical methods appropriate?
- Do the conclusions fit the results derived from the analytical methods?
- Is the work peer reviewed?

Not all good or high quality evidence is published in scientific journals and it should not be automatically rejected from consideration simply because it has not been published. If the authors have undertaken some form of peer review of the published material this might increase confidence in the quality of the evidence. In addition, other parties may afford a greater confidence to JNCC's evidence if published material has been subject to a peer review process. Conversely, peer review does not necessarily guarantee quality of information and JNCC staff should remain open-minded and inquiring about the evidence being drawn upon.

Valuable sources of evidence exist that will not have gone through any formal peer review process (e.g. industry activity data, site reports from SNCBs, EIA casework, data from trade associations etc.) and in some cases it may not be clear whether peer review has been undertaken. Review and inclusion of these kinds of evidence will require some assessment of quality, which will take more time.

1.4 Judging relevance of third party evidence

[2] Dicks, L.V., Haddaway, N., Hernández-Morcillo, M., Mattsson, B., Randall, N., Failler, P., Ferretti, J., Livoreil, B., Saarikoski, H., Santamaria, L., Rodela, R., Velizarova, E., and Wittmer, H. (2017). Knowledge synthesis for environmental decisions: an evaluation of existing methods, and guidance for their selection, use and development – a report from the EKLIPSE project.

When assessing the relevance of evidence that has been collected by individuals or organisations outside JNCC the key points that should be considered are:

- Are your work or project objectives clear and correct? In some cases, it might be decided that objectives require peer review to ensure that they will deliver evidence that is fit for purpose.
- Do the hypotheses or objectives of third-party reports match some or all of the project's objectives? Evidence that is only marginally relevant would normally be excluded and records should be kept of the decision to exclude certain evidence.
- When was the evidence gathered? Evidence gathered some time ago might be less valuable than recently collected evidence due to the dynamic nature of natural systems, the policy landscape and other variable factors that may affect validity of historical evidence. However, historic data should not be automatically discounted and staff will need to make a judgement as to whether the evidence is still reliable and relevant.

It is important not to reject evidence solely because it conflicts with other evidence; this is not a valid way to judge relevance.

1.5 Using third-party evidence for assessments

Scoring and ranking quality and relevance can help determine how to use the evidence in an assessment or review. For example, lower quality evidence might be given less weight in an assessment if there is better quality very similar evidence available (see Table 1 for example rankings, on a scale of 1-5, with 5 being highest ranked). An attempt should be made to collate all relevant evidence of good or high quality to include in an assessment. Evidence of more marginal relevance that is included in an assessment might be weighted as less important for the findings than evidence more closely matching the project's objectives.

Table 1: An example of how to rank or score third party evidence

Relevance →	Quality →		
	High relevance Low quality Rank 2	High relevance Medium quality Rank 4	High relevance High quality Rank 5
	Medium relevance Low quality Rank 2	Medium relevance Medium quality Rank 3	Medium relevance High quality Rank 4
	Low relevance Low quality Rank 1	Low relevance Medium quality Rank 2	Low relevance High quality Rank 2

A scored/ranked approach to using evidence will be important for assessment of certainty in the overall findings from a review (also see tables in Section 3 'Assessing Certainty'). It will be up to project managers to decide when, and if, evidence should be discounted based upon ranking or scoring. Under some circumstances it may be the case that all available evidence is of low quality; this limitation must be clearly communicated in the final product or advice provided.

1.6 Meta-analysis of third-party evidence

Data sourced from other studies can be re-analysed statistically, including when combined with data generated by JNCC. This type of analysis should follow good scientific practice, including an assessment of the quality of the external data prior to use, and the resulting JNCC outputs.

Whichever methods of meta-analysis are chosen, it is useful to test that they are appropriate for your objectives through peer review, and have the findings peer reviewed (see Appendix 2).

1.7 Documenting the search and selection process

When undertaking the search and selection process it is recommended that the following steps are followed to enable an audit of the evidence selection process to be undertaken if necessary:

- Produce a simple plan for searching, collation and review of third-party evidence;
- Record the methods that have been utilised during this process;
- Keep a list of literature returned from specified search terms or the combination of terms and their different permutations;
- Record the risk assessment undertaken to judge the effort needed in collating and reviewing evidence;
- Assess the quality and relevance of collated evidence and record reasons for rejecting specified evidence, including material meeting quality and relevance judgements;
- If using a scoring method for assessing quality and relevance of evidence then a record should be kept of the method applied and outcomes;
- Record any peer review methods and outcomes for determining meta-analysis approaches and testing outcomes of analyses undertaken;
- Correctly cite all evidence used in an assessment; staff should refer to the JNCC Design Identity Manual for guidance on correct citation style.

1.8 Expert opinion and judgement

See Appendix 3.

2. Dealing with conflicting evidence

Systematic reviews of evidence typically demonstrate that divergent conclusions emerge from different studies of similar ecological processes or the effects of the same, or similar, interventions. There is a risk that in undertaking selective reviews of evidence conflicting evidence will be missed from an evaluation and not included in final conclusions (findings). This may result in assessments of certainty and response options being erroneous.

As best practice, the assessment of relevance and quality should be applied to all evidence, irrespective of whether there are conflicting findings. In practice this can be difficult, as literature searches may not include grey literature and fail to pick up on evidence that is deemed ‘un-publishable’ for various reasons.

There may be occasions when JNCC staff are faced with a situation where evidence sources give conflicting conclusions and will have to judge which evidence is the most reliable. If there is a case where 2 or 3 high quality evidence sources conflict with numerous

low quality evidence sources it will be imperative that staff judge the reliability of the sources and it would be likely that the high quality evidence would take precedence over low quality evidence.

The weighting of evidence, if used, should be applied consistently to ensure that the outcome of an assessment of findings is repeatable and that the certainty of the overall finding can be qualified. To enable this, JNCC staff are required to document the decision-making processes that have been used to select what is deemed to be the highest quality evidence.

3. Assessing certainty

Uncertainty can arise from lack of evidence or disagreement about what evidence conveys. Evidence types can vary and may be measured (quantifiable) or descriptive (qualitative).

Assigning certainty terms to findings from the review or assessment of multiple sources of evidence can be done using the UK National Ecosystem Assessment (UK NEA) 4-box model and likelihood scale (Appendix 3.1, page 61, of the Technical Report), which is modelled on the Intergovernmental Panel on Climate Change (IPCC) approach.

The IPCC approach uses two ways of communicating certainty in findings:

- **Confidence in the validity of a finding** based on type, amount, quality and consistency of evidence and the degree of agreement. Confidence is expressed qualitatively;
- **Quantified measures of uncertainty** in a finding expressed probabilistically (as likelihood).

Findings or conclusions can be assigned a standard term using criteria:

Confidence	Standard term	Criteria
High	Well established	High agreement between evidence and plenty of good to high quality relevant evidence available
Medium	Established but incomplete evidence	High agreement but limited evidence
Low	Competing explanations	Low agreement albeit with plenty of evidence
Very low	Speculative	Low agreement based on limited evidence

Or assigned with a level of likelihood using the scale:

Likelihood terminology	Probability of occurrence
Virtually certain	>99%
Very likely	>90%
Likely	>66%
About as likely as not	33-66%
Unlikely	<33%
Very unlikely	<10%
Exceptionally unlikely	<1%

Examples of how this is used can be found in the UK NEA:

> ***"Agri-environment schemes are critical to maintain and enhance the biodiversity of ecosystem service of semi-natural grassland****. Maintenance of the biodiversity and cultural value of semi-natural grassland requires low intensity management related to traditional farming (**well established**) ... protected and restored semi-natural grasslands also have potential to provide recreational and tourism services, and pollinator and pest control services for adjacent intensive farmland (**likely**)."* (Key findings, Chapter 6 of Technical Report, page 163):

In this example, the statement about low intensity management is derived from a mixture of quantitative and qualitative evidence; hence a qualitative assessment of confidence has been made. The statement on pollinator and pest control services is based on quantitative evidence only.

> *"**Marine microbial organisms play a key role in cycling nutrients that are essential for other marine organisms and the services and benefits they provide** (**well established** and **virtually certain**). Microbial processing of nutrients in the sediment depends on invertebrates disturbing and irrigating the sediment (**established but incomplete evidence**). Without this recycling, most nutrients would be lost from the ecosystem to the seabed as they would sink from the water column and then be buried (**virtually certain**). In open water, planktonic coccolithophores make a major contribution to the global carbon sink (**virtually certain**). Climate change may affect internal nutrient cycling by changing nutrient exchange processes between the open waters and the open ocean and altering water stratification, but the likely direction and extent of these changes is still poorly understood (**likely**)."* (Key finding, Chapter 12 of Technical Report, page 461)

In this example, the statement that microbial organisms play a key role in nutrient cycling is derived from well established, peer reviewed, quantitative evidence. Therefore, a quantitative assessment of confidence has been applied. The statement on the effect of climate change on internal nutrient cycling was derived from interdisciplinary reviews of evidence that was emerging in 2006/07. Due to the limited amount of established evidence a lower confidence has been applied.

It is important to note that there is a marked difference between confidence assigned from statistical analysis and certainty assessments based on the conclusions from multiple sources of evidence. This should be kept in mind and clearly communicated in the final evidence outputs.