Baltic and North Sea Coordination and Support Action

Work package 4

Deliverable: D4.3 Developing mechanisms for research synthesis and scalable knowledge transfer

Lead organization: Swedish Agency for Marine and Water Management (SWAM), Task Leader: Floor ten Hoopen

Submission date: 28 October 2021

The main outcomes

- Evidence (or research) synthesis forms a critical step in translating research findings for use in policy.
- Reviews and maps of the literature summarise and describe the body of evidence (data and primary studies) on a topic, providing information in a digestible and transparent way that facilitates its use in policy and practice decision making.
- Evidence synthesis has to date not been the topic of many calls or programmes, but there have been several examples (including the BONUS synthesis call) from which valuable lessons can be learned.
- This report discusses different synthesis methods and call structures that could be relevant in a future programme; including considerations and guidelines for choosing and using the different methods for future projects and thoughts on topics relevant for future calls.
- Stakeholder engagement is critical for synthesis to be successful. Ideally, stakeholders should be actively involved in the development of the call, and also in successful projects. The report discusses different aspects of stakeholder engagement

To be cited as Haddaway, N. R, and ten Hoopen, F. (2021). Developing mechanisms for research synthesis and scalable knowledge transfer. BANOS CSA/D4.3.

Authors: Neal R. Haddaway[1,2], Floor ten Hoopen[3]

Affiliations: [1]Stockholm Environment Institute, Linnégatan 87D, Stockholm, Sweden

[2]Africa Centre for Evidence, University of Johannesburg, Johannesburg, South Africa

[3]Swedish Agency for Marine and Water Management (SwAM), Sweden

Contact regarding administrative questions: Floor.hoopen@havochvatten.se

Contact regarding technical questions: Neal.haddaway@sei.org

This report was produced on behalf of the Swedish Agency of Marine and Water Management within the framework of the Baltic and North Sea Coordination and Support Action (BANOS CSA) project funded from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 817574.

Task 4.2 Developing mechanisms for research synthesis and scalable knowledge transfer. Task leader: SWAM

Effective scientific synthesis, dissemination and knowledge transfer are essential elements to crucial steps towards transferring research results to societal benefits, including improved environmental policies and management, technical innovation and sustainable development. This task consists of two activities: (i) The first activity will develop an improved process to synthesise primary scientific information. The synthesis process will use transparent and evidence-based methodologies for scoping, collecting, assessing and synthesising research in a given field. The research synthesis process will take into account scientific findings from existing BONUS-projects as well as from a wider survey of the respective scientific field. (ii) The second activity will create a process for improved dissemination and knowledge transfer of scientific results. It will develop methods/mechanisms to effectively communicate the outputs from syntheses to a wide range of stakeholders. This will include identification of target groups, further tailoring of syntheses to target audiences and delivery of syntheses' results using appropriate strategies and channels. Thus, the two activities in this task are complementary and will proceed in parallel. The processes and methods developed in activities (i) and (ii) will be documented and made openly available to users, including research funders and applicants, in the form of education material.

The task will build on previous experiences of e.g. centres for evidence-based environmental management, outcomes from (ongoing) BONUS synthesis projects and other selected projects on successful knowledge transfer. As a starting point, the development of the synthesis process will use four projects under Societal Challenge 2: Food Security initiatives on Knowledge Transfer and Societal Impact, directly relevant to BONUS Mission: H2020 COLUMBUS4, BioLINX5, CommBeBiz6 and ProBIO7.

The work on developing mechanisms for research synthesis will commence as soon as the scope of the future programme is delineated (M8). During the initial key areas for scientific synthesis will be identified and delimited, and relevant data systematically collected, analysed and compiled. In M18 this task will contribute to the SOW. During the second period, the focus will be on means of dissemination and knowledge transfer to stakeholders, policy makers at various levels outside the scientific realm. By this we ensure that pathways will exist for a more efficient and stringent of knowledge transfer that can be used for the future, thus shortening the gap between science and policy. The final output of this this task will be produced by completion of the project (M30).

Contents

Executive summary

The Baltic and North Sea Coordination and Support Action (BANOS CSA) is preparing a framework for the joint Baltic Sea and North Sea Research and Innovation Programme (BANOS). Systematic synthesis of research results has a potential for enlarging the societal impact of BANOS, but its implementation in research programmes has yet only been realised partially and in a fragmented way. The main objective of the task presented in this report (D4.3) is to develop mechanisms for research synthesis and scalable knowledge transfer that are applicable for BANOS and to provide guidance towards their implementation.

In this report, different synthesis methods are discussed that could be relevant in a future programme; considerations and guidelines are provided for choosing and using the different methods for future projects, and thoughts are included on topics relevant for future calls, and what considerations need to be made for implementing them successfully. The report includes an analysis of the results from the synthesis call by the predecessor programme BONUS.

Introduction

In this section, some key definitions and background are introduced, before moving on to explain the potential of evidence synthesis for evidence-informed decision making. It also illustrates why there is a need to discuss knowledge translation specifically for evidence synthesis projects.

Definitions

Although the original title of this deliverable as presented in the BANOS CSA description of work was "Developing mechanisms for *research synthesis* and scalable knowledge transfer", '*evidence synthesis'*, is the most widely used synonym for synthesis of primary research studies – the process explicitly referred to in the BANOS CSA description of work (European Union's Horizon 2020 Research and Innovation Programme, grant agreement No. 817574; see Figure 1). In addition, the majority of guidance and support for synthesis of primary research refers to 'evidence synthesis', as does literature and guidance on 'evidence-informed decision making'. Evidence synthesis is therefore the terminology used in this report.

Another term covering the synthesis of information from a wide variety of knowledge systems is the term '*knowledge synthesis*'. This term covers not just primary research but also includes expert consultation, indigenous knowledge, local knowledge and focus groups (Dicks et al, 2017). Since knowledge synthesis is not mentioned in the task description and is not a widely used term (see Figure 1), this report does not consider such a broad definition. We do however provide an extensive summary of methods covering also evidence synthesis for overview in Annex A.

Evidence synthesis is defined as the process of identifying and summarising a body of research literature (CEE 2018), typically primary research (although synthesis of secondary research/reviews may also be included). Evidence synthesis methods include traditional literature reviews, scoping reviews, rapid reviews, systematic maps and systematic reviews. For a detailed presentation of each of these methods, see Annex A.

To include "lessons learned" on language used in previous synthesis activities, call material was analysed and interviews were undertaken with project coordinators of synthesis projects funded under the predecessor programme BONUS, which published a call for synthesis projects (BONUS call 2017: Synthesis). The BONUS Synthesis call text explicitly called for systematic reviews 'wherever possible and appropriate', and mentioned both 'knowledge synthesis' and 'critical review of research outputs'. The aim of the call has been to fund synthesis in a very loose definition. However, it has been interpreted by end-users and funders as well as most applicants to mean that evidence syntheses were a compulsory part of the projects. The term 'systematic review' in the call text was interpreted very differently by project teams, as demonstrated by interviews and an evaluation of the review methods used (see Annex B Table 4).

The variability regarding interpretation of the terminology in the BONUS synthesis call process could indicate that there may also be a substantial diversity of understanding amongst the coordinators and applicants of a future BANOS programme. The interpretation of the term 'synthesis' ranged from the rigorous assembly and appraisal of bodies of empirical research evidence, to the process of learning from the results of an organisations completed projects (although this is typically referred to as 'evaluation'). These differences in understanding could originate from a diversity in specificity and reference to internationally accepted methodological guidance in the call text (see Annex B), which again may be related to the multiple definitions of the term 'synthesis' in broader environmental research communities where evidence synthesis is not yet a well-established practice. An analysis of project outputs has been included, investigating how synthesis methodology was conducted. This evaluation indicates how future synthesis projects could improve the rigour of their outputs in any future synthesis call. This deliverable thus seeks to clarify and provide guidance tailored to the needs of future applicants.

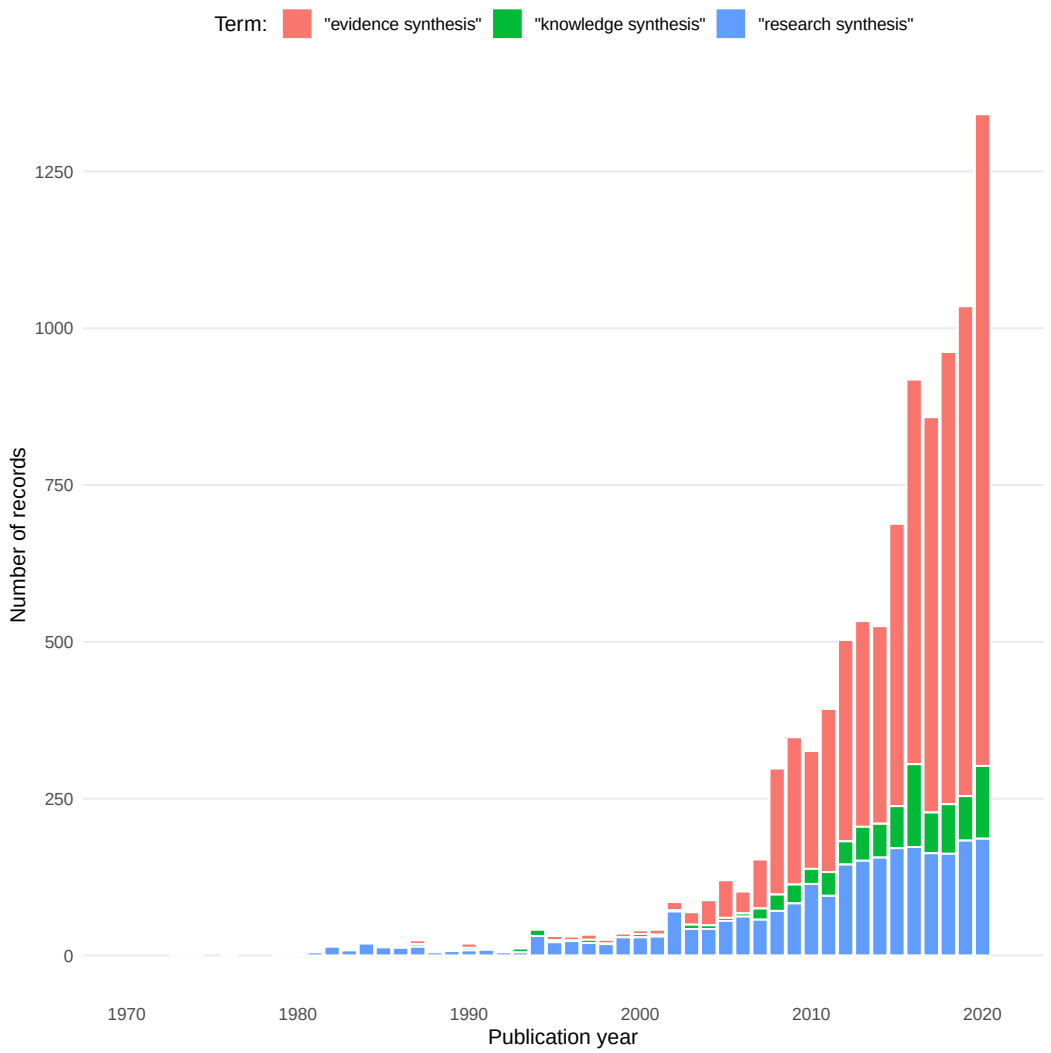

Figure 1. Relative frequencies of the synonymous terms 'evidence synthesis', 'knowledge synthesis', and 'research synthesis' in published research showing changes in popularity of the terms in the academic literature. Based on a Lens.org search on 07/07/2021 of title, abstract, keyword and field of study fields for the following terms: "research synthesis", "knowledge synthesis", and "evidence synthesis".

The role of evidence synthesis in decision making

Evidence synthesis forms a critical step in the process of translating research findings for use in policy (White 2019; Figure 2). Reviews and maps of the literature summarise and describe the body of evidence (data and primary studies) on a topic, providing information in a digestible and transparent way that facilitates its use in policy and practice decision making through summary in platforms (allowing easy location of relevant reviews), portals (allowing for identification of potential interventions and their underlying evidence), and then true decision support tools to suggest recommended actions in a given context (referred to as 'guidelines', and 'checklists' in Figure 2).

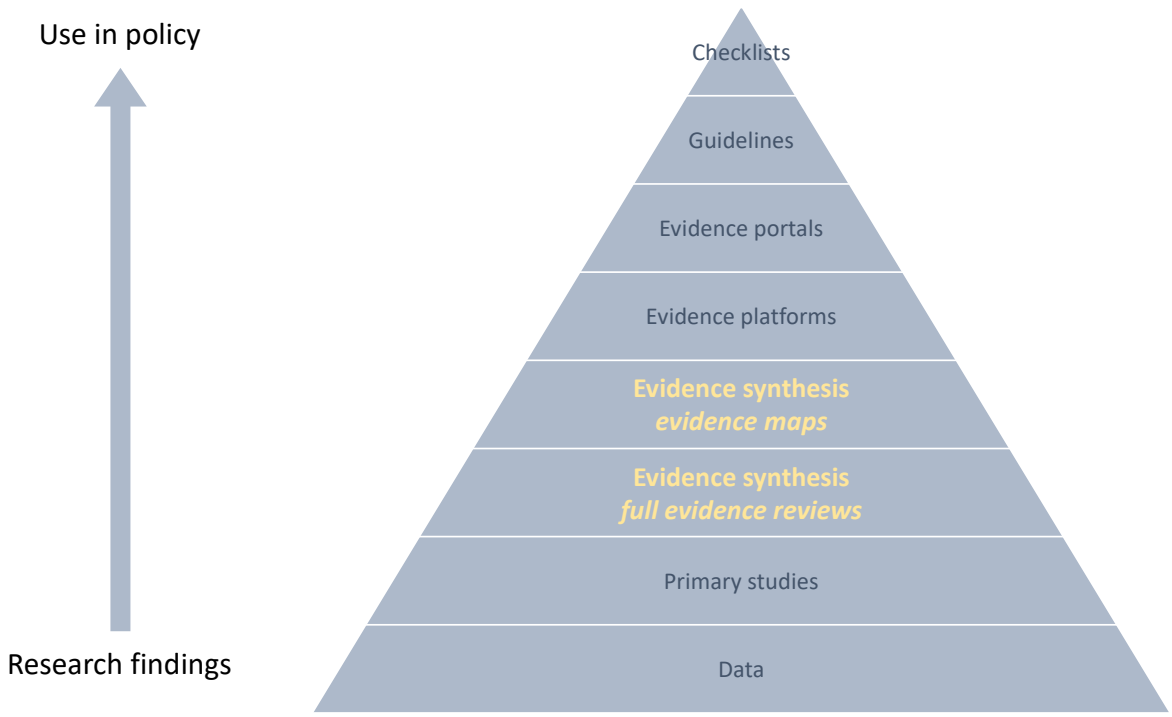

Figure 2. Evidence architecture that forms the framework for evidence-informed decision making, from data to decision-support checklists. Adapted from White (2019).

Evidence synthesis allows a body of evidence that may consist of disparate or conflicting studies to be identified, appraised, described and combined in a consistent and reliable way (CEE 2018). It might otherwise be impossible for decision makers to encompass an entire evidence base. Evidence synthesis is especially relevant for answering more complex questions, or issues where summary of a broad evidence base is relevant and necessary.

Evidence synthesis can be integrated into policy, management and research in any way single studies otherwise would be used (e.g. citations for statements in IPCC or IPBES assessment reports). In this way it is a tool for single projects or for working groups of for example SCOR, ICES, HELCOM or OSPAR. summarising large bodies of research in an easily digestible format, thereby saving time and providing a transparent, reliable basis for policymakers.

Evidence synthesis can be used to summarise broad evidence bases to ask what is the nature of research on a given topic (evidence maps), or a narrower evidence base to ask what the effect of an intervention or impact is (full evidence reviews). In this way, evidence synthesis describes a body of research or summarises the findings of a collection of studies (James et al. 2016). Evidence synthesis is not appropriate for topics where the information to be synthesised is not research evidence (for example, where the knowledge is in the form of Indigenous and local knowledge and not documented in formal research studies): for these topics, another form of knowledge synthesis is a more appropriate tool (see Dicks et al. 2017). This is because evidence synthesis methods are only appropriate for documented scientific research evidence, and cannot be used to synthesise information from other sources.

Furthermore, evidence syntheses are typically focused on narrow topics, even for relatively broad evidence maps. As such, it is highly unlikely that one review alone is an appropriate basis for decision-making. It is likely that a suite of related maps and reviews would be needed to examine different aspects of the issue at hand (Haddaway et al. 2016). For example, it may be appropriate to conduct at least one review on each of the

following: the potential effectiveness of an intervention under controlled conditions; the barriers and facilitators to intervention implementation; the factors affecting intervention effectiveness in practice; the comparative cost-effectiveness of different interventions of known efficacy.

The importance of evidence synthesis

Evidence synthesis can have an important function for three main reasons:

- Firstly, evidence synthesis allows us to make the most of an existing body of research studies, rather than commissioning new research (Haddaway and Bilotta 2016). It also provides more power than any single study would alone (Gurevich et al. 2018) – increasing the effective sample size and study contexts, and allowing the investigation of variables that cannot be examined within any one study.
- Secondly, it allows a large, potentially disparate evidence base to be summarised in a single document that streamlines the process of translating research into policy/practice (Haddaway and Pullin 2014). This in turn means that, assuming rigorous methods have been used for the synthesis, those in need of evidence need only read the synthesis rather than attempting to identify and digest the entire body of literature themselves.
- Thirdly, evidence synthesis is an independent, robust, critical overview of an evidence base that involves a transparent and procedurally objective appraisal of each included study. Again, assuming the methods used and the expertise of the group were appropriate, this means that the reader can trust the findings of the synthesis – something that may be particularly important if the underlying evidence base is conflicting or known to vary in its reliability (Kohl et al. 2015).

In the field of health, evidence synthesis has become an integral part of developing practitioner and policy guidelines and recommendations, so-called evidence-based medicine (Djulbegovic and Guyatt 2017). Cochrane, the evidence synthesis coordinating body in clinical health, has published over 7,500 systematic reviews, with members and supporters from over 220 countries (https://www.cochrane.org/about-us), and demonstrable impact on health research and clinical care (Bunn et al. 2015). Since evidence synthesis has proven useful across the field of health (from clinical medicine (Rooney et al. 2014) to public health policy (Jackson and Waters 2005)) where there exists a high degree of complexity, and given the broad and complex challenges that exist in environmental management (Polasky et al. 2011), it is likely to be a useful tool for environmental management and policy (Stewart et al. 2005).

The role of knowledge transfer in evidence synthesis

Evidence synthesis is used to compile and summarise existing research, with the explicit aim of translating previous research findings into a format that provides a suitable base for decision-making (CEE 2018). As such, it is a form of knowledge translation itself. That said, evidence syntheses require thoughtful communication, since the review should answer relevant questions and the reports themselves are long and technical (Sundin et al. 2018). This technicality results from the cornerstones of synthesis methodology that call for a suite of specific and tailored methods for finding, assessing and synthesising diverse sources of research literature, and for highly detailed and transparent reporting of all activities to allow verification and replication (Gough et al. 2017). Furthermore, single evidence syntheses are likely to only cover a part of a theory of change or conceptual model for an intervention or system, and so multiple syntheses must be communicated and considered together in decision making. Communication is therefore vital to ensure that stakeholders can be provided with concise and clear summaries in a variety of formats suiting their needs, clearly based on and linking to the underlying evidence that has been carefully synthesised.

Evidence syntheses have several important differences relative to other forms of research that warrant specific considerations in relation to knowledge transfer. In short, they relate to: the complex methodology involved; the standard practice of publishing an *a priori* methods protocol first followed by a final report; the need for detailed documentation and transparency leading to long and technical documentation; the

importance and benefit of engaging with stakeholders from question formulation and scope setting through to communication of the review findings; and the typical aim of appealing to a range of audiences from researchers to practitioners, requiring a suite of different communications.

Report aims and objectives in the BANOS context

Firstly, this report aims to support the development of an '*improved process to synthesise primary scientific information*' by learning from the BONUS experiences with synthesis and external projects attempting to support evidence-informed decision making through the commissioning, conduct and communication of evidence syntheses.

Secondly, it aims to '*create a process for improved dissemination and knowledge transfer of scientific results*', by focusing specifically on the nuances of evidence synthesis that warrant targeted approaches to communication.

The report provides recommendations and a range of potential frameworks that could support an efficient evidence synthesis production system with meaningfully engaged stakeholders and strong and effective communication.

This deliverable is based on a compilation of literature that is available on synthesis and an evaluation of the synthesis call in the BONUS programme (BONUS call 2017: Synthesis) by interviewing BONUS-funded projects' coordinators and evaluating projects' outcomes. The aim is to investigate how synthesis could be valuable tool for a potential future programme.

Choosing and using evidence synthesis methods

This section describes the reasons for conducting evidence syntheses and provides the prerequisites for commissioning evidence syntheses. It then outlines the main methods of evidence synthesis available and describes their key characteristics, advantages and disadvantages. Finally, tools are provided for selecting an appropriate synthesis method given a specific context and set of objectives.

There are a number of key reasons that evidence synthesis may be an appropriate course of action given a specific context and evidence need (Land et al. 2017): if any of these are true for an evidence need, it is likely that an evidence synthesis is warranted. These needs are outlined in Table 1, along with the methods that aim to overcome them and the main procedural cornerstone that can be used to achieve the objective. These objectives and the problems they aim to overcome may occur together, or they may occur in isolation. For high-risk policy making (for example, areas where the costs of a ineffective or damaging policy change or inaction are high), many of the objectives will apply.

Table 1. Evidence needs, objectives, and procedural cornerstones of evidence synthesis.

Evidence need	Objective	Procedural cornerstone of evidence synthesis
Continued funding of primary research when existing research may be sufficient for decision-making	Make the most of existing research, rather than conducting novel primary research	All aspects/procedures
Lack of overview of a research area	Understand the nature of a broad evidence base	Evidence mapping
Lack of precise/accurate estimate of effect	Improve precision and accuracy	Aggregative quantitative synthesis
Disparate evidence base	Provide comprehensiveness	Systematic searching of academic and grey literature
Lack of clarity on role of context	Better understand effect modifiers/influence of context	Configurative qualitative/quantitative/mixed synthesis
Lack of rigour in previous syntheses / need for high level of rigour in evidence use	Provide rigorous evidence	Full systematic review or systematic map
Primary research studies disagree	Provide a definitive answer where there is a disagreement between primary research studies	Full systematic review or systematic map
Dispute over the quality of primary research or previous syntheses	Independently appraise the validity of an evidence base	Rigorous critical appraisal
Known cluster of primary research but no synthesis exists	Summarise a known cluster of evidence (e.g. from an evidence map)	Quantitative synthesis (i.e. meta-analysis) within systematic review
Previous synthesis is out-of-date	Update an existing synthesis	Review or map update
Dispute over reproducibility of previous synthesis	Replicate an existing synthesis	Review or map replication
Previous syntheses disagree	Provide a definitive answer where there is a disagreement between previous syntheses	Review of reviews (sometimes referred to as an umbrella review)
Dispute over whether a gap in primary research exists	Demonstrate a lack of evidence	Full systematic review or systematic map

The cornerstones mentioned in Table 1 are integral parts of evidence synthesis methodology, whether it is a rapid review or a full systematic review. More details on how to ensure reviews follow these appropriate methods are provided below.

Prerequisites for evidence synthesis

The following factors that must be met for an evidence synthesis to be appropriate:

- The review question must be answerable using (typically empirical) research studies published in the academic literature or available as grey literature (organisational reports, preprints, theses, white papers, etc.)
- The review question must be well-framed and clear. This may be aided by structuring the question into 'key components', often done using what is referred to as 'PICO/PECO' frameworks (Morgan et al. 2018):

P, population (not necessary a biological population but the unit being examples, e.g. 'the Baltic Sea', or 'fishing rights in the North Sea'); I, intervention (i.e. a policy or management action) or E, exposure; C, comparator; O, outcome. When described explicitly this helps to refine the scope and definitions of the question to avoid ambiguity, but also helps to plan the synthesis activities.

- The question must also be focused. It can still be relatively broad for evidence maps, but must be narrow for systematic reviews.
    - If an aggregative review is needed in order to more precisely estimate effectiveness or impact, the review question should be answerable by equivalent primary studies. If the question is too broad it should be split into multiple (interrelated) reviews.
    - If a configurative review is needed in order to better understand a study system and develop a theory of change or conceptual model (e.g. asking what factors might affect effectiveness of an intervention), then the question may be more broad.
    - If an evidence mapping is needed (i.e. asking what research exists and what methods or systems have been studied, but NOT the findings of the primary studies) the question may be rather broad, but system boundaries for the question's key elements should still be defined.
- Adequate resources should be provided for the evidence synthesis method required: a sufficient time window to plan and complete all tasks (Haddaway and Westgate 2019), including meaningful stakeholder engagement (Haddaway and Crowe 2018); sufficient budget to cover staff salaries for necessary tasks to be completed in a robust way (e.g. dual screening and consistency checking); inclusion of team members with experience of evidence synthesis, advisory support from an evidence synthesis methodologist/specialist, and/or appropriate training in evidence synthesis methods (Eales et al. 2017); expert peer-review oversight (either internal or independent) for robust and timely feedback on the planned methods in order to improve methodology prior to conduct (if external, independent review, then sufficient time and budget must be assigned).
- Commissioners and end-users should have a clear understanding of what the synthesis will provide (Taylor et al. 2017). Evidence syntheses should not suggest policy actions, but rather indicate what might work and when (reviews) or what evidence exists (maps). Policy is broader than any review: a suite of reviews needed for different aspects that must be considered (e.g. financial, biological, social, potential effectiveness, implementation effectiveness), and other information is often vital in considering policy actions other than evidence alone (i.e. evidence-INFORMED policy making).
- The evidence synthesis should include meaningful stakeholder engagement to optimise the review focus and maximise communication and endorsement/acceptance (Haddaway and Crowe 2018). This should be integrated into: scope setting and question formulation; co-design of the *a priori* protocol outlining the planned methods; meaningful responses to stakeholder inputs; communication with key stakeholders before publication; feedback on presentation of the findings and other communication media; balanced engagement that empowers underrepresented/marginalised voices.

Evidence synthesis methods available

There are a range of related evidence synthesis methods that could be relevant for the BANOS programme. We highlight them briefly here and describe the situations in which they may be appropriate. Table 2 outlines the current major types of evidence synthesis methods and indicates whether they are appropriate for use in guiding environmental decision-making.

This assessment is based on widely accepted definitions of evidence-informed decision-making (Higgins et al. 2019; CEE 2018) that have carefully sought to identify limitations and biases in these different forms of review. The higher value of rigorous evidence synthesis for evidence-informed decision making has been widely discussed elsewhere and is therefore not repeated here at length. For further details on why only more rigorous forms of evidence synthesis are suitable for decision making, see (Moat and Lavis 2013; Shlonsky et al 2011; Litell and Shlonsky 2010; Head BW 2016).

In summary:

1. Good quality, reliable reviews are more expensive and take more time.
2. Faster, cheaper reviews are possible, but significantly cheaper and faster options are likely to be very unreliable.
3. Where corners are cut they need to be highlighted clearly so that the reader can weigh up the risk in interpreting the findings.
4. Reviews where corners are cut should not be intended to be used in other contexts – the corners were cut because of local constraints – so these reviews should not be published in international academic journals.

Table 2. Summary of main types of evidence synthesis methods, their aims, methodological cornerstones, descriptions, synthesis aspect (what is novel), general notes and links to guidance. Adapted from Haddaway, Lotfi, Mbuagbaw. 2021. Systematic reviews: a glossary for public health. Scandinavian Journal of Public Health, invited manuscript, in submission. Rows are shaded according to their rigour and appropriateness for supporting decision-making: grey = narrative introductions unlikely to be relevant to decision-making; red = often conducted and/or claimed to be relevant to policy making but reliant on unspecified, limited and/or biased methodology; yellow = often relevant to policy and some formal methodology may be used but subject to limitations and/or bias; green = often relevant to policy and should be conducted and reported according to strict methodological standards that aim to mitigate limitations and bias.

Evidence synthesis type	Synonyms	Aim to improve	Methodological cornerstones	Description	The synthesis (what is novel)	Notes	Guidance
Primer	'Overview', 'literature review', 'traditional literature review'	General understanding	None	Generic term: summary of the [medical] literature that attempts to survey the literature and describe its characteristics	The collation of concepts across a broad evidence base and/or discussion of an evidence base through a novel analytical lens.	A generic term for reviews that do not attempt to be systematic but rather introduce the reader to a new topic	No formal guidance
State-of-the-art review		Understanding of recent methodological or subject-specific developments	None, emphasis on recent research	Tend to address more current matters in contrast to other combined retrospective and current approaches. May offer new perspectives on issue or point out area for further research	Focuses more on recent research developments and may discuss a recent evidence base through a novel analytical lens.		No formal guidance
Scoping review	'Pilot review'	Understanding of the nature of an evidence base (rapidly)	None	Preliminary assessment of potential size and scope of available research literature. Aims to identify nature and extent of research evidence (usually including ongoing research)	The application of some systematic approaches to evidence synthesis in a very restricted way, usually lacking critical appraisal or full synthesis in an attempt to demonstrate comprehensiveness and transparency and minimise susceptibility to bias. Restricted methods are used to save substantial time and resources.	Conducted to rapidly understand the nature of research evidence on a topic, sometimes to assess feasibility of a full systematic review, test the 'answerability' of the review question. Scoping reviews have, in some instances, been given formalised methodologies (e.g. UK civil service), but these are generally still open to interpretation in different ways, leading to substantial differences in the quality of conduct.	Pham, M.T., Rajić, A., Greig, J.D., Sargeant, J.M., Papadopoulos, A., McEwen, S.A. A scoping review of scoping reviews: advancing the approach and enhancing the consistency. Res Synth Method 5 371-385 (2014). https://doi.org/10.1002/jrsm.1123
Systematic map	'Synopsis', 'evidence map', 'systematic mapping review'	Understanding of the nature of an evidence base	Comprehensiveness, transparency (for accountability/repeatability), bias mitigation	Map out and categorise existing literature on a broad subject from which to commission further reviews (knowledge clusters) and/or primary research by identifying gaps in research literature (knowledge gaps)	A formal and fully systematic approach to all aspects of the way in which studies are found, screened for relevance, catalogued, and combined in an attempt to demonstrate comprehensiveness and transparency and minimise susceptibility to bias. A searchable database of studies described across different variables is a typical novel output for such a review.	Generally accepted as the most reliable or 'gold standard' form of evidence synthesis. Attempts to describe the nature of an evidence base (what research has been done where, how, when and by whom?). Does not fully synthesis study findings as does a systematic review'	James, K.L., Randall, N.P. & Haddaway, N.R. A methodology for systematic mapping in environmental sciences. Environ Evid 5, 7 (2016). https://doi.org/10.1186/s13750-016-0059-6
Rapid review	'Responsive review', 'rapid	Understanding of the effect or effectiveness of an action/impact	**Some** aspects of: comprehensiveness, transparency (for	Aggregative or configurative assessment of a specific body of evidence, attempting to	A systematic approach to some (but not all) aspects of the way in which studies are	A broad category of reviews varying substantially in reliability depending on the nature of the corners that are cut. Rapid	Tricco, A.C., Antony, J., Zarin, W. et al. A scoping review of rapid review methods. BMC Med 13, 224 (2015).

	evidence assessment'		accountability/repeatability), bias mitigation	understand an 'average' effect across studies. Attempt to include elements of systematic review process while stopping short of systematic review	found, screened for relevance, appraised for validity, and combined in an attempt to demonstrate comprehensiveness and transparency and minimise susceptibility to bias. Methodological corners often cut in response to resource or time constraints.	evidence assessments (REAs; also termed rapid reviews) have, in some instances, been given formalised methodologies (e.g. UK civil service), but these are generally still open to interpretation in different ways, leading to substantial differences in the quality of conduct.	https://doi.org/10.1186/s12916-015-0465-6
So-called 'systematic review'	'Systematic-style review', 'Systematic literature review', 'literature review with a systematic search', inappropriately labelled 'systematic review'	Understanding of the effect or effectiveness of an action/impact	**Some** aspects of: comprehensiveness, transparency (for accountability/repeatability), bias mitigation	Aggregative or configurative assessment of a specific body of evidence, attempting to understand an 'average' effect across studies. Attempt to include elements of systematic review process while stopping short of systematic review	A systematic approach to some (but not all) aspects of the way in which studies are found, screened for relevance, appraised for validity, and combined in an attempt to demonstrate comprehensiveness and transparency and minimise susceptibility to bias. Methodological corners often cut in response to resource or time constraints.	A broad category of reviews varying substantially in reliability depending on the nature of the corners that are cut.	Haddaway, N.R., Bethel, A., Dicks, L.V. et al. Eight problems with literature reviews and how to fix them. Nat Ecol Evol 4, 1582–1589 (2020). https://doi.org/10.1038/s41559-020-01295-x
Systematic review	May be additionally labelled as 'qualitative evidence synthesis', quantitative synthesis' or 'mixed methods review'	Understanding of the effect or effectiveness of an action/impact. Precision and power of an analysis of impact/effectiveness (via quantitative synthesis), examine the effect of contextual factors across studies (via qualitative synthesis)	**All** aspects of: comprehensiveness, transparency (for accountability/repeatability), bias mitigation Also considers: precision, power	Seeks to systematically search for, include, appraise and synthesis research evidence, often adhering to guidelines on the conduct of a review	A formal and fully systematic approach to all aspects of the way in which studies are found, screened for relevance, appraised for validity, and combined in an attempt to demonstrate comprehensiveness and transparency and minimise susceptibility to bias.	Generally accepted as the most reliable or 'gold standard' form of evidence synthesis	Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook.
Realist review	'Realist synthesis'	Understanding regarding an intervention's underlying causal mechanisms, how they work, and under what conditions	Stakeholder engagement, purposive searching and inclusion driven by theory, iterative processes of searching, synthesis and evaluation, focus on understanding causal mechanisms rather than predicting outcomes	'...focuses on providing explanations for why interventions may or may not work, in what contexts, how and in what circumstances' (Greenhalgh T, Wong G, Westhorp G, Pawson R: Protocol – realist and meta narrative evidence synthesis: Evolving standards (RAMESES). Protocol – realist and meta narrative evidence synthesis: Evolving standards (RAMESES) BMC Med Res Methodol. 2011, 11:)	Understanding of causal mechanisms under different contexts		Rycroft-Malone, J., McCormack, B., Hutchinson, A.M. et al. Realist synthesis: illustrating the method for implementation research. Implementation Sci 7, 33 (2012). https://doi.org/10.1186/1748-5908-7-33

Overview of reviews	'Systematic review of systematic reviews', 'review of reviews', 'umbrella review'	Understanding of the effect or effectiveness of an action/impact. Precision and power of an analysis of impact/effectiveness, examine the effect of contextual factors across studies	**All** aspects of: comprehensiveness, transparency (for accountability/repeatability), bias mitigation. Also considers interactivity, usability	Specifically refers to review compiling evidence from multiple reviews into one accessible and usable document. Focuses on broad condition or problem for which there are competing interventions and highlights reviews that address these interventions and their results	Reviews (sometimes only formal systematic reviews) are collated, described, appraised and combined (sometimes statistically) in a formal systematic approach (see 'Systematic review', above).	Conducted regularly within healthcare where many systematic reviews have been published on the same topic	Pollock M, Fernandes RM, Becker LA, Pieper D, Hartling L. Chapter V: Overviews of Reviews. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook.
Literature review	'Narrative review', 'literature review', 'evidence review', 'evidence synthesis'	(Usually) understanding of the effect or effectiveness of an action/impact	**May** possess some aspects of: comprehensiveness, transparency (for accountability/repeatability), bias mitigation	Aggregative or configurative assessment of a specific body of evidence, attempting to understand an 'average' effect across studies. Typically lacking in transparency/methodological detail, may include a meta-analysis.	The collation of concepts across a broad evidence base and/or discussion of an evidence base through a novel analytical lens. May involve a systematic approach to some (but not all) aspects of the way in which studies are found, screened for relevance, appraised for validity, and combined in an attempt to demonstrate comprehensiveness and transparency and minimise susceptibility to bias. Methodological corners often cut in response to resource or time constraints.	A broad category of reviews varying substantially in reliability depending on the nature of the corners that are cut. Typically not seen as research items in their own right and lack a methodology describing their conduct.	Haddaway, N.R., Woodcock, P., Macura, B., Collins, A. Making literature reviews more reliable through application of lessons from systematic reviews. Cons Biol 29 1596-1605 (2015). https://doi.org/10.1111/cobi.12541
Meta-analysis	'Quantitative synthesis'	Understanding of the effect or effectiveness of an action/impact. Precision and power of an analysis of impact/effectiveness, examine the effect of contextual factors across studies (e.g. through meta-regression)	Precision, power	Technique that statistically combines the results of quantitative studies to provide a more precise effect of the results	Use of powerful statistical tools to combine study findings together to produce a summary effect estimate and measure of variability for groups of studies measuring similar outcomes.	Many researchers and organisations (e.g. Cochrane, CEE, Campbell Collaboration) believe that meta-analysis should only be conducted in the context of a formal systematic review (see 'Systematic review', below). Meta-analysis is a statistical tool for combining study findings, and the methods do not specify how the studies are located in the first place. Often no critical appraisal of study validity is conducted.	e.g. Schmid CH, Stijnen T, White I, editors. Handbook of Meta-analysis. CRC Press; 2020 Sep 8.
Critical review	'Critique', 'critical commentary', 'critical appraisal', 'critical analysis'	Understanding of evidence validity	(Informal) Critical appraisal	Aims to demonstrate writer has extensively researched literature and critically evaluated its quality. Goes beyond mere description to include degree of analysis and conceptual innovation. Typically results in hypothesis or model	The collation of concepts across a broad evidence base in combination with critical appraisal of the validity of studies and the evidence base.	Sometimes used to describe an evaluation of one particular text, which is not a form of evidence synthesis.	No formal guidance

Considerations supporting evidence synthesis method choice

There are a number of considerations that can help to select an evidence synthesis method, as follows:

- The type of 'good' that the review seeks to fulfil:
  - '*Global public good*' – The synthesis should be made available to all (i.e. via publication in an academic journal), often generalised across multiple contexts but ideally also providing explicit context-based conclusions. The review can be taken and used 'off-the-shelf', without the need for local adaptation or further synthesis/analysis. For these reviews, international accepted standards for evidence synthesis should be used, since they are likely to be read by multiple end users. Furthermore, methodological corners should not be cut in conducting the review, since it will be archived and used by others who may assume a high level of rigour has been used. Ideally, these should be regularly updated as new evidence becomes available.
  - 'Private good' – The synthesis is designed for a specific end-user, and is applicable to a limited number of specific contexts relevant to the commissioner or end-user. These reviews are often not published in international journals because of their specificity, although they are often made publicly available on institutional websites. The review is not aimed at being used in other contexts. For these reviews, the repeatability of the methodology may not be a key cornerstone, since some included evidence may not be publicly sharable or accessible to researchers outside the project. As a result, formalised evidence synthesis methods may not be necessary/appropriate. It may also be appropriate to cut corners, methodologically in order to complete the task within a given time frame.
- The objective of the synthesis:
  - *Evidence mapping* – Sometimes, the intention of an evidence synthesis is to improve understanding of what research exists in the literature and studies are catalogued and described according to their contexts, methods and settings. In these cases study findings are not extracted and synthesised – this is instead left for any subsequent full synthesis. The questions being answered relate to what we know about an evidence base and NOT what we know about an intervention or impact.
  - *Full synthesis* – In contrast, we may wish to better understand an intervention, impact or system, rather than the research studies themselves. In these cases, syntheses should seek to perform a full synthesis of the findings of a set of studies. This may be to either ask about the effectiveness or efficacy of an intervention or impact (*aggregative* synthesis), or it may be to better understand how a system functions or what factors affect a pathways to impact (*configurative* synthesis).
- The availability and type of evidence to be synthesised:
  - *Research studies and data* – This is traditionally the main focus of evidence syntheses. These studies may be published in traditional academic journals, but a considerable proportion of the evidence base may consist of grey literature (i.e. reports and papers published by non-academic organisations on all levels, and also theses and preprints). Evidence syntheses should also seek to integrate not-yet-published studies (so-called 'file-drawer research') through contact with stakeholders, trial registries and data repositories.
  - *Non-research information* – Some questions may call for the synthesis of information not primarily in the research literature, for example policies, patents, or information from other knowledge systems, such as practitioner knowledge (e.g. biodiversity action plans or national data reporting for example in Helcom context). Evidence synthesis methods may still be useful for these kinds of questions if comprehensiveness, representativeness, accuracy, precision, procedural objectivity and transparency/repeatability are important. Systematic review

methods can be easily adapted for any information system by following the core principles despite differences in data sources. An example of this can be seen in the published protocol for a systematic map of environmental and social impact assessments from the Formas-funded 3MK project at the Stockholm Environment Institute (SEI); Macura et al. (2019).

- *Non-public data* – Where most or all of the information relevant to a question is not publicly available, formal evidence synthesis methods are likely to be inappropriate, since they rely on transparency and repeatability. Here, the review may not itself be published in the end (see 'private goods' reviews above). Alternatively, if non-public information is gathered from a disparate and select group of stakeholders or experts, for example, the repeatability may be deemed to be so low that another formal knowledge synthesis method should be used.

- The availability of time and resources:
  - *Time available* – Ideally, knowledge needs are identified sufficiently far in advance that they would not affect the choice of method of evidence synthesis, but often windows for evidence use are, for whatever reason, narrow. Where years are available, full and formal evidence synthesis should always be the preferred choice, given the conditions above. If only months are available, a condensed form of rapid review should be considered – for example excluding dual screening or consistency checking, or focusing on a random sample of the evidence base. Where weeks are available, existing reviews can be searched for and appraised, condensing multiple reviews using a form of 'review of reviews' approach (see Table 2). If only days are available, efforts to identify rigorous evidence syntheses may prove useful – potentially highlighting existing systematic reviews, for example. In fact, this final approach is advisable before commencing any novel evidence synthesis.
  - *Resources available* – Where available budgets are low, considerations regarding the most suitable method should be based on a translation of funding into staff time, since this is the core component of an evidence synthesis project. However, other considerations relating to resourcing could affect planning of a synthesis. Explicitly, the need for training, mentoring or inclusion of expertise on the project has associated costs, and will always result in a more robust output. If these resources are not available, then evidence synthesis projects should arguably not take place.

Considerations relating to frameworks and systems for conducting evidence syntheses; lessons learned from previous calls

This section considers frameworks for conducting synthesis to support decision-making in policy and practice. To include lessons learned, an evaluation is included of the BONUS synthesis projects format (BONUS call 2017: Synthesis) as well as some key examples of alternative approaches to commissioning, conducting and communicating evidence syntheses for a particular topic.

Different formats have been tested for evidence synthesis calls. All of them have advantages and disadvantages, which are listed below.

1. **Multiple specific-topic (open) call**

The BONUS synthesis call is an example of an open call which requested proposals for synthesis on 9 topics, 8 of which were funded. The call was relatively unique in the level of specificity provided concerning the desired tasks for each topic. For each topic there were between 3 and 6 sub-tasks to be completed, with some topics therefore requiring multiple systematic reviews. Successful project teams were required to use systematic review methods 'wherever possible and appropriate', and to produce at least three review papers, each with their own policy brief. Each funded project was awarded 450 000 Euro over 1.5 years.

Evaluation of the call material supplemented with interviews with funded BONUS synthesis project coordinators and users of some of the generated results (for more detail see Annex B) suggested that there was a diverse understanding of synthesis terminology and a resultant application of a diverse set of synthesis methods. It also highlighted that it could be useful to provide tools to applicants and carefully refine the call topics in close collaboration with end-users already before publishing to ensure that project outputs are available on an appropriate time line and in a format that is easily digestible and understandable. No protocols (i.e. *a priori* detailed plans of the synthesis methods that would be used) were made available other than for 1 project (Storie et al. 2020), and formal evidence synthesis methods were very rarely used (see Annex B Table 4), indicating the novelty of synthesis calls and limited awareness or application of formal evidence synthesis methodologies amongst environmental scientists in the Baltic Sea region. This document seeks to provide some tools that would support the commissioning and conduct of more rigorous evidence syntheses.

Strengths: Resources used to plan, coordinate and review the call together with stakeholders, ensure that the call can be very targeted, increasing the impact of the outcome, submissions and funded projects can be more efficiently used across multiple simultaneous projects. Collaboration across projects is possible because of the similar timing.

Weaknesses: The topics are very specific and must be carefully planned and scoped prior to the call to ensure they are appropriately defined and feasible with the given resources. In addition, it is crucial that the topics are fitting within the management time-frame.

2. **Appointment of an approved supplier**

The Department for Environment, Farming and Rural Affairs (Defra) in the United Kingdom until recently had an internal initiative to produce and provide fit-for-purpose evidence synthesis for use in decision-making within the organisation; the Joint Water Evidence Group (JWEG). JWEG commissioned various

partner organisations to conduct different types of evidence synthesis (including systematic reviews, systematic maps, rapid evidence assessments and scoping reviews) on various environmental topics. As part of this process, a public call was made for applications to act as one of several 'approved suppliers' who would be commissioned directly to conduct evidence syntheses as and when they were needed. JWEG worked closely with each supplier (who had been vetted as being competent in evidence synthesis methods) to assemble an advisory group to set the scope, definitions and question. They also reviewed the *a priori* protocol setting out the planned methods and reviewed the final report before communicating the findings to the commissioners within Defra.

Strengths: This approach allows syntheses to be initiated in a responsive, targeted and efficient way, since a competent team has already been assembled. It also allows oversight of the planning and conduct of the syntheses. Furthermore, communication of the findings can be centralised, which is particularly useful for overlapping topics with end-users who have limited time to engage with workshops and other forms of direct contact.

Weaknesses: The oversight needed to form advisory groups, review protocols and final reports requires additional resources and methodological awareness and expertise from the core team (JWEG in the case above). Approved suppliers are assumed to have sufficient time and flexibility to respond to synthesis requests within a short time period. This may be more likely for teams from consultancies than research institutes, which may affect the competencies available.

3. Core funding for multiple reviews led by a central coordinating team

The Swedish Mistra Council for Evidence-Based Environmental Management was established in 2012 with a large core funding of 60 million SEK and the remit to conduct c. 15 evidence syntheses pertaining to the Swedish environmental objectives. Mistra established a Board, which in turn hired a Secretariat of 5 project managers who would assemble review teams of internationally renowned subject experts to conduct each review, with the supervision of a project manager as the methodologists. The topics for the reviews were suggested by a diverse group of national stakeholders, after which the Secretariat reformulated the topics into review questions where possible and conducted pilot studies after consultation with the Board. The Board then reviewed the pilot studies and selected topics to take forwards as systematic reviews or systematic maps. Each topic then assembled its own stakeholder group for final question formulation and protocol co-design and assembled an expert review team.

The UK-based "What Works Centres" were established as clearinghouses for supporting evidence-informed decision making in social policy through assembling research evidence on the effectiveness of social interventions. Each What Works Centre was established and run in a slightly different way, with a different focus on the stage of evidence production and communication. The What Works Centre for Crime Reduction (WWCCR), for example, was established with dedicated funding to conduct needs-based evidence syntheses and also primary research relating to crime reduction interventions. The Centre also developed user-centric decision-support tools based on the evidence produced (see final report for the WWCCR here; https://whatworks.college.police.uk/About/Documents/ICPR_Final_Evaluation_WWCCR.pdf).

Strengths: The centralisation of knowledge need assessment, stakeholder engagement, project management, and communication can increase efficiency of projects and reduce burden on a limited group of stakeholders. Expertise and experience in evidence synthesis can also be retained through a

core coordinating team/secretariat working in collaboration and across multiple reviews. This helps to ensure a common and high standard for all outputs, and also helps to increase the efficiency of review conduct.

Weaknesses: These initiatives require substantial funding for coordination of a core coordinating team/secretariat, stakeholder engagement to identify evidence needs, conduct of the reviews, and communication of the findings.

4. Individual topic-specific procurement

The Green Climate Fund (GCF) has commissioned specific systematic reviews in recent years by having individual open calls. Each review topic is fixed and specific, with applicants open to propose their own methods and refinement processes within this relatively narrow scope. The calls typically include detailed key elements and inclusion criteria that cannot be revised by the applicants. For the GCF, the calls often have a very short application period of 3 to 4 weeks (see a recent example of a tender here; https://www.greenclimate.fund/procurement/rfp-2020-017) and budgets are typically not announced, placing onus on the applicant to suggest a feasible set of required resources. The evaluation of proposals includes significant assessment of the applicant's ability to conduct evidence syntheses.

Strengths: Applicants must demonstrate an understanding of evidence synthesis and provide evidence of the ability to conduct systematic reviews, reducing risks and increasing the likelihood of methodological success. Timelines are typically short and rigid, helping to ensure results are available within a known policy window.

Weaknesses: These calls assume that teams with sufficient methodological experience and expertise exist and will apply. There is a risk, therefore, that these highly subject-specific calls with short application periods may not attract a suitable calibre of applicants. Risks associated with selecting an experienced team are high, since timelines for the projects are short and rigid. Considerable effort must be spent by the funder to ensure the call topics are sensible and feasible within the resources and time constraints provided.

5. Multiple broad-topic call

The Swedish Environmental Protection Agency recently opened a call for funding for evidence synthesis pertaining to wildlife management (for more details see; https://www.swedishepa.se/Guidance/Research/For-applicants-the-Wildlife-Management-Fund/). This is a two part fund, one part of which will fund a review of current knowledge on wildlife management. The call is broad and non-specific, so open to applicant's own suggestions of topics from across natural science, social science or the humanities.

The UK National Environmental Research Council made a call for proposals for systematic maps in a very broad remit relating to the environment in the context of the UK's exit from the EU (for more details see; https://nerc.ukri.org/research/partnerships/national/programmes/eef/news/ao-initiative/ao/). A total of 400,000 GBP was made available and 5 projects were funded. The call was made in partnership with the Collaboration for Environmental Evidence (CEE; https://environmentalevidence.org), who aided in reviewing the text for the call and peer-reviewing applications. All projects were required to submit their protocols and final reviews for publication in Environmental Evidence, CEE's journal dedicated to publishing environmental systematic reviews.

Strengths: If emphasis is made on topics being selected in consultation with stakeholders, this can be an efficient way to solicit topics that does not require considerable effort from the funder. Support from experts in evidence synthesis can help to ensure high quality, particularly if previous experience in conducting evidence syntheses is an evaluation criterion.

Weaknesses: The call has little control over the topics proposed within the broad scope, and efforts to adjust the applicants' scope require expertise from funders and are likely to be challenging.

Communication, stakeholder engagement and knowledge transfer in BANOS context

In this section, special considerations are described around communication, stakeholders and knowledge transfer when working with evidence syntheses in the context of BANOS. Considerations are illustrated with experiences from the BONUS synthesis call (2017) and interviews with managers at the Swedish Agency for Marine and Water Management.

Several issues are important to take into account regarding stakeholder engagement and communication when undertaking evidence syntheses. Without these considerations, the syntheses will not reach their full potential:

Language difficulties

Evidence syntheses are by their very nature procedurally complex undertakings and require the use of established methods and procedures to minimise bias and maximise objectivity, comprehensiveness, and transparency. As such, evidence syntheses are associated with a suite of complex terminology that may not be immediately understandable to a lay reader. Translating the findings of evidence syntheses may therefore include an additional step in order to summarise the process and findings for an academic audience prior to usual knowledge transfer practices. It was confirmed by interviews with several BONUS synthesis project leaders that the complexity and size of the produced material could be challenging for stakeholders. Furthermore, there may be a semantic gap between researchers and end users, meaning that careful planning and meaningful early stage engagement is vital to ensure terminology is understood in the same way and that final reports use a broadly shared understanding of key concepts and definitions.

Stakeholder engagement for communication

Robust evidence syntheses involve stakeholders from the outset; during the formulation of the question, setting the review's scope, and defining key terms and concepts. This provides vital input to the review methods plan (the protocol), that should ideally be made public (published), preferably in a journal with capacity to provide formalised and supportive peer-review feedback prior to the review commencing. This engagement of stakeholders from the beginning of the review supports buy-in and awareness raising for the review: it generates a community to whom the final results of the review can be easily communicated. Furthermore, the feeling of involvement from the outset (if done meaningfully and tokenism is avoided) can help to communicate the review amongst the stakeholders' own networks, increasing the likelihood of impact. In this way, communication is integrally linked with stakeholder engagement and should commence from the planning stages of the evidence syntheses. Herein, we refer to stakeholder engagement as a means of both improving the validity and relevance of the review and facilitating and optimising communication. Several BONUS synthesis projects involved stakeholders from the very beginning of the projects, whilst others only involved them later during the project lifetime. Much of the stakeholder engagement took place through already existing contacts (participation of project partners in working groups, correspondence groups or through already existing contacts), several indicated difficulty to take in new stakeholders, many indicated that engagement needs particular attention and sufficient allocated resources and time. Several stakeholders from the Swedish Agency for Marine and Water management indicated that it would have been relevant with an increased stakeholder engagement (towards NGOs, policy and management organisations), and additional (digital) workshops and seminars.

Further information, recommendations and guidance on stakeholder engagement in evidence synthesis can be found in the special series of papers published in Environmental Evidence (https://www.biomedcentral.com/collections/SESRM) and the stakeholder engagement and synthesis web portal (https://stakeholdersandsynthesis.github.io/). The main principles of stakeholder engagement in evidence synthesis are outlined below.

- *Stakeholder identification and mapping*

Stakeholders in the context of evidence synthesis represent any individual or group who might affect or be affected by an evidence synthesis project (Haddaway et al. 2017). Examples of stakeholders were summarised in Lescroart, et al (2019). Stakeholder mapping can help to identify and prioritise stakeholders to ensure that the engagement and communication activities are as effective and efficient as possible (Haddaway et al. 2017). Interest-influence matrices, for example, help to identify which stakeholders to focus efforts on (i.e. those with high interest and high influence), and in which way (e.g. high interest low impact can be targeted with newsletters, low interest high impact can be targeted with one-on-one meetings) (Reed and Curzon 2015). It has to be noted that the stakeholder landscape greatly differs in different countries, with some structures being accessible with more ready procedures for collaboration, whereas the system in other countries is so fragmented that identification and engagement are challenging (interviews)

Several frameworks exist for understanding the role of different stakeholders in evidence synthesis, but the tripartite framework in Haddaway et al. (2017) proposes that stakeholders can be thought of in terms of the actor (the who), their role (the what) and the actions they may perform in the review (the how). This can also be useful in identifying when stakeholders can be contacted to the greatest mutual benefit (so-called phased contact), to avoid burdening all stakeholders throughout the process and ensuring there are sufficient resources to have meaningful stakeholder engagement.

- *Question formulation and scope setting*

Involving stakeholders in formulating the research question and setting the scope (i.e. deciding the system boundaries and key definitions) not only helps to ensure relevance of the review to end users, but it also helps to ensure a shared understanding of those key definitions (Haddaway et al. 2017), particularly in diverse or multidisciplinary/cross-disciplinary topics. As a result, this helps to ensure effective communication. Furthermore, involving stakeholders at this stage helps to foster a feeling of ownership and increase the likelihood of 'endorsement' of the review findings (thereby supporting communication) (Spök et al. 2018). It was indicated in several interviews with BONUS PIs that stakeholders had completely different values and mindset, and did not always understand methods and results, illustrating the relevance of discussion. Several project leaders indicated challenges understanding management processes, which can be an important prerequisite for producing relevant outputs.

- *Protocol and planning*

It may also be advantageous to involve some stakeholders in the development of the *a priori* protocol (planning document) for the synthesis, by soliciting suggestions of search terms, sources of grey (and other) literature, and comments on understandability of terminology in the protocol (Land et al. 2017). This also helps to raise awareness and understanding of the project, thereby facilitating clear communication of the findings, particularly the importance of rigour in the review process (Haddaway et

al. 2017). For several of the BONUS synthesis projects, it was indicated by managers at the Swedish Agency for Marine and Water Management that the projects highlighted areas for new research and confirmed the relevance of ongoing practice and shared views, but several did not provide new insights. This was partly because the timing of the projects was in discrepancy with management needs, because changes in reporting practice or political decisions were not taken into account, partly because the projects missed practical aspects of relevance in relation to their approach. A closer link to management could therefore potentially have benefitted project planning.

- *Providing evidence*

Stakeholders may be an important source of information for the review, not only in terms of suggestions of scope, definitions, search terms and sources of literature, but also providing the literature themselves. This may be particularly useful where there is likely to be so-called 'file-drawer' research (i.e. research studies that have not been published in the literature for some reason), or where an organisation is unlikely to have the facilities to publish, archive, or signpost their research reports in an easily discoverable way (Haddaway and Bayliss 2015). Furthermore, stakeholders (including researchers) can be an important source of missing information where details are lacking from published research (Selph et al. 2014).

- *Tailoring communications*

Communication in evidence syntheses can be vastly improved by engaging with target end users to ascertain what type of media are most likely to be read and easily understood (Sundin et al. 2018). In this way, a subset of stakeholders may be willing to help provide advice and comments to tailor communications media to the right audiences. As mentioned above, this may also foster a sense of involvement that broadens the reach of communication by making use of their direct networks. For tailoring communications, understanding of management processes and providing exactly the right input is essential. Several interviews with BONUS synthesis projects indicated that there are lacking pathways to implementation of new methods in management, that knowledge transfer to management is challenging and slow. Also, time constraints at the management side can be a critical issue.

- *Further considerations*

Along with the above, some further considerations that are important to bear in mind during stakeholder engagement and communication in relation to evidence syntheses are as follows. Firstly, terminology associated with evidence syntheses can be confusing and complex, and review authors may wish to avoid describing evidence synthesis methods in detail to some stakeholders. A stakeholder mapping of the understanding of science, research methods, and evidence informed policy can help to decide which stakeholders to communicate to in which language.

Secondly, certain stakeholders are increasingly approached by researchers with requests to contribute to projects. As a result, review authors should be encouraged to avoid overburdening their stakeholders with contacts and requests for participation or input. It may be possible to compensate stakeholders for their time financially, but where it is not feasible, it may be advisable to target certain stakeholders at particular stages of the review, using a combination of passive and active communication. Furthermore, related projects should ideally combine engagement and communication efforts to make it easier to contribute – this is particularly the case for multiple related systematic review projects: shared engagement could benefit multiple reviews through a single point of interaction and input. Combining

stakeholder communication efforts has been central and highly appreciated in the BONUS programme (Snoeijs-Leijonmalm et al 2017).

Thirdly, effective stakeholder engagement can require substantial resources, and it is easily done poorly (Haddaway et al. 2017). Since reviews are conducted in often complex and nuanced stakeholder 'ecosystems', review authors should be very careful to avoid 'tokenism' in their engagement (Hahn et al. 2017). Tokenistic engagement could be conducted for the sake of appearing to involve, when the opportunities and incentives for being involved are unclear or limited. Care should be taken to appropriately plan and budget for meaningful engagement, particularly where topics are highly controversial, and/or where vulnerable, marginalised or minority groups are involved (Kløcker Larsen and Nilsson 2017).

Recommendations

Based on the literature and other material that is available, as well as on an analysis of the call text, the project outputs and interviews with project coordinators and potential users of the project findings within the BONUS synthesis call, this section provides recommendations and alternative courses of action for eventual future BANOS evidence synthesis efforts.

The following recommendations may increase the efficiency, rigour, and success of future synthesis calls.

1. *Consistent and clear terminology and standards for synthesis throughout the call*

Due to the inconsistent understanding and application of the term 'synthesis' it would be prudent to be specific in what is requested of applicants, as well as using widely accepted and clear terminology describing the exact output required.

The awareness and capacity for rigorous evidence synthesis amongst Baltic environmental researchers, managers and funders appears to be highly variable, with some interviewed project coordinators clearly aware of robust systematic review guidance and standards. The same is expected for researchers from the North Sea region. Future calls would benefit from attempts to ensure all applicants share an awareness of evidence synthesis methods by spelling this out clearly in the call. Choosing when systematic review methods are appropriate is a complex matter that requires considerable methodological expertise, so it is advisable to consider carefully when choosing call topics.

Where multiple, prespecified call topics are outlined again in future calls (as was done for the BONUS Synthesis call), we suggest that the desired methodology for each task should be specified clearly in the main call text. For example, where the synthesis desired relates to recent technological developments, the following text might be used: "Applicants should use systematic approach to collating and assessing technologies from across web-based sources and stakeholder suggestions". Where evidence synthesis of primary research studies is appropriate, the following text may be useful: "Applicants should employ robust systematic mapping methods using international accepted guidance and standards (e.g. CEE 2018)".

Because of the possibility for confusion, in general, the term 'synthesis' alone should be avoided unless the possible or necessary methods also specified clearly for each project/task.

2. *Where evidence synthesis (i.e. systematic reviews) is not what is requested from applicants, future calls should be explicit about what methods should be used and what evidence is expected to be analysed*

Where future synthesis calls are planned for projects or tasks for which evidence synthesis is already known not to be suitable (for example for current technologies not yet in the published research literature), the call should be explicit about what methods should be used. For example, where previous BONUS projects are to be summarised and analysed, formal project evaluation methods should be promoted. The timing of these projects should also be carefully considered, since one interviewee stated that previous relevant BONUS-funded primary research projects were not eligible to be included in their synthesis, since the projects had not been completed and the results not available before the synthesis project began.

3. *Call topics amenable to evidence synthesis and timed to coincide with timing of policy/management decisions*

If evidence synthesis in particular is the desired method for call topics, it is vital that the topics and tasks themselves are amenable to systematic reviewing or mapping. The following considerations may help to assess and ensure this:

- The question should be sufficiently narrow to be answerable using evidence synthesis methods. In practice, this means that the review question should, at least in theory, be answerable itself by one or a small number of research studies. For example, 'How effective are agricultural interventions in reducing pollution in the Baltic?' is likely to be too broad for a review, and can be broken down into multiple review questions focused on the effectiveness or pathways to impact of groups of interventions or single interventions. A suitable systematic mapping question for this topic could be; "What evidence exists on the effectiveness of agricultural interventions for reducing pollution in the Baltic?", although this is still a rather broad topic.
- The question should be well-defined, and the scope set according to key elements that can be used to structure systematic searching, inclusion criteria for study eligibility assessment, and extraction of study data and meta-data (descriptive information). In practice, this means ensuring that all core concepts are defined in a way that is widely accepted by the end users.
- Suitable evidence of sufficient validity should exist before a systematic review is commissioned. Although there are notable examples of important 'empty reviews' being conducted to a high standard that finding no evidence, it is likely that this would otherwise be deemed a high risk where evidence is needed to inform policy. Where an evidence base is unknown, broad evidence mapping (e.g. using systematic mapping methods (James et al. 2016)) may be more appropriate to initially identify gaps and clusters in research knowledge, before selecting subtopics where a sufficient validity and volume of evidence exists to permit full synthesis via systematic reviews (Haddaway et al. 2016). This could either be done as two staggered calls (the map results informing the topic for the second call), or as a single call with a request for a map and a subsequent review.
- Broad stakeholder engagement may be particularly useful in both setting the scope and definitions of a review question in a call, and ensuring that the timing of the milestones required match the evidence needs of end users (Haddaway et al. 2017).

4. *Links to methodological guidance, training and support*

Providing links to guidance, training and support in methodologies desired to be part of the call could be beneficial for two reasons. Firstly, it may increase the quality of applications by explicitly demonstrating what is meant by 'systematic review' for example, setting expectations. Secondly, it may also support those wishing to increase their expertise in and capacity for evidence synthesis, increasing capacity for rigorous reviews in the body of available applicants.

A vast amount of guidance and support is available for evidence synthesis methods online, and a concise list follows here. In addition, a range of free training resources are available online. These resources have been developed across disciplines but are equally as applicable for the focus on BONUS and BANOS.

Table 3. Examples of guidance, support and training freely available online. Green, training; blue, guidance; yellow, reporting standards.

Resource	Organisation	Link
Systematic review and mapping methods course (self-paced)	Stockholm Environment Institute, Collaboration for Environmental Evidence	https://systematicreviewmethods.github.io/
Stakeholder engagement and evidence synthesis methods course (self-paced)	Stockholm Environment Institute, Collaboration for Environmental Evidence	https://stakeholderengagementtraining.github.io/
Systematic mapping methods course (self-paced)	Stockholm Environment Institute, Collaboration for Environmental Evidence	https://systematicmappingmethods.github.io/
Introduction to systematic reviews and meta-analysis (self-paced)	John Hopkins University	https://www.coursera.org/learn/systematic-review
Evidence Synthesis Academy (various, self-paced courses)	Brown University, AHRQ	https://evsynthacademy.org/
Guidelines and Standards for Evidence Synthesis in Environmental Management (environment)	Collaboration for Environmental Evidence	https://environmentalevidence.org/information-for-authors/
A methodology for systematic mapping in environmental sciences (environment)	Collaboration for Environmental Evidence	https://environmentalevidencejournal.biomedcentral.com/articles/10.1186/s13750-016-0059-6
Campbell systematic reviews: Policies and Guidelines (social policy)	Campbell Collaboration	https://onlinelibrary.wiley.com/pb-assets/Campbell%20Policies%20and%20Guidelines%20Dec2020-1608292090217.pdf
Guidance for producing a Campbell evidence and gap map	Campbell Collaboration	https://onlinelibrary.wiley.com/doi/10.1002/cl2.1125
Cochrane Handbook for Systematic Reviews of Interventions (health)	Cochrane	https://training.cochrane.org/handbook
ROSES RepOrting standards for Systematic Evidence Syntheses: pro forma, flow-diagram and descriptive summary of the plan and conduct of environmental systematic reviews and systematic maps	Collaboration for Environmental Evidence	https://environmentalevidencejournal.biomedcentral.com/articles/10.1186/s13750-018-0121-7
The PRISMA 2020 statement: an updated guideline for reporting systematic reviews	PRISMA	https://www.bmj.com/content/372/bmj.n71
Methodological Expectations of Cochrane Intervention Reviews (MECIR)	Cochrane	https://methods.cochrane.org/methodological-expectations-cochrane-intervention-reviews
Methodological Expectations of Campbell Collaboration Intervention Reviews (MECCIR)	Campbell Collaboration	https://www.campbellcollaboration.org/about-meccir.html

5. *Endorsing or enforcing certain tools and technologies, for example systematic review management tools (e.g. EPPI Reviewer, CADIMA)*

According to our interviews with project coordinators, two BONUS synthesis projects made use of tools specifically designed to support evidence synthesis conduct; EPPI-Reviewer (Thomas and Brunton 2007) and CADIMA (Kohl et al. 2018). Other projects, however, were unaware of such tools, nor the frameworks for conduct and reporting of evidence syntheses. The teams that made use of the review management tools reported that they found them very easy to use and that they greatly aided their project management.

This highlights that there would be use in 'endorsing' (i.e. recommending) certain tools that might help project teams to manage their reviews, increase transparency via efficient data recording, and maximise rigour. Since systematic review management tools are designed specifically to follow robust best practice, their recommendation could also build capacity and increase the quality of review conduct and reporting. Review management tools allow the user to upload and document their searches across multiple platforms, remove deduplicate records, conduct eligibility screening keeping track of all decisions, assess consistency between multiple reviewers, assign tasks across a review team, upload and read full text PDF documents, extract data and code studies, track the fate of each record included in the review, export record lists for reporting of review activities, produce visualisations summarising the methods and findings. Examples of systematic review management tools include:

- EPPI-Reviewer – a very powerful, low-cost web-based tool with excellent customer support developed and maintained by the EPPI-Centre at University College London (https://eppi.ioe.ac.uk/cms/Default.aspx?alias=eppi.ioe.ac.uk/cms/er4)
- SysRev – a freemium web-based tool with excellent customer support (https://sysrev.com/)
- CADIMA – a free web-based tool built around the CEE guidelines for systematic reviews in environmental management (https://www.cadima.info/)
- Rayyan – a free web-based tool with a complementary mobile app focusing mainly on record management and screening (but not latter review stages) (https://www.rayyan.ai/)

Several of these tools (e.g. EPPI-Reviewer and Rayyan) now also integrate machine learning to increase screening efficiency (Tsou et al. 2020): users assess records as relevant or not, and an algorithm continually attempts to score records for relevance based on the words present in their abstracts, with abstracts then ordered according to their perceived relevance (Bannach-Brown et al. 2019). In some cases, this technology has been shown to drastically reduce the time needed to perform screening (Marshall and Wallace 2019), although there are reservations over the risk of omitting relevant but divergently described studies.

In addition there are a suite of other tools and technologies that can increase efficiency, transparency and rigour. The Systematic Review Toolbox provides a descriptive searchable database of tools; http://systematicreviewtools.com/.

6. *Requiring minimum standards in evidence synthesis and stakeholder engagement (e.g. publishing a protocol)*

Several organisations provide guidance and standards in evidence synthesis methods (as outlined in Table 3). In most cases, these also constitute minimum standards that should be followed to ensure that reviews are reliable and fit-for-purpose in decision-making. Reviews conducted to these high-standards have been demonstrated to be of greater validity and reliability than other reviews (Woodcock et al. 2017).

Where systematic reviews and maps are the focus of future synthesis calls, it would be advisable to ensure that the systematic reviews and maps are conducted according to the guidance and minimum standards established by the Collaboration for Environmental Evidence (CEE 2018). At a minimum, this would need the guidance to be specified in the call text. There are several ways in which this could be facilitated:

- A central group of experts (e.g. external methodology and subject experts) assesses the methodology proposed in the funding applications, during the conduct of the reviews, and at the end of the projects, prior to publication of the review manuscripts.
- Methodological experts could be commissioned to peer review applications with explicit evaluation criteria relating to the planned synthesis methodology, ongoing projects, and review manuscripts at the end of the project, prior to publication.
- Project teams could be required to publish their reviews in a specialist venue dedicated to peer-reviewing and publishing systematic reviews and maps, for example, the Collaboration for Environmental Evidence journal (*Environmental Evidence*).

One key standard for evidence syntheses is that they should write and publish an *a priori* plan for their synthesis methods, also referred to as a *protocol*. This document is finalised before work commences, giving an important opportunity for feedback to improve methodology and avoid or limit possible bias and other limitations (CEE 2018; Higgins et al. 2019). Typically, these protocols are designed closely with stakeholders to ensure relevance and acceptance, and also foster a sense of inclusion and shared ownership (Haddaway et al. 2017). These protocols could be reviewed in the same way the final manuscript is reviewed, but importantly giving an opportunity to ensure methodological rigour.

Similarly, projects could be explicitly expected to conduct meaningful stakeholder engagement during the planning, conduct and communication of their reviews. Interviews conducted here with project managers indicate that stakeholder engagement in the BONUS synthesis call was often very much dependent on the project teams' existing networks, which may have resulted in variable communication and impact. Ensuring reviewers follow particular guidance in stakeholder engagement for their reviews (e.g. CEE 2018; James et al. 2016) may help to increase impact.

7. *Partnering with an evidence synthesis organisation for support in drafting the call, peer-reviewing applications, protocols and final reports, and for advice/training for funded project teams*

As mentioned above, the rigour of systematic reviews conducted as part of a future synthesis call could be assessed at multiple stages in cooperation with individuals or an organisation with expertise in evidence synthesis methodology. Applications, protocols and final review papers could be peer-reviewed to ensure adequate and appropriate reporting. If this happens from an early stage, the rigour of the projects could be substantially improved.

Furthermore, such a partnership could involve one or more training or mentoring workshops, providing a highly applied capacity building opportunity across teams. Such an event could also support sharing of experiences across funding project teams.

8. *Evaluating the project team's prior evidence synthesis experience as part of the application review*

One aspect that appears not to have been evaluated as part of the BONUS Synthesis call was either the rigorousness of the planned evidence synthesis methodology, or the prior experience of the applicants

in synthesis methods. If systematic reviews are to be requested in future synthesis calls, evidence synthesis expertise and experience should be assessed formally in applicants curricula vitae.

9. *Evaluating funded project teams' progress with obligatory milestones: e.g. protocol at month 2*

Another opportunity to increase the likely success of the funded projects would be to require all projects to submit an *a priori* protocol (also referred to in primary research as a registered report) that outlines their planned methods in detail at the start of the project. This would allow for feedback on the methods they intend to use and support communication and stakeholder engagement. Systematic review and map protocols are increasingly published in environmental journals such as Environmental Evidence (https://environmentalevidencejournal.biomedcentral.com/), but registered reports that outlined planned methods for primary research are also becoming increasingly recognised for their benefits to rigour (Soderberg et al. 2021).

10. *Supporting/encouraging cross fertilisation across projects*

Future synthesis calls could consider formalising opportunities for cross fertilisation across projects in a number of ways. As mentioned above, methodological training events and workshops could be scheduled for project teams involved in evidence synthesis or indeed other methods. These events might not only build capacity, but also provide an opportunity for partners involved in the conduct activities to share experiences, *tips and tricks*, and tools and software that can increase efficiency and accessibility, such as some of the review management tools mentioned above. Although some projects informally cooperated during the BONUS Synthesis projects, this was done on an ad hoc basis, primarily for the projects that shared scope. This cross fertilisation would also support shared stakeholder engagement planning and activities, reducing the burden on stakeholders and widening the contactable networks across all projects with overlapping end users.

11. *Consider centralising stakeholder engagement (including communication); mandatory communication requirements for a dissemination strategy*

On a related note, future Synthesis funding calls could consider centralising stakeholder engagement in a number of ways. For example, identifying, networking with and communicating results to relevant Baltic and North Sea stakeholders could be taken on as a role for the BANOS Secretariat in close collaboration with project teams. This could reduce the burden on stakeholders who are end users of multiple projects and avoid reliance on project teams' existing stakeholder networks. A key finding from interviews with BONUS Synthesis project coordinators related to communication of their findings: several interviewees mentioned that budgets and timelines did not allow much time for communication, and that the national communication was often dependent on their partners' existing networks.

A different approach could involve the Secretariat supporting the coordination of each project's stakeholder engagement and communication, ensuring consistency in approaches and availability of support for identifying and communicating with national and regional stakeholders. For example, the BANOS Secretariat could formalise contact with regional stakeholders like HELCOM, arranging shared communication channels across all projects, removing the need for each project to make contacts independently (with the exception of a small number of projects who arranged a collaborative communication with HELCOM, as reported in one interview).

12. *Support longer term communication through extended project timelines and tapered funding (i.e. extra time with minimum budget to continue communicating, if communication not centralised)*

As mentioned above, communication of the project findings was mentioned by several interviewees to be challenged by the short duration of the projects and the limited funding: in one case a project could not use the allotted open access funds because the article processing fee invoice was not received within the chargeability period. Future synthesis calls could include a tapering of project funding over a longer project period, with a longer deadline for ongoing communication deliverables.

Conclusions

This section summarises the findings of our evaluations and recommendations for future synthesis calls.

Synthesis of primary research has a vital role to play in supporting evidence-informed decision making in research, policy and practice. This report has outlined a set of practical recommendations for appropriately designing future calls for evidence syntheses to support decision making in the Baltic and North Sea regions. The report highlights different evidence synthesis frameworks that can support the production of rigorous evidence, providing a range of methods depending on the desired objectives (understanding an evidence base versus understanding impacts and effectiveness), along with practical constraints, such as time, resources and acceptable level of risk in subsequent decision making.

The report also outlines methods for rigorous, meaningful and effective engagement of stakeholders in evidence syntheses to support and facilitate effective communication and update of synthesis findings. It also outlines options available for designing and coordinating calls, learning from a variety of existing frameworks. No single call framework has been recommended here, since a suite of external factors, including legal issues, will affect the suitability of each.

A key issue relating to evidence synthesis is ambiguity in the appreciation and understanding of the methods and associate terminology, which is understandably varied and confusing. Future synthesis calls could reduce ambiguity by explicitly stating what methods are desired for each call/task, and the tables and descriptions of different evidence synthesis methods and their central tenets provided here could prove useful in selecting either a formal evidence synthesis method or a similar alternative.

Where evidence synthesis is appropriate, however, future calls could improve rigour, transparency and reliability in decision-making by encouraging and ensuring funded projects follow internationally accepted best practice guidance and standards for the conduct of evidence syntheses. Where these standards are not followed, researchers should be encouraged to explicitly describe why they have chosen not to follow these standards. Ensuring that these standards are efficiently and successfully built into funding texts, application review and project oversight can be greatly facilitated by engaging with methodology experts from the community at large. Furthermore, by more closely including experts in synthesis methodology from the outset of the call design, the rigour of the outputs and capacity of the funded teams can be maximised.

Finally, by centralising and extending support for communication activities across multiple funded projects, the impacts and legacy of the syntheses could be maximised.

References main report

Bannach-Brown, A., Przybyła, P., Thomas, J., Rice, A.S., Ananiadou, S., Liao, J. and Macleod, M.R., 2019. Machine learning algorithms for systematic review: reducing workload in a preclinical review of animal studies and reducing human screening error. Systematic reviews, 8(1), pp.1-12.

BONUS call 2017: Synthesis: BONUS call 2017: Synthesis - Bonus EEIG (bonusportal.org) [04.10.21]

Bunn F, Trivedi D, Alderson P, Hamilton L, Martin A, Pinkney E, Iliffe S. The impact of Cochrane Reviews: a mixed-methods evaluation of outputs from Cochrane Review Groups supported by the National Institute for Health Research. Health Technol Assess. 2015 Apr;19(28):1-99, v-vi. doi: 10.3310/hta19280. PMID: 25875129; PMCID: PMC4781054

Collaboration for Environmental Evidence (2018). Guidelines and Standards for Evidence synthesis in Environmental Management. Version 5.0 (Pullin AS, Frampton GK, Livoreil B and Petrokofsky, Eds). www.environmentalevidence.org/information-for-authors. [14.09.21]

Dicks LV, 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

Djulbegovic B, Guyatt GH, 2017. Progress in evidence-based medicine: a quarter century on, The Lancet, Volume 390, Issue 10092, Pages 415-423, https://doi.org/10.1016/S0140-6736(16)31592-6.

Eales, J., Haddaway, N.R. and Webb, J.A., 2017. Much at stake: the importance of training and capacity building for stakeholder engagement in evidence synthesis. Environmental Evidence, 6(1), pp.1-8.

Gough, D., Oliver, S. and Thomas, J. eds., 2017. An introduction to systematic reviews. Sage.

Gurevitch, J., Koricheva, J., Nakagawa, S. and Stewart, G., 2018. Meta-analysis and the science of research synthesis. Nature, 555(7695), pp.175-182.

Haddaway, N.R. and Bayliss, H.R., 2015. Shades of grey: two forms of grey literature important for reviews in conservation. Biological Conservation, 191, pp.827-829.

Haddaway, N.R., Bernes, C., Jonsson, B.G. and Hedlund, K., 2016. The benefits of systematic mapping to evidence-based environmental management. Ambio, 45(5), pp.613-620.

Haddaway, N.R. and Bilotta, G.S., 2016. Systematic reviews: Separating fact from fiction. Environment international, 92, pp.578-584.

Haddaway, N.R., Crowe, S. 2018 Experiences and lessons in stakeholder engagement in environmental evidence synthesis: a truly special series. *Environ Evid* **7,** 11. https://doi.org/10.1186/s13750-018-0123-5

Haddaway, N.R., Kohl, C., Rebelo da Silva, N. et al. 2017 A framework for stakeholder engagement during systematic reviews and maps in environmental management. Environ Evid 6, 11. https://doi.org/10.1186/s13750-017-0089-8

Haddaway, N.R. and Pullin, A.S., 2014. The policy role of systematic reviews: past, present and future. Springer Science Reviews, 2(1-2), pp.179-183.

Haddaway, N.R. and Westgate, M.J., 2019. Predicting the time needed for environmental systematic reviews and systematic maps. Conservation Biology, 33(2), pp.434-443.

Hahn DL, Hoffmann AE, Felzien M, LeMaster JW, Xu J, Fagnan LJ, 2017 Tokenism in patient engagement, Family Practice, Volume 34, Issue 3, Pages 290–295, https://doi.org/10.1093/fampra/cmw097

Head, B.W. (2016), Toward More "Evidence-Informed" Policy Making?. Public Admin Rev, 76: 472-484. https://doi.org/10.1111/puar.12475

Higgins, J.P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J. and Welch, V.A. eds., 2019. Cochrane handbook for systematic reviews of interventions. John Wiley & Sons.

Jackson, N. and Waters, E., 2005. Criteria for the systematic review of health promotion and public health interventions. Health promotion international, 20(4), pp.367-374.

James, K.L., Randall, N.P. & Haddaway, N.R. 2016 A methodology for systematic mapping in environmental sciences. Environ Evid 5, 7. https://doi.org/10.1186/s13750-016-0059-6

Kohl, C., Frampton, G., Sweet, J., Spök, A., Haddaway, N.R., Wilhelm, R., Unger, S. and Schiemann, J., 2015. Can systematic reviews inform GMO risk assessment and risk management?. Frontiers in bioengineering and biotechnology, 3, p.113.

Kohl, C., McIntosh, E.J., Unger, S., Haddaway, N.R., Kecke, S., Schiemann, J. and Wilhelm, R., 2018. Online tools supporting the conduct and reporting of systematic reviews and systematic maps: a case study on CADIMA and review of existing tools. Environmental evidence, 7(1), pp.1-17.

Larsen, R.K., Nilsson, A.E. 2017 Knowledge production and environmental conflict: managing systematic reviews and maps for constructive outcomes. Environ Evid 6, 17. https://doi.org/10.1186/s13750-017-0095-x

Land, M., Macura, B., Bernes, C. and Johansson, S., 2017. A five-step approach for stakeholder engagement in prioritisation and planning of environmental evidence syntheses. Environmental Evidence, 6(1), pp.1-7.

Lescroart, J., F. De Raedemaecker, H. Pirlet, A-K. Lescrauwaet, M. Sirola 2019. A holistic map of the programme's stakeholders. BANOS CSA/D3.2.

Littell JH, Shlonsky A. 2010 Toward Evidence-Informed Policy and Practice in Child Welfare. Research on Social Work Practice 20(6):723-725. doi:10.1177/1049731509347886

Macura, B., Haddaway, N.R., Lesser, P. et al. 2019 Mapping the predicted and potential impacts of metal mining and its mitigation measures in Arctic and boreal regions using environmental and social impact assessments: a systematic map protocol. Environ Evid 8, 21. https://doi.org/10.1186/s13750-019-0166-2

Marshall, I.J., Wallace, B.C.2019 Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. Syst Rev 8, 163. https://doi.org/10.1186/s13643-019-1074-9

Moat KA, Lavis, JN 1013 10 best resources for … evidence-informed health policy making. Health Policy and Planning, Volume 28, Issue 2, March 2013, Pages 215–218, https://doi.org/10.1093/heapol/czs050

Morgan, R.L., Whaley, P., Thayer, K.A. and Schünemann, H.J., 2018. Identifying the PECO: a framework for formulating good questions to explore the association of environmental and other exposures with health outcomes. Environment international, 121(Pt 1), p.1027.

Polasky S, Carpenter SR, Folke C, Keeler B, 2011. Decision-making under great uncertainty: environmental management in an era of global change, Trends in Ecology & Evolution, Volume 26, Issue 8, Pages 398-404, https://doi.org/10.1016/j.tree.2011.04.007

Reed, M.S. and Curzon, R., 2015. Stakeholder mapping for the governance of biosecurity: a literature review. Journal of Integrative Environmental Sciences, 12(1), pp.15-38.

Selph, S.S., Ginsburg, A.D. & Chou, R. 2014 Impact of contacting study authors to obtain additional data for systematic reviews: diagnostic accuracy studies for hepatic fibrosis. Syst Rev 3, 107. https://doi.org/10.1186/2046-4053-3-107

Shlonsky, A., Noonan, E., Littell, J.H. et al. 2011 The Role of Systematic Reviews and the Campbell Collaboration in the Realization of Evidence-Informed Practice. Clin Soc Work J 39, 362–368. https://doi.org/10.1007/s10615-010-0307-0

Snoeijs-Leijonmalm P, Barnard S, Elliott M, Andrusaitis A, Kononen K, Sirola M, 2017, Towards better integration of environmental science in society: Lessons from BONUS, the joint Baltic Sea environmental research and development programme, Environmental Science & Policy, Volume 78, 2017, Pages 193-209

Soderberg, C.K., Errington, T.M., Schiavone, S.R., Bottesini, J., Thorn, F.S., Vazire, S., Esterling, K.M. and Nosek, B.A., 2021. Initial evidence of research quality of registered reports compared with the standard publishing model. Nature Human Behaviour, 5(8), pp.990-997.

Spök, A., Racovita, M., Karner, S. et al. 2018 Inclusive development and prioritization of review questions in a highly controversial field of regulatory science. Environ Evid 7, 1. https://doi.org/10.1186/s13750-017-0113-z

Stewart, G.B., Coles, C.F. and Pullin, A.S., 2005. Applying evidence-based practice in conservation management: lessons from the first systematic review and dissemination projects. Biological Conservation, 126(2), pp.270-278.

Storie J; Suškevis M; Külvik M; Lehtoranta V; Vikström S; Riikonen S; Kuosa H; Kuhn K, Oinonen S. (2020) What evidence exists for the impact of Baltic Sea ecosystems on human health and well-being? A systematic map protocol. Environ Evid 9, 5 (2020). https://doi.org/10.1186/s13750-020-00189-6

Sundin, A., Andersson, K. and Watt, R., 2018. Rethinking communication: integrating storytelling for increased stakeholder engagement in environmental evidence synthesis. Environmental Evidence, 7(1), pp.1-6.

Taylor, J.J., Rytwinski, T., Bennett, J.R. et al. 2017 Lessons for introducing stakeholders to environmental evidence synthesis. Environ Evid 6, 26. https://doi.org/10.1186/s13750-017-0105-z

Thomas and Brunton 2007 https://discovery.ucl.ac.uk/id/eprint/10003909/1/Thomas2007EPPI-Reviewer%28Poster%29.pdf

Tsou, A.Y., Treadwell, J.R., Erinoff, E. and Schoelles, K., 2020. Machine learning for screening prioritization in systematic reviews: comparative performance of Abstrackr and EPPI-Reviewer. Systematic reviews, 9(1), pp.1-14.

Rooney AA, Boyles AL, Wolfe MS, Bucher JR, and Thayer KA 2014 Systematic Review and Evidence Integration for Literature-Based Environmental Health Science Assessments Environmental Health Perspectives 122:7 CID: https://doi.org/10.1289/ehp.1307972

White, H. (2019). The twenty-first century experimenting society: the four waves of the evidence revolution. Palgrave Commun 5, 47. https://doi.org/10.1057/s41599-019-0253-6

Woodcock, P., O'Leary, B.C., Kaiser, M.J. and Pullin, A.S., 2017. Your evidence or mine? Systematic evaluation of reviews of marine protected area effectiveness. Fish and Fisheries, 18(4), pp.668-681.

Annex A: The 21 forms of knowledge synthesis – a spectrum of framing, rigour, external validity and resource requirements

This annex aims to explain the meaning of the terms 'knowledge synthesis' and 'evidence synthesis' in more detail, and identify and categorise the main types of evidence synthesis methodology. The overarching purpose is to support decision-making around selection of evidence synthesis methods by clearly explaining the features, suitability, advantages and disadvantages of a suite of alternative methods for identifying, collating, processing, appraising and combining research data for particular topics.

This addendum is based on two other existing reports that have aimed to describe synthesis methods in a transparent and systematic way. The first of these is a report from the Methods Expert Group of the EU-funded EKLIPSE project (http://www.eklipse-mechanism.eu/) on 21 knowledge synthesis methods. The second report is the Defra/NERC-funded WT1552 produced by the Centre for Evidence-Based Conservation at Bangor University in collaboration with the Joint Water Evidence Group at Defra (http://sciencesearch.defra.gov.uk/Document.aspx?Document=13630_WT1552ToolsandTechniquesforEvidenceReviews-FinalReport.pdf). This document combines the findings of these reports, building on their work, taking into account also other relevant literature.

1.1. The methods

The 21 methods identified by the EKLIPSE authors are detailed below and summarised in Table 1. These methods overlap in their purpose and share similar characteristics, but have been identified to cover a broad spectrum of methodologies.

Some of the methods included are viewed by the authors as not being suitable standalone methods. The authors recognised that previous work has included some methods that should not have been included (e.g. meta-analysis and vote-counting), but include them in their assessment for completeness.

The level of participation and risk of bias are a relative ranking, and the assessment is subjective because of the lack of universal definitions that work across all methods.

In general, low resource cost, high participation and low risk of bias are assumed to be the most important factors in selecting knowledge synthesis methods, but this should be judged on a case-by-case basis.

Table 1. Definitions of knowledge synthesis methods described by the EKLIPSE Methods Expert Group.

Method	**Definition**
Systematic Review	A structured, step-wise methodology following an a priori protocol to comprehensively collate, critically appraise and synthesise existing research evidence (traditional academic and grey literature). Systematic reviews should be conducted according to the rigorous standards demanded by review coordinating bodies such as the Cochrane Collaboration, the Collaboration for Environmental Evidence and the Campbell Collaboration, as well as the ROSES reporting standards; tools such as PredicTER enable the calculation of the time needed to conduct systematic reviews (see references). Reporting requirements include: protocol of methods; fates of all articles screened at full text; transparent documenting of all methods used. The method includes tertiary reviews, or systematic reviews of reviews.
Solutions Scanning	A structured, step-wise methodology to identify a long list of available actions, interventions or approaches, in response to a broad challenge. A list is gathered through consultation with a wide range of stakeholders, and continues to be circulated through networks until five new people have seen it and add nothing.
Subject-wide Evidence Syntheses (Summaries and Synopses)	Flexible, transparent approach to collate and summarise existing research evidence over a broad topic in a standard format. Interventions, actions, or impacts are first listed (can use a process of Solution Scanning). Review methods are flexible and pragmatic, selecting and reporting the best available search methodology, with a focus on existing systematic reviews and systematic maps where possible.
Meta-Analysis	A statistical tool to reanalyse existing data from multiple studies. Meta-analysis is not an independent type of review. It relies on data extracted from an existing set of studies resulting from a review.
Rapid Evidence Assessment	A structured, step-wise methodology, usually following an a priori protocol to comprehensively collate, critically appraise and synthesise existing research evidence (traditional academic and grey literature), following systematic review methodology but with components of the process simplified or omitted to produce information in a short period of time. The method is sometimes called 'rapid review' (Tricco et al. 2015). The exact set of methods used, or the components of systematic review omitted, are flexible, and the method itself is not well defined internationally. A standardised version of Rapid Evidence Assessment has been defined by the UK Government (Collins et al. 2014).
Scoping Review	A structured, step-wise methodology, preferably following an a priori protocol to collate and describe existing research evidence (traditional academic and grey literature) in a broad topic area, following a systematic map methodology but with components of the process simplified or omitted to produce information in a short period of time. This is not the same as the scoping stage of a systematic review. The method has been called 'Quick Scoping Review' (Collins et al. 2015). The exact set of methods used, or the components of systematic map that are left out is flexible, and the method itself is not standardised internationally. A standardised version of Quick Scoping Review has been defined by the UK Government (Collins et al. 2014).
Systematic Map	Structured, step-wise methodology following an a priori protocol to comprehensively collate and describe existing research evidence (traditional academic and grey literature). Systematic maps should be conducted according to the rigorous standards demanded by review coordinating bodies such as the Collaboration for Environmental Evidence12 and the Social Care Institute for Excellence SCIE13, as well as ROSES reporting standards (Haddaway et al. 2018). Reporting requirements include: protocol of methods, fates of all articles screened at full text, transparent documenting of all methods used. Tools such as PredicTER enable the calculation of the time needed to conduct systematic maps (Haddaway and Westgate 2019).
Vote-Counting	A simple tool used to synthesise findings from multiple studies, by counting the numbers of studies finding positive and negative results. This method is based only on the direction and sometimes significance of the result, and does not critically appraise or differentially weight the studies. Vote counting is limited to answering the question "is there any evidence of an effect?".
Non-Systematic Literature Review	Literature review that describes (and may appraise) the state/nature of existing evidence, but does not follow a standardised, systematic method.
Expert Consultation	The consultation of a designated set of experts, either individually or in a group, to gather judgement, evaluation or opinion. This can use online consultation, in-person meetings, individual interviews, written consultation or group meetings.
Multiple Expert Consultation with Formal Consensus Method such as Delphi	This method is a subset of expert consultation, representing the most rigorous approach to eliciting expert knowledge. It combines the knowledge of multiple, carefully selected experts into either quantitative or qualitative assessments, using formal consensus methods such as the Delphi process (described and reviewed by Mukherjee et al. 2016), or other elicitation techniques, including Cooke's method of weighting experts for their accuracy, described in Martin et al. (2012). Such approaches have been empirically demonstrated to generate estimates for ecological parameters that are more accurate than the estimates of the best-regarded expert in the group (Burgman et al. 2011).

Causal Criteria Analysis (typically a combination of methods)	Causal Criteria Analysis synthesizes understanding of causal linkages in a system, by testing against a set of pre-defined criteria for causality. It combines pictorial relationships between factors depicting hypothesized or known causal linkages in a system, with literature review to synthesize evidence for specific links in the chain. The diagrams (called influence diagrams, if they include management actions or policy options) are used as scaffolds to synthesize and present evidence. They can also serve as a first step to more elaborate modelling approaches. The review stage preferably employs the systematic review or rapid evidence assessment method, in which studies are critically appraised and weighted. It could also employ expert consultation, using formal consensus method such as Delphi, or a Bayesian Belief Network approach to elicit knowledge.
Bayesian Belief Networks (typically a combination of methods)	A semi-quantitative modelling approach that combines empirical data with expert knowledge to calculate the probability of a specific outcome or set of outcomes. Similar to the Causal Criteria Analysis, the method first builds a visual representation of the system. Probabilities for each link can be based on expert judgement, literature review, or a prescribed mechanistic model. The BBN model can then generate a range of probabilities for the final outcome, based on the underlying system. The main output is a diagrammatic interpretation of a system showing probabilistic relationships and outcomes within a causal chain. This method explicitly incorporates uncertainty about linkages in a causal chain via conditional probabilities. For example, a BBN could quantify likelihood of storm events large enough to impact coastal ecosystems.
Focus Groups	Structured discussion of an issue by a small group six to ten of people, led by a skilled moderator. The group is purposively selected usually to involve different stakeholders and/or potentially differing perspectives. The joint discussion allows participants to consider and react to arguments put forward by other participants so it allows examination of group dynamics and opinion formation. Focus groups are regarded as an appropriate method for evaluating attitudes, knowledge and experiences, although features of the focus group method should be reported to allow better interpretation of results (Orvik et al. 2013). Focus groups can also be used to gather information form a specific group, to build scenarios in a choice experiment method for instance, or test questions or issues for a quantitative survey.
Discourse Analysis	Discourse analysis is a structured method for investigating conflicts and alliances among different knowledge holders or stocks of knowledge when discourses are emerging. The aim is to identify the key issues and actors, distinguish between certain and uncertain knowledge, and determine which knowledge claims are points of conflict between different groups in society and the sciences. The focus is on arguments, procedures or putative facts that are seen as correct or true by the actors under analysis, rather than on whether they are true. Discourse analysis can therefore reveal why a particular understanding of a given environmental problem at some point gains dominance and is seen as authoritative, while other understandings are discredited.
Joint Fact Finding (JFF)	Joint Fact-Finding is a process in which separate coalitions of scientists, policy-makers and other stakeholders with differing viewpoints and interests work together to develop data and information, analyse facts and forecasts, and develop common assumptions and informed opinions (van Buuren et al. 2007). Finally, they can use the information they have developed to reach decisions together. A comparatively small group can be involved, but all opposing positions need to be represented.
Scenario Analysis	Scenario Analysis formulates assumptions about future developments in one connected storyline. Scenarios are consistent and coherent descriptions of alternative hypothetical futures that reflect different perspectives on past, present, and future developments. Qualitative storylines for the future development of complex systems can be integrated with quantitative modelling. “Scenarios and models play complementary roles, with scenarios describing possible futures for drivers of change or policy interventions and models translating those scenarios into projected consequences for nature and nature’s benefits to people.” IPBES (2016). Scenarios are more likely to lead to read policy outcomes if they use participatory approaches to involve stakeholders throughout, from the initial phase of problem definition and feature frequent exchanges between scientists and stakeholders. Participatory scenario development aims to supplement and synthesize existing data and formalized knowledge with other relevant forms of stakeholder knowledge.
Structured Decision Making	Structured Decision Making (SDM) is a well-defined method for analyzing a decision by breaking it into components including the objectives, possible actions, and models linking actions to objectives. It aims to compare possible actions in terms of one or more objectives. It provides transparency by specifying each of these components and providing information that a decisionmaker can use to implement and defend a decision. This method can incorporate other knowledge synthesis methods. Expert consultation with elicitation is often used to quantify predictive relationships as part of SDM. SDM is founded on principles of value-focused thinking and decision analysis and can be conducted in a participatory manner with decision-makers, stakeholders, and experts. It can also provide a basis for adaptive management. Structured Decision Making typically involves a series of iterative steps called PrOACT (Problem framing, Objectives, Actions, Consequences, and Tradeoffs).
Collaborative Adaptive Management (typically a combination of methods)	Collaborative Adaptive Management (CAM) is a structured/flexible, stepwise, transparent approach that includes the iteration of knowledge synthesis, most often using collaborative methodologies, such as participatory scenario building, joint fact-finding and/or multi-criteria analysis. New knowledge is then generated, through the selection, application and monitoring of policies or management strategies. CAM differs from other knowledge synthesis methods in a key aspect. Instead of aiming to identify single, broadly-applicable, optimal solutions, it aims to identify flexible solutions that are resilient to errors and uncertainty. The initial phase of CAM represents a specific type of knowledge synthesis, but the overall approach goes beyond synthesis to locally or specifically relevant knowledge generation.

Participatory Mapping	Participatory mapping defines a set of approaches and techniques that combine the tools of modern cartography with participatory methods to represent the spatial knowledge of local communities. It is based on the premise that local inhabitants possess expert knowledge of their local environments that can be expressed in a geographical framework, which is easily understandable and universally recognised.
Multi Criteria Decision Analysis	Multi-Criteria Decision Analysis (MCDA) evaluates the performance of alternative courses of action with respect to criteria that capture the key dimensions of the decision-making problem, involving human judgment and preferences (Belton and Stewart 2002).

1.2. Factors considered

The EKLIPSE Methods Expert Group classified all 21 methods according to the following criteria:

Cost – a relatively comprehensive assessment in non-monetary terms of the requirements of the method. Staff time is estimated in full time equivalent person months (typically as a range in months), whilst other additional costs are described (e.g. subscription fees, software, meetings etc.). EKLIPSE also describes for some methods where costs are significantly affected by other factors.

Time required – an estimate of the time window needed for completion of the method (typically a range in months). EKLIPSE also describes for some methods where the time requirement is significantly affected by other factors.

Repeatability – a relative qualitative appraisal of how feasible it is to replicate the methods used for a typical project. Categorised as low/moderate/high, with brief comments.

Transparency - a relative qualitative appraisal of the level of transparency in the conduct and reporting of the methods for a typical project. Categorised as low/ moderate /high, with brief comments.

Risk of bias – a relative qualitative appraisal of the potential risk of bias inherent in the method, with brief qualifying explanations. Categorised as low/ moderate /high, with brief comments.

Scale (or level of detail) – a description of the geographical scale and/or level of detail obtainable in the answer to a question, with brief comments.

Capacity for participation – a relative qualitative appraisal of the potential for stakeholders to participate in the method. Categorised as low/ moderate /high, with brief comments.

Data demand – a relative qualitative appraisal of the requirements for existing data. Categorised as low/ moderate /high, with brief comments.

Types of knowledge that can be synthesized – a brief summary of the knowledge systems that can be assessed with the method. Knowledge system; scientific, technical, opinion, Indigenous and local knowledge: knowledge type; explicit, tacit. Categorical statement with brief comments.

Types of output that can be produced – a description of the type of documents and/or outputs that can be/are typically produced from the method.

Specific expertise required – a brief description of the requirements of the method for subject/methodology expert involvement.

They also classified each method as being exploratory, stakeholder engagement-centric, analytical, and evaluation, defined as follows:

Exploratory methods review and collate evidence in a more or less intense way depending on time and resources availability. Relevant policy cycle phase: agenda-setting.

Engagement methods collect stakeholder opinions and the expert contributions through focus groups, expert consultations or Delphi processes. This group of methods can potentially be applied throughout the policy cycle, including implementation and evaluation, but they are more often indicated for the earlier phases in the cycle. Relevant policy cycle phases: all.

Analytical methods use knowledge synthesis methods to analyse and compare possible courses of action, for example through scenario building, collaborative adaptive management or multi-criteria analysis. Relevant policy cycle phase: policy formulation.

Evaluation methods. For analysing policy impacts, the set of methodologies with predictive power, such as scenario building, are useful. To systematically evaluate policy structures, acceptance and narratives the most indicated KSM were: structured decision making, joint fact finding and discourse analysis. Relevant policy cycle phase: policy evaluation.

Some of this information is summarised in Table 2, and further details can be found in the individual methods guidance notes (http://www.eklipse-mechanism.eu/expert_group_on_methods#guidance%20notes). Each knowledge synthesis method has also been classified as to whether it also belongs to a subset of evidence synthesis methods.

Table 2. Summary of knowledge synthesis methods detailed in the EKLIPSE Methods Expert Group report. 'Submethod' signifies methods that combine a listed evidence synthesis method with other knowledge synthesis methods.

Method	Code	Time and resource requirement	Capacity for participation	Risk of bias	Explore	Engage	Analyse	Evaluate	Evidence synthesis method
Systematic Review	SystR	High	Medium	Low	x		x	x	Yes
Solutions Scanning	SolS	Low	High	Medium	x				No
Subject-wide Evidence Syntheses (Summaries and Synopses)	Sum	High	Medium	Low	x				Yes
Meta-Analysis (not a standalone method)	MA	Low	Low	Low	x		x		Yes
Rapid Evidence Assessment	REA	Medium	Medium	Medium	x		x	x	Yes
Scoping Review	ScopR	Medium	Medium	Medium	x				Yes
Systematic Map	SM	High	Medium	Low	x		x	x	Yes
Vote-Counting	VC	Low	Low	High	x				Yes
Non-Systematic Literature Review	NSystR	Medium	Low	High	x		x	x	Yes
Expert Consultation	ExC	Low	Medium	High		x		x	No
Multiple Expert Consultation with Formal Consensus Method such as Delphi	ECwD	Low	Medium	Medium	x	x	x	x	No
Causal Criteria Analysis (typically a combination of methods)	CCA	Low	Medium	Medium	x	x			Submethod
Bayesian Belief Networks (typically a combination of methods)	BBN	Medium	Medium	Medium	x	x	x		Submethod
Focus Groups	FG	Low	Medium	High		x		x	No
Discourse Analysis	DA	Medium	Low	Medium				x	No
Joint Fact Finding (JFF)	JFF	Medium	Medium	High		x		x	No
Scenario Analysis	Scen	Low	High	Medium	x	x	x	x	No
Structured Decision Making	SDM	Medium	High	Medium			x	x	No
Collaborative Adaptive Management (typically a combination of methods)	CAM	High	High	Low	x	x	x	x	No
Participatory Mapping	PM	Medium	High	Medium	x	x		x	No
Multi Criteria Decision Analysis	MCDA	Medium	High	Medium			x		No

1.3. Visualisations

The ELKIPSE Methods Expert Group selected what it viewed to be the three most important criteria relating to each knowledge synthesis method for decision-makers with knowledge needs: the time and resources required; the capacity for participation; and the risk of bias.

Each knowledge synthesis method has been visualised across a spectrum individually for each of these criteria in Figure 1. Figure 2 then plots each method across two axes; firstly, for time and resources versus risk of bias; and secondly, for capacity for participation versus risk of bias, highlighting those methods that are also *evidence synthesis methods*.

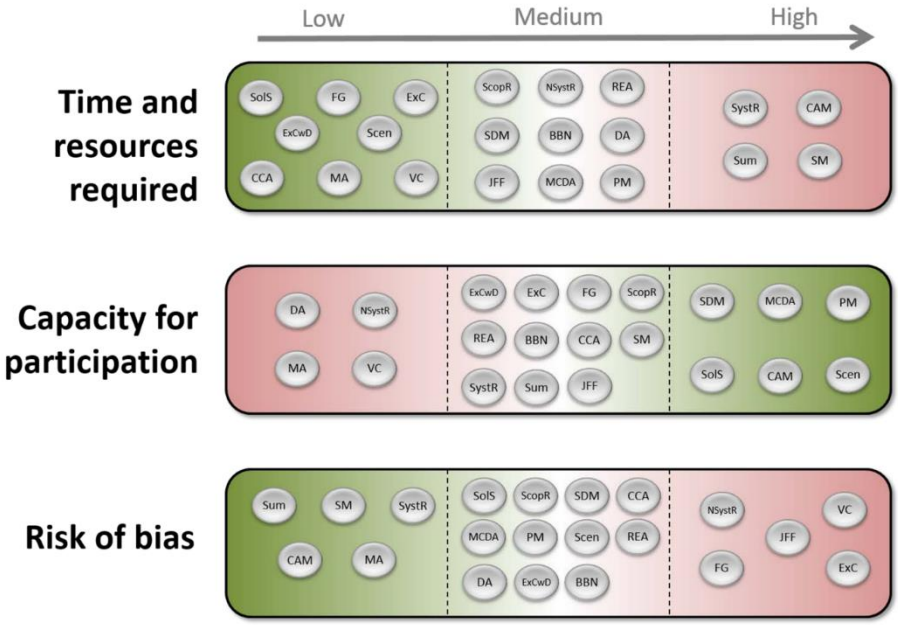

Figure 1. How the 21 KSMs fall among the categories. The colour scheme is used to anticipate which end of each scale might be considered more desirable, where green = desirable (low cost, high participation, low risk of bias) and red = undesirable. This scheme is speculative, and will differ among policy contexts. Taken from Dicks, L., Haddaway, N., Hernández-Morcillo, M., Mattsson, B., Randall, N., Failler, P., Ferretti, J., Livoreil, B., Saarikoski, H., Santamaria, L. and Rodela, R., 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.

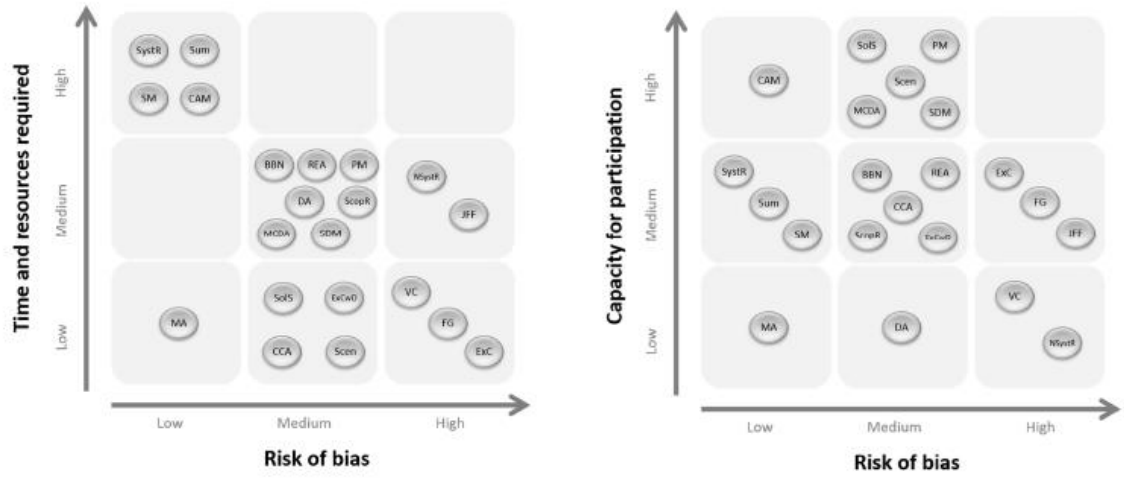

Figure 2. The distribution of the 21 KSMs. The distribution of the 21 KSMs when two strength/weakness factors are seen simultaneously, to illustrate trade-offs: a) Time/resource against risk of bias; b) Capacity for participation against risk of bias. Red circles indicate *evidence synthesis methods*. Adapted from Dicks et al, 2017.

These visualisations demonstrate that there is a trade-off between risk of bias and time and resources required, with meta-analysis being the only method to be low for both. The authors of the EKLIPSE report note that meta-analysis is not a standalone method, however, and that it must be based on a previous review and data extraction phase – in essence this is part of a literature review. The authors do not go into detail about the relationship between risk of bias and the capacity for participation, but no real pattern appears to exist, with most methods capable of involving stakeholders to some degree.

Of the 21 knowledge synthesis methods described above, 9 involve evidence synthesis to some degree: systematic review; subject-wide evidence syntheses; meta-analysis; rapid evidence assessment; scoping review; systematic map; vote-counting; non-systematic literature review; and causal criteria analysis.

Some of these methods above should not be seen as standalone evidence synthesis methods, due to severe failings in their methodological design. As noted above, meta-analysis necessitates a prior literature review before data can be synthesised, so should not be considered alone. Vote-counting was noted by the EKLIPSE project as a flawed synthesis method, since it ignores study magnitude, instead fixating on statistical significance, and also fails to consider study validity, treating all research as equally valid. As such, vote-counting should always be discouraged. Indeed, meta-analysis and vote-counting are data synthesis methods, rather than types of review. For these reasons, they will not be further discussed. Causal criteria analysis is a combination of methods that includes both the production of a visual logic model with literature reviews to synthesise evidence for particular linkages within it. As such, it relies on one of the other literature review methods, such as systematic review, and will not be considered separately as an evidence synthesis method.

Non-systematic literature reviews can be considered to include summaries/primers that do not intend to support decision-making, but rather intend to summarise broad concepts for non-expert readers. These primers are not aiming to be comprehensive/representative, accurate or transparent in their methodology, and so should not be considered as 'evidence synthesis' methods, per se. **We focus instead on those sorts of literature review that aim to comprehensively or reliably summarise a body of evidence and its findings in an accurate and reliable way.**

1.4. Main groups of evidence synthesis methodology

We are thus left with six methods: systematic review, subject-wide evidence synthesis, rapid evidence assessment, scoping review, systematic map, and non-systematic literature review.

These methods can be broadly separated into those that aim to summarise the state of evidence on a topic (i.e. "what do we know about…?"), also referred to as 'exploratory', and those that aim to synthesise findings of the evidence to answer questions about impacts or causality in a system, also referred to as 'analytical'. Systematic mapping and scoping reviews are exploratory methods, whilst systematic reviews and rapid evidence assessments are analytical. Subject-wide evidence syntheses aim to be both exploratory and analytical, whilst non-systematic literature reviews are essentially low-quality/-validity methods that can be exploratory, analytical or both. These factors are visualised in Figure 3 that describes the risk of bias, exploratory-vs-analytical nature and the presence of formalised methodology.

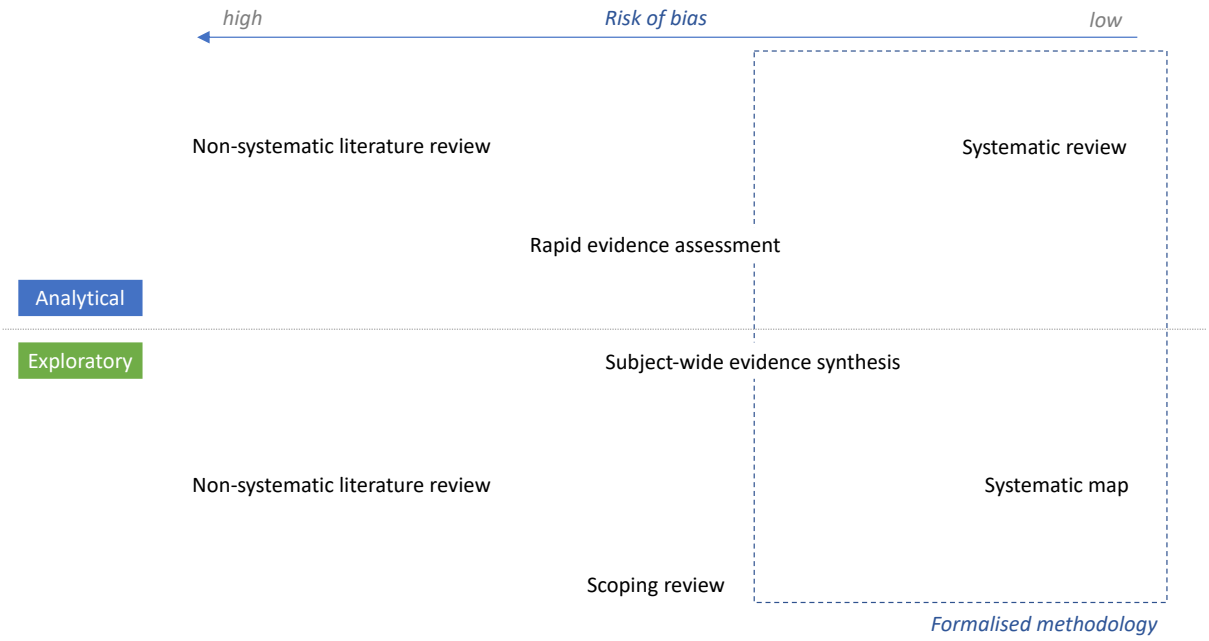

Figure 3. Conceptual model placing the six evidence synthesis methods identified in this report according to the risk of bias, the analytical-vs-exploratory approach, and the presence of formalised methodology.

For all but one of the methods (non-systematic literature review), formalised methodological guidance exists (e.g. Pullin and Stewart 2006 and see Table 2), but many authors publish reviews under the guise of these 5 methods without adhering to these guiding principles (Haddaway et al. 2020). Key problems associated with reviews that do not pay sufficient attention to methodological guidance include: selection bias; publication bias; discussion bias; and a lack of critical appraisal of study validity (Haddaway et al. 2016).

Table 2. Key guidance documents outlining methodologies for the 5 formalised evidence synthesis methods.

Evidence synthesis method	**Formalised guidance published by**
Systematic review	Various organisations, including; Cochrane, the Campbell Collaboration, the Collaboration for Environmental Evidence, and the Joanna Briggs Institute
Systematic map	The Collaboration for Environmental Evidence
Rapid evidence assessment	UK Department for Environment, Food and Rural Affairs
Scoping review	UK Department for Environment, Food and Rural Affairs
Subject-wide evidence synthesis	Conservation Evidence

In the following section, the key principles, advantages and disadvantages of each of these 5 forms of evidence synthesis will be discussed

1.4.1.Systematic review

Systematic review methodology was first developed in the field of social science (Smith and Glass 1977) but was first widely adopted and institutionalised in the field of healthcare in the 1990s, with the establishment of the Cochrane Collaboration (Bero and Rennie 1995). This formalised methodology was used for the purposes of eliminating and/or mitigating all sources of bias in the process of identifying, collating, processing, appraising, synthesising, and reporting research for a particular evidence base (Higgins et al. 2019). the major steps in the systematic review process are outlined in Figure 4.

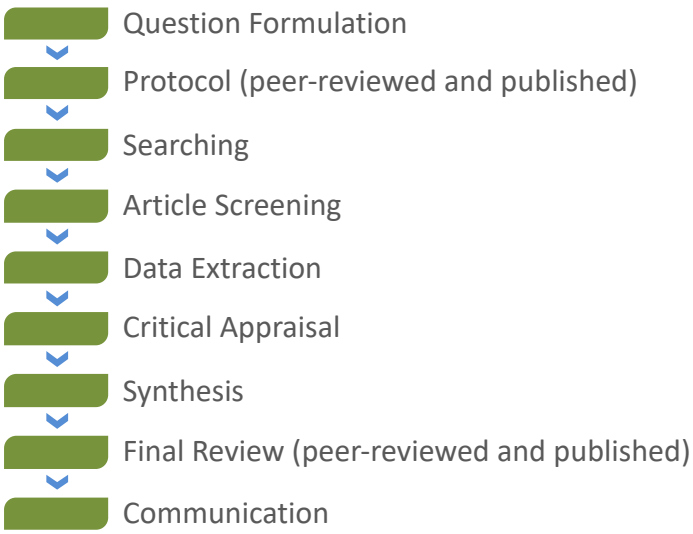

Figure 4. Schematic of the key steps in a systematic review

Systematic reviews were intentionally first used to synthesise randomised control trials for single clinical interventions with an aim of successfully assessing their effectiveness. In clinical medicine, systematic reviews soon became relatively widely accepted as the highest reliability evidence for use in the production of practice guidelines for healthcare practitioners (Cook et al. 1997; Haddaway and Pullin 2014).

In recent decades, other fields have adopted formalised/institutionalised systematic review methodology, including: social welfare (Pearson 2007); international development (Mallett et al. 2012); conservation and environmental management (Pullin and Stewart 2006); environmental health (Whaley et al. 2021); and computer science (Budgen and Brereton 2006). Concurrently, there has been a realisation that randomised control trials may not always be feasible or desirable in certain disciplines or contexts, and that observational and qualitative research methods have much to offer in terms of evidence to support decision-making. Qualitative synthesis, in particular, has grown in popularity as a means of elucidating the how and why of intervention effectiveness (Barnett-Page and Thomas 2009), where quantitative synthesis is well suited at identifying if and by how much something works. Finally, there has been a recent push to assess complex interventions in systematic reviews (Pawson et al. 2005), based on the understanding that multiple interventions are typically put in place under a variety of interdependent contextual, locally specific parameters. As a result of all of these factors, systematic review methodology has continued to develop at an increasing rate (Haddaway and Pullin 2014).

Systematic reviews are a substantial undertaking in terms of time and resources, as a result of the detailed nature of the methods behind the planning and conduct phases, all intended to minimise bias and increase reliability. Recent estimates of the time needed to complete a systematic review in healthcare have been quoted as: 1) 67.3 weeks from start to finish of a review project (Borah et al. 2017); and 2) 23.4 weeks full time equivalent of staff time (Haddaway and Westgate 2019). They also require methodological expertise from experienced systematic reviewers and librarians/information specialists in order to be efficient.

Many co-called systematic reviews are published outside of the review coordinating bodies, such as Cochrane, that fail to sufficiently mitigate bias but not adhering to the guidance and standard provided in the guidance from these organisations. As a result, these reviews are unreliable.

Furthermore, they can cause confusion in the broader research and practice community about the value and utility of systematic reviews (see, for example, Haddaway et al. 2016).

Reporting standards (e.g. ROSES (Haddaway et al. 2018) and PRISMA (Page et al. 2021)) are frequently used alongside systematic reviews during publication of the protocol and final report, which aim to increase the level of transparency and accountability of the review (Haddaway and Macura 2018).

1.4.2.Rapid evidence assessment

Rapid evidence assessment is the name given by the UK Civil Service (Collins et al. 2015) and others to a type of literature review that aims to be systematic but that typically cuts methodological corners in an effort to provide a more timely or cost-effective answer than it is perceived to be possible using systematic review (Tricco et al. 2015). Defra defines rapid evidence assessments as "a type of evidence review that aims to provide; an informed conclusion on the volume and characteristics of an evidence base, a synthesis of what that evidence indicates and a critical appraisal of that evidence". Other names used for this kind of approach are rapid review, rapid syntheses, evidence summaries, brief reviews, etc. (Haby et al 2016).

There is no consensus over which corners should be cut (i.e. which aspects of systematic review methodology can be dropped or adapted), and no evidence regarding the impact of particular shortcuts that can be generalised.

From a methodological perspective, rapid evidence assessments may involve less scoping during the planning phases, search fewer bibliographic resources, exclude grey literature, include no consistency checking with a single reviewer conducting all of the screening, limited extraction of study meta-data (descriptive information) or study findings, include only basic synthesis (with the potential to focus on vote-counting and no meta-analysis, for example).

Typical methodological guidance aspects that are different between rapid evidence assessments and systematic reviews are summarised in Table 3 and are detailed in the CEBC report WT1552 (CEBC 2015).

Table 3. Key methodological differences between rapid evidence assessments (as defined by Civil Service Guidance from Defra) and systematic reviews (as defined by the Collaboration for Environmental Evidence). Table taken from Collins et al. 2015

Stage	**Rapid evidence assessment (REA)**	**Systematic review (SR)**
Protocol development	Much less focus on scoping, no suggestion of trial critical appraisal to determine whether the information retrieved would contain extractable data of a suitable quality.	Detailed guidance on scoping to assess volume of evidence, quality of evidence and likely resources needed.
Protocol finalised	In REA, the guidance confirms that the protocol may be iteratively revised during the review process	CEE guidance acknowledges that there may be deviations from the protocol but that it is supposed to be a priori to minimise bias. Any deviations should be justified in the final report.
Searching, screening and critical appraisal stages	For REA these suggest greater methodological quality than demanded in early civil service guidance (for example, by recommending that published, grey and unpublished literature are searched). There is less emphasis on double checking for repeatability in the REA guidance.	Suggests more stringent checking of stages where reviewer bias may affect inclusion e.g. at filtering against inclusion criteria
Synthesis and reporting	REA guidance places more emphasis on describing the volume and characteristics of the evidence base, although this is commonly	SR more focused on presenting the findings, less prescriptive about providing an overview of the volume and characteristics of the evidence base.

	done for CEE systematic reviews and in particular systematic maps.	
Outputs	REAs are expected to produce a map of the evidence as a by-product, rather than as a standalone activity as in the CEE guidance. Otherwise, requirements for producing 'supplementary materials' are similar, including lists of search results etc.	SR guidance does not explicitly ask for a 'map' of the literature for systematic reviews, only for systematic maps.

In general, the formalised rapid evidence assessment guidance follows the CEE guidelines for systematic reviews (CEE 2018) quite closely, although far less detail is given and more scope for methodological flexibility is permitted. This results in some rapid evidence assessments that are equal in rigour and similar in methodological standards to CEE systematic reviews. However, the lack of an endorsing body for rapid evidence assessments allows for a broad spectrum of methodological rigour. As such, some rapid evidence assessments may be of particularly low validity.

1.4.3.Systematic map

Systematic mapping was first developed as a complementary method to systematic reviews for the purposes of gaining an overview of the evidence landscape for a particular topic (James et al. 2016). Rather than fully synthesising study findings, as is done in a systematic review, systematic maps describe the evidence base by summarising the number of studies (and less often the validity of the methodological designs) across key descriptive categorical and continuous variables; for example, publication year, measured outcomes, study design and sample size.

The method was first developed in the field of social science (Clapton et al. 2009), and in the early 2010s was adopted by the field of conservation and environmental management (Randall et al. 2012). Concurrently, the methodology was adopted in the field of software engineering (Petersen et al. 2008) in isolation from other developing fields. In the environmental field, systematic mapping is now more frequent than systematic reviews (see www.environmentalevidencejournal.org).

Systematic maps are typically more broad in scope than systematic reviews - often, several key elements will have multiple levels (for example, more than one intervention), or may not be prespecified at all in the *a priori* protocol, being instead catalogued iteratively (James et al. 2016).

Systematic mapping can also combine multiple different types of evidence within one review, unlike systematic reviews, that typically focus on one study type, such as randomised control trials or observational studies. Systematic maps can catalogue any study type, and are particularly useful for generating bibliographies of grey literature (Haddaway et al. 2016).

Key outputs from systematic mapping must include: i) an interactive, searchable database of all relevant studies complete with tabulated descriptive information for each; ii) a narrative report describing the evidence base. In addition, good systematic maps also include; iii) visualisations of the evidence base, for example as interactive heat maps (cross tabulations of the volume of evidence across two categorical variables), evidence atlases (studies plotted across geographical space), and other graphics; iv) lists of knowledge gaps (where few/no studies exist for particular subtopics) and knowledge clusters (where sufficient similar studies exist to allow full synthesis in a systematic review). Systematic maps can thus form a first step in the 'evidence synthesis pathway' from broad concern to multiple, focused systematic reviews: doing so can increase the efficiency and success of synthesis projects, and can better integrate stakeholder engagement in the process of choosing topics for full review following systematic mapping (Haddaway et al. 2016).

Systematic maps follow the same stringent methodology as systematic reviews (although are typically larger in volume), until the point of data extraction and critical appraisal (James et al. 2016).

From here, systematic mapping extracts only meta-data (descriptive information) and not full study findings, and no critical appraisal is typically performed (although meta-data relating to validity are often extracted, for example sample size).

1.4.4.Scoping review

Scoping reviews are, like rapid evidence assessments, a poorly defined group of methods: indeed, the term is sometime used synonymously with rapid review. However, in general, scoping reviews are seen as preparatory or exploratory evidence synthesis method that aims to assess the volume and nature of evidence in advance of conducting a full systematic review (or in place of a more thorough synthesis where time constraints are particularly challenging) (Collins et al. 2015). Defra defines *quick scoping reviews* as "a type of evidence review that aims to provide an informed conclusion on the volume and characteristics of an evidence base and a synthesis of what that evidence indicates in relation to a question" (Collins et al. 2015).

Scoping reviews can thus be considered to be a rapid and less systematic version of a systematic map, although less emphasis is made on the use of outputs from the reviewing process (i.e. a database, visualisations, etc.) and instead on a qualitative description regarding the nature of the evidence base. Scoping reviews may be conducted to inform decisions about topics for systematic review in order to avoid focusing on an empty evidence base or uncovering surprises along the way.

As with rapid evidence reviews, there is no formal guidance for scoping reviews, particularly in where to cut corners. In healthcare, however, some scoping reviews would be classified as systematic maps, for example in the field of environment. Healthcare evidence synthesis communities do not frequently use systematic mapping, and the term 'scoping review' thus covers a spectrum from non-systematic reviews right through to equivalents of robust systematic reviews (e.g. Archer et al. 2011).

1.4.5.Subject-wide evidence synthesis

Subject-wide evidence synthesis is a specific method conducted by an organisation called Conservation Evidence (www.conservationevidence.org) based at Cambridge University and led by Professor William Sutherland. This evidence synthesis method is a 'practice-centric' approach that aims to be comprehensive in its listing of conservation and environmental management interventions. It aims to collate comparative experimental evidence in order to identify all possible actions that practitioners could put in place, and to report the results of each study to support environmental decision-making.

Previously referred to as 'synopses', subject-wide evidence syntheses share similarities with systematic reviewing and mapping, but differ from these methods in that they are incredibly broad in scope, for example, focusing on bird conservation. They also differ fundamentally in their methodological rigour: subject-wide evidence syntheses screen for evidence only within a predetermined list of journals deemed to be relevant to the topic, whereas systematic reviews search across multiple bibliographic databases and supplementary resources. A major difference in the methodology relates to the level of transparency and reporting of the methods used: systematic reviews and maps include an *a priori* protocol, detailed and transparent methods, and minimum standards for each step of the synthesis process. Subject-wide evidence syntheses, on the other hand, do not have a publicly accessible standard methodology, and no *a priori* protocol. Instead, methods are more *ad hoc*[1]. Subject-wide evidence syntheses summarise study findings (i.e. the

[1] However, two subject-wide evidence syntheses have used systematic maps as a basis for their evidence bases, but these syntheses went beyond the process of systematic mapping to extract and describe study findings.

results of any statistical tests and the authors' conclusions regarding intervention success) for all included studies in a fairly standardised short paragraph of text that includes statements linked to the sample size and some aspects of study design (e.g. treatment allocation and the nature of comparators). However, they do not include a full critical appraisal of study validity, and do not reanalyse study findings: they rely on the primary research authors' analyses and conclusions.

Subject-wide evidence syntheses are conceptually far more broad than systematic reviews and maps, and provide on a user-friendly interface and document style focused on supporting practitioners to make decisions regarding alternative interventions. However, as a result of the trade-offs between subject breadth and methodological rigour, they are more susceptible to bias than systematic reviews or maps.

References Annex A

Archer, N., Fevrier-Thomas, U., Lokker, C., McKibbon, K.A. and Straus, S.E., 2011. Personal health records: a scoping review. Journal of the American Medical Informatics Association, 18(4), pp.515-522.

Barnett-Page, E. and Thomas, J., 2009. Methods for the synthesis of qualitative research: a critical review. BMC medical research methodology, 9(1), pp.1-11.

Belton V and Stewart TJ (2002) Multiple criteria decision analysis: an integrated approach. Kluwer, London.

Bero, L. and Rennie, D., 1995. The Cochrane Collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. Jama, 274(24), pp.1935-1938.

Borah R, Brown AW, Capers PL, Kaiser KA. 2017 Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. BMJ Open. Feb 27;7(2):e012545. doi: 10.1136/bmjopen-2016-012545. PMID: 28242767; PMCID: PMC5337708.

Budgen, D. and Brereton, P., 2006, May. Performing systematic literature reviews in software engineering. In Proceedings of the 28th international conference on Software engineering (pp. 1051-1052).

Burgman MA et al (2011) Expert status and performance. PLoS ONE 6, e22998.

van Buuren A, Edelenbos J and Klijn EH (2007) Managing knowledge in policy networks. Organising joint fact-finding in the Scheldt Estuary. URL: https://repub.eur.nl/pub/10165

CEBC Defra 2015 Emerging Tools and Techniques to Deliver Timely and Cost Effective Evidence Reviews Final report WT1552 Defra, UK - Science Search file:///C:/Users/flohoo/Downloads/13630_WT1552ToolsandTechniquesforEvidenceReviews-FinalReport.pdf

Clapton J, Rutter D, Sharif N. SCIE Systematic mapping guidance; April 2009. http://www.scie.org.uk/publications/researchresources/rr03.pdf Accessed 19 Oct 2015.

Collins A et al (2014) The Production of Quick Scoping Reviews and Rapid Evidence Assessments: A How to Guide. Joint Water Evidence Group, UK.

Cook DJ, Mulrow CD, Haynes RB 1997. Systematic Reviews: Synthesis of Best Evidence for Clinical Decisions. Ann Intern Med126:376-380. doi:10.7326/0003-4819-126-5-199703010-00006

Haby, M.M., Chapman, E., Clark, R. et al. 2016 What are the best methodologies for rapid reviews of the research evidence for evidence-informed decision making in health policy and practice: a rapid review. Health Res Policy Sys 14, 83. https://doi.org/10.1186/s12961-016-0155-7

Haddaway, N.R., Macura, B., Whaley, P. and Pullin, A.S., 2018. ROSES RepOrting standards for Systematic Evidence Syntheses: pro forma, flow-diagram and descriptive summary of the plan and conduct of environmental systematic reviews and systematic maps. Environmental Evidence, 7(1), pp.1-8.

Haddaway, N.R., Bethel, A., Dicks, L.V., Koricheva, J., Macura, B., Petrokofsky, G., Pullin, A.S., Savilaakso, S. and Stewart, G.B., 2020. Eight problems with literature reviews and how to fix them. Nature Ecology & Evolution, 4(12), pp.1582-1589.

Haddaway, N.R., Land, M, Macura, B, 2016 "A little learning is a dangerous thing": A call for better understanding of the term 'systematic review', Environ Int, http://dx.doi.org/10.1016/j.envint.2016.12.020

Haddaway, N.R. and Macura, B., 2018. The role of reporting standards in producing robust literature reviews. Nature Climate Change, 8(6), pp.444-447.

Haddaway, N.R., Pullin, A.S. 2014 The Policy Role of Systematic Reviews: Past, Present and Future. Springer Science Reviews 2, 179–183. https://doi.org/10.1007/s40362-014-0023-1

Haddaway NR and Westgate MJ (2018) Predicting the time needed for environmental systematic reviews and systematic maps. Conservation Biology. 33: 434-443. https://doi.org/10.1111/cobi.13231 The PredicTER tool is available at https://estech.shinyapps.io/predicter/

Haddaway NR, Woodcock P, Macura B, Collins A. 2015 Making literature reviews more reliable through application of lessons from systematic reviews. Conserv Biol. Dec;29(6):1596-605. doi: 10.1111/cobi.12541. Epub 2015 Jun 1. PMID: 26032263.

Higgins, J.P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J. and Welch, V.A. eds., 2019. Cochrane handbook for systematic reviews of interventions. John Wiley & Sons.

IPBES (2016). The methodological assessment report on scenarios and models of biodiversity and ecosystem services. S. Ferrier, K. N. Ninan, P. Leadley, R. Alkemade, L. A. Acosta, H. R. Akçakaya, L. Brotons, W. W. L. Cheung, V. Christensen, K. A. Harhash, J. Kabubo-Mariara, C. Lundquist, M. Obersteiner, H. M. Pereira, G. Peterson, R. Pichs-Madruga, N. Ravindranath, C. Rondinini and B. A. Wintle (eds.). Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Bonn, Germany. 348 pages.

James, K.L., Randall, N.P. and Haddaway, N.R., 2016. A methodology for systematic mapping in environmental sciences. Environmental evidence, 5(1), pp.1-13.

Mallett, R., Hagen-Zanker, J., Slater, R. and Duvendack, M., 2012. The benefits and challenges of using systematic reviews in international development research. Journal of development effectiveness, 4(3), pp.445-455.

Martin TG et al (2012) Eliciting expert knowledge in conservation science. Conservation Biology 26, 29-38.

Mukherjee N et al (2015). The Delphi technique in ecology and biological conservation: applications and guidelines. Methods in Ecology and Evolution 6, 1097-1109.

Orvik A et al (2013) Situational factors in focus group studies: a systematic review. Int J Qual Methods 12:338–358.

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M. and Moher, D., 2021. Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. Journal of Clinical Epidemiology, 134, pp.103-112.

Pawson, R., Greenhalgh, T., Harvey, G. and Walshe, K., 2005. Realist review-a new method of systematic review designed for complex policy interventions. Journal of health services research & policy, 10(1_suppl), pp.21-34.

Pearson, M., 2007. Systematic reviews in social policy: to go forward, do we first need to look back?. Evidence & Policy: A Journal of Research, Debate and Practice, 3(4), pp.505-526.

Petersen, K., Feldt, R., Mujtaba, S. and Mattsson, M., 2008, June. Systematic mapping studies in software engineering. In 12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12 (pp. 1-10).

Pullin, A.S. and Stewart, G.B., 2006. Guidelines for systematic review in conservation and environmental management. Conservation biology, 20(6), pp.1647-1656.

Randall NP, James KL. 2012 The effectiveness of integrated farm management, organic farming and agri-environment schemes for conserving biodiversity in temperate Europe—A systematic map. Environ Evid.1:4.

Smith, M.L. and Glass, G.V., 1977. Meta-analysis of psychotherapy outcome studies. American psychologist, 32(9), p.752.

Tricco, A.C., Antony, J., Zarin, W., Strifler, L., Ghassemi, M., Ivory, J., Perrier, L., Hutton, B., Moher, D. and Straus, S.E., 2015. A scoping review of rapid review methods. *BMC medicine*, *13*(1), pp.1-15.

Whaley, P., Blaauboer, B.J., Brozek, J., Hubal, E.A.C., Hair, K., Kacew, S., Knudsen, T.B., Kwiatkowski, C.F., Mellor, D.T., Olshan, A.F. and Page, M.J., 2021. Improving the quality of toxicology and environmental health systematic reviews: What journal editors can do. ALTEX-Alternatives to animal experimentation, 38(3), pp.513-522.

Annex B: Lessons learned from the BONUS synthesis call (BONUS synthesis 2017)

In this section, the results of a set of key informant interviews and an output analysis are described, that were conducted to explore barriers and facilitators to efficient commissioning, conduct and communication of evidence syntheses. The aim of the interviews was to see whether there were any gaps in relation to understanding the very complex terminology around synthesis, whether there were any tools that could have made work easier that were unknown to applicants, whether there were results from the projects that were taken up by management, whether the call guidelines were clear and the procedures effective and whether there are general bottlenecks that could be removed or mitigated to make future calls more effective. The interviews and assessment of project outputs have provided vital insights for the main report on understandings, awareness and application of systematic review methods and stakeholder engagement.

The BONUS programme has had a focus on compiling project results already early in its history, starting with a synopsis of the results of the 16 projects funded during the BONUS+ programme, which ran from 2009-2012. The project results were presented in a special issue of the AMBIO journal (2015). Subsequently, the potential of synthesis was pointed out in an analysis of the BONUS outcome in Pauline Snoeijs-Leijonmalm, S. B. Towards better integration of environmental science in society: Lessons from BONUS, the joint Baltic Sea environmental research and development programme (Environmental Science & Policy Volume 78, December 2017), leading to a synthesis call in 2017.

For understanding of the aims of the call, we analysed the terminology of the call text. The call brief (BONUS Briefing 28) refers to 'knowledge synthesis' and 'synthesising research outputs', implying the need for systematic approaches ("Projects are expected to analyse as broad body of research outputs as necessary for robust and unbiased review and credible gap analysis"), and explicitly calling for systematic reviews: "Wherever possible and appropriate, projects are expected to apply systematic review methodology, combined, where necessary, with a narrative approach". As such, the language is relatively explicit that evidence synthesis should be the most appropriate method. However, when looking at the tasks within each of the 9 topics listed, it occurred that they are not all well suited to evidence synthesis, which in turn indicates that the call did not have as strict a focus on synthesis as the title and the overarching description suggests. We explore this in Table 3. Some topics do not appear suited to evidence synthesis at all (e.g. topic 1), some have certain tasks for which evidence synthesis may be appropriate methods (e.g. topic 2), whilst others appear entirely answerable with evidence synthesis (e.g. topic 7). For other topics the relevance of evidence synthesis is unclear (e.g. topic 5) since the framing of the type of knowledge to be synthesised is not given – more specifically, evidence synthesis is possible, but would include only published empirical research, which would likely miss vital information from other sources.

Table 3. Call topics and their suitability for evidence synthesis as defined by the major evidence synthesis coordinating bodies (Cochrane, the Campbell Collaboration, Joanna Briggs Institute and the Collaboration for Environmental Evidence; e.g. https://www.cochrane.org/news/evidence-synthesis-what-it-and-why-do-we-need-it).

Call text	**Suitability for evidence synthesis**
1. Sustainable marine and freshwater aquaculture development perspectives in the Baltic Sea region	
A comprehensive analysis of the state of aquaculture in the Baltic Sea region	**None apparent.** Evidence synthesis only relevant if the scope is aquaculture research of research papers investigating aquaculture facilities.
A detailed policy framework analysis at national, Baltic Sea region and EU levels in areas relevant to development of sustainable aquaculture production	**None apparent.** Evidence synthesis only relevant if published research on policies are intended to be synthesised (i.e. not an analysis of policy itself).
Scientifically justified criteria to enable environmentally balanced, economically viable and socially acceptable aquaculture production across the Baltic Sea region	**Unclear.** Evidence synthesis may be appropriate if criteria are to be selected from the research literature (i.e. if 'scientifically justified' means 'empirically tested or described in the research literature).
An outlook of future aquaculture development in the Baltic Sea region based on these criteria to identify the key needs for new knowledge and innovative solutions	**None apparent.** Evidence synthesis only relevant if future developments are to be collated from within the research literature.
2. A synthesis of knowledge on the Baltic Sea food webs including an outlook for priority future studies	
A synthesis of our knowledge on the Baltic Sea food webs covering all trophic levels and their linkage to nutrient cycling	**Clear.** Call asks for synthesis of knowledge - the question is rather broad but would lend itself to mapping.
An outlook of food web and ecosystem functioning research in the Baltic Sea, taking into account the priority knowledge needs (gaps), new methodological approaches and diversity of the Baltic Sea ecosystem	**Clear.** Calls asks for assessment of knowledge needs, methodologies used (in primary research) and ecosystems (implied) investigated.
A feasibility study on application of food web models to validate the use of indicators and associated threshold values under the EU Marine Strategy Framework Directive Descriptor 4 and to implementation of the EU Common Fisheries Policy	**None apparent.** Evidence synthesis is only relevant if multiple feasibility studies are to be collated and summarised.
An analysis of how the knowledge gained in this area has been and could be used more efficiently in order to ensure sustainable use of the Baltic Sea ecosystem services and biological resources	**None apparent.** Evidence synthesis is only relevant if multiple impact evaluations are collated and summarised.
3. Towards improved environmental status assessment and monitoring systems for the Baltic Sea	
A critical analysis of currently applied monitoring systems, their correspondence with the assessment requirements under different policies and identification of the priority ways of improvement	**Unclear.** Evidence synthesis is only relevant if only published evaluations of monitoring systems are to be collated and appraised.
A review of current development of novel cost-efficient marine monitoring and assessment approaches and methodologies applicable in the Baltic Sea	**Unclear.** Evidence synthesis may be appropriate if novel approaches and methods are only available in the empirical literature.
A proposal for a renewed monitoring system for the Baltic Sea, that closes the existing information gaps and cross-sector data gathering and handling disconnects, takes into account the state-of-the-art observation methodologies and harmonises the requirements within the EU data collection, HELCOM Baltic Sea Action Plan, EU Water Framework Directive, EU Marine Strategy Framework Directive and EU Common Fisheries Policy frameworks	**None apparent.** Evidence synthesis not appropriate – summary of other tasks and proposal from the project team.
4. Policy instruments and institutions for nutrient abatement	
A knowledge synthesis on designing instruments and institutions for nutrient abatement	**Unclear.** The topic is too broad for a single evidence synthesis.
An assessment of implementing nutrient abatement instruments and institutions worldwide	**Clear.** Evidence mapping of interventions possible, although rather broad still.

Both knowledge synthesis and assessment of practical experience shall analyse such characteristics as e.g. cost-effectiveness and targeting, transaction costs for e.g. verification, monitoring and administration, distribution of costs, fairness and equity concerns, the spatial scale of policy instrument implementation and its effect, cross-sectoral character, multi-functionality of abatement measures, creativity in fostering new solution(s) for nutrient reduction and ability to invite actors to participate in protection activities	**Clear.** In part evidence synthesis is appropriate, although the included factors make the task very large – this would typically warrant multiple systematic reviews for each factor or a large realist review.
An outlook on future priority work directions in developing fit-for-purpose nutrient abatement instruments and institutions for the Baltic Sea region	**Unclear.** Could be a summary of gaps identified in the evidence mapping above, depending on whether the source of priority areas could be restricted to published literature (or whether other knowledge systems should be included, e.g. expert opinion).
5. High frequency automated in situ observations in the Baltic Sea	
A review of the experiences with high frequency automated observations in the Baltic Sea	**Unclear.** Evidence synthesis appropriate only if primary qualitative studies of experiences exist in the literature, which seems unlikely.
An analysis of applicability of the outputs of high frequency automated observations (e.g. buoys, drifters, gliders, ARGOfloats, ferryboxes and other sensors on ships of opportunity, seabed observatories etc.) for assessment of the Baltic Sea ecosystem status	**Unclear.** Unclear what is meant by 'applicability', but evidence synthesis could be appropriate if intervention effectiveness is the subject of the task and multiple primary studies could be collated.
Suggestion on ways of assimilation of the high frequency observation data into data pools used for environmental state assessment, and other practical applications	**Unclear.** Evidence synthesis only appropriate if suggestions are collated from primary research that made suggestions, which seems unlikely.
An analysis of technical challenges to be resolved to ensure broader use of high frequency observations for obtaining new data about the Baltic Sea, including cross-sector and cross-border cooperation and coordination	**Unclear.** Evidence synthesis only appropriate if technical challenges are collated from primary research that made suggestions, which seems unlikely.
Proposals for better integration of outputs of high frequency observations into international marine data nodes e.g. EMODNET	**Unclear.** Evidence synthesis only appropriate if proposals are collated from primary research that made suggestions, which seems unlikely.
Suggestions for future development of high frequency in situ observations in the Baltic Sea	**Unclear.** Evidence synthesis potentially appropriate if future developments are identified from gaps in the primary research literature.
6. Development of a unified access point for science-based virtual decision support tools for ecosystem-based management in the Baltic Sea and its drainage	
A review of the existing virtual decision support tools for ecosystem-based management in the Baltic Sea and its drainage	**Clear.** Evidence synthesis appropriate, although clear need for strong grey literature component to capture novel tools not in the published academic literature: could either be a mapping of tools or multiple reviews of effectiveness/implementation.
Creation of a comprehensive catalogue of the existing science-based virtual decision support tools with unified access point and proposal of a mechanism for its sustainability and future development. Elaboration of a set of definitions, as well as quality and performance criteria for such tools	**Clear.** Evidence synthesis appropriate, assuming 'science-based' implies empirical assessment of effectiveness.
Suggestions for the most needed decision support tools to be developed by future projects	**Not apparent.** Evidence synthesis (mapping) only appropriate if needs arise from gaps in the primary literature, but this would miss general needs of practitioners (likely very important).
7. Non-monetary values of the Baltic Sea ecosystem goods and services provided to human lifestyles and well-being	
A review of the ecosystem services provided to humans by the Baltic Sea and linking them to human lifestyles and well-being	**Clear.** Evidence synthesis appropriate in the form of evidence mapping.

Identifying the links and possible positive and negative synergies between different marine ecosystem services	**Clear.** Evidence synthesis appropriate in the form of thematic analysis of the results of evidence mapping (above) to create a framework/conceptual model.
A review of knowledge on human health and well-being effects of the Baltic Sea ecosystem	**Clear.** Evidence synthesis appropriate in the form of evidence mapping, although this seems to be a core part of the task above.
An analysis of methods and models available to estimate nonmonetary values of marine ecosystem services	**Clear.** Evidence synthesis appropriate in the form of evidence mapping.
An analysis of methods and models available to estimate monetary values of marine ecosystem services	**Clear.** Evidence synthesis appropriate in the form of evidence mapping.
8. Improved maritime risk analysis and mitigation	
A review of the risk analysis methods and decision support tools available for authorities in the Baltic Sea region to cost-effectively decrease maritime and environmental risks (especially associated with shipping)	**Clear.** Evidence synthesis appropriate in the form of evidence mapping.
A review of the underlying scientific basis and efficacy of the available methods and decision support tools for the prevention of maritime accidents and crisis management, in particular considering the human behaviour in various accidental scenarios, enhancing the safety cultures, effects of proper training and various aspects of interaction between humans and technical tools	**Clear.** Evidence synthesis appropriate in the form of systematic reviews of method/tool effectiveness.
A review of the existing knowledge and analysis of future research priorities and innovation needs for advancement of e-navigation in daily shipping practices, including such aspects as acquiring and exchanging situational information and safest routing both in open water and ice conditions	**Clear.** Evidence synthesis appropriate in the form of evidence mapping.
A review of the existing knowledge and analysis of future research priorities and innovation needs for mitigation of the risks associated with the future maritime activities related to the renewable offshore energy developments	**Clear.** Evidence synthesis appropriate in the form of evidence mapping.
9. Cumulative effects of human activities: linear and non-linear interactions and knowledge gaps	
A synthesis of the existing knowledge about cumulative (including additive, antagonistic and synergistic) effects of human activities and multiple stressors on different elements and the whole Baltic Sea ecosystem	**Clear.** Evidence synthesis appropriate in the form of evidence mapping and a very large or multiple systematic review(s).
A gap analysis to identify the most pressing research questions for the future	**Clear.** Evidence synthesis appropriate in the form of gaps identified from evidence mapping.
Suggestions for overcoming the key challenges in cumulative effects' assessment, e.g. incorporating uncertainty, determining significance of different stressors, choosing appropriate spatial and temporal scales	**Unclear.** Evidence synthesis appropriate if suggestions or challenges come from collating them within the primary literature.
Identification of the management challenges related to cumulative effects of human activities and suggestions for solving them	**Unclear.** Evidence synthesis appropriate if suggestions or challenges come from collating them within the primary literature.

Interviews with PIs of BONUS synthesis projects and managers from the Swedish Agency for Marine and Water Management

We conducted interviews with the project coordinators for all of the 8 funded BONUS synthesis projects (interview questions may be found in the end of this annex). Project coordinators will have a good overview of the project as a whole, and may have been involved in the conduct of specific synthesis tasks, but it should be noted that it was not feasible to interview the entire project team, and coordinators may be unaware of the specifics of a particular task that they did not execute. Interviews were also undertaken with managers at the Swedish Agency for Marine and Water Management, one of many important stakeholders for the results of BONUS projects. The interviews were conducted to get an idea especially of potential understanding, awareness and application of rigorous evidence synthesis methods for a future programme.

Understanding of the definition of "synthesis"
We asked the interviewees to describe how they understood the definition of 'synthesis' in the context of the call. Most (but not all) of the responses clearly referred to the combination of existing knowledge rather than novel research. For example: *"Deepening the understanding of already existing research, trying to connect different research and summarise what is actually available". In one project, "synthesis" was not mentioned after the project had started*

For three of the respondents, synthesis appeared to mean the combination of findings from multiple work packages or different knowledge sources (e.g. specifically, previous BONUS projects). These respondents did not associate the word 'synthesis' with their literature reviews, but rather the process after completion of the reviews when findings from project partners were combined.

The use of formalised synthesis methods
On further questioning, only 2 of the 8 interviewees claimed to have used formal methods for the project's evidence syntheses. Both of these respondents made use of CADIMA, a free-to-use web-based review management tool that applies the Collaboration for Environmental Evidence (CEE) guidelines for systematic reviews and maps. One of these respondents was also aware of and familiar with the guidelines, whilst the other had not heard of guidelines and standards for evidence synthesis, but was very familiar with the review management tool.

One interviewee told us they used a novel method that they had developed for their literature reviews and synthesis that they called 'meta-evaluation'. This was not done according to formally accepted methodology, but the interviewee described some principles common to evidence synthesis (i.e. retrieving potentially relevant records using complex search strings).

Assessment of project publications
Along with the interviews, we have analysed the published papers that were produced by each team, in order to assess whether formal evidence methodology was used, and if not, what aspects (eg. transparency, rigour) could have been improved using evidence synthesis methodology. Also, not applying methodology despite suitability of call topics in combination with a requirement in the call text (to use systematic reviews wherever appropriate and feasible) can indicate issues for the scientific target group of a future programme to execute evidence syntheses. It can indicate specific needs regarding the call processes, the call material or for providing specific support.

For each paper, the following was examined and clear deviations from accepted best practice for literature reviews are noted: 1) presence of an *a priori* protocol outlining planned review methods; 2) use of systematic searching across multiple resources to comprehensively assemble a set of potentially relevant records; 3) systematic screening of search results against a predetermined set of inclusion criteria; 4) standardised extraction of data and descriptive information from the assembled, relevant studies; 5) critical appraisal of study validity for all included papers (relevant for full syntheses alone, e.g. meta-analyses, but not relevant for mapping reviews). This is a concise form of appraisal adapted

from AMSTAR 2.0 (Shea et al 2007 and CEESAT (Woodcock et al. 2014) tools for appraising evidence syntheses.

The results are described in Table 4. The following criteria were used for evaluation, with the following observations:

1) Presence of an *a priori* protocol outlining planned review methods
Several projects applied traditional literature review with very limited or no description of the methods used. Protocols were provided for only one of the papers. Providing a protocol would have increased transparency (as illustrated in the Communication, stakeholder engagement and knowledge transfer in BANOS context chapter), and is a very low cost action to implement, since the reviewers have a plan of how to conduct their review.

2) Use of systematic searching across multiple resources to comprehensively assemble a set of potentially relevant records
Very few studies used accepted evidence synthesis (i.e. systematic review) approaches in their literature reviews, and those that did performed systematic mapping rather than a full synthesis of study findings (e.g. meta-analysis).
Systematic searching and study screening were carried out for some of the review projects, but the details of the methods used are not described in sufficient detail to verify the methods used or repeat the work (a central tenet of the scientific method). Reporting further details of the methods used requires very little effort but greatly increases the reliability of the reviews.

Conducting systematic searching, screening, and data extraction requires careful planning and effort, but where not employed, the results of a review cannot be trusted, since the evidence base assembled may not be comprehensive or representative. Many of the review outputs could be improved in rigour by having a detailed plan for these stages peer-reviewed by methodology experts to provide constructive feedback prior to conduct.

3) Systematic screening of search results against a predetermined set of inclusion criteria

4) Standardised extraction of data and descriptive information from the assembled, relevant studies

5) Critical appraisal of study validity for all included papers
Critical appraisal is a challenging step in any review, but vital for ensuring that reliable primary studies are given more weight than those with issues (such as low sample size or inappropriate methodological design). Very few reviews considered critical appraisal, but doing so would have provided more faith in the findings, since not all research is equally as valid. Again, planning the critical appraisal could have been improved by the use of protocols, giving constructive feedback prior to conduct. Guidance exists to support reviewers in designing critical appraisal (e.g. CEE 2018).

General comments
A brief look at the CEEDER database of appraised reviews in ecology (https://environmentalevidence.org/ceeder-search/) demonstrates that the vast majority of reviews in the field of environmental science are conducted and reported poorly, but equally can be easily remedied with appropriate guidance and methodological expertise; such considerations would be extremely relevant for future calls. Several sources for guidance are presented in the recommendations section, table 3. Several examples of *gold-standard* systematic reviews were conducted within the projects and can be used as examples of rigorous evidence synthesis tasks for future calls, but in general, the level of detail reported in the review outputs of the BONUS synthesis call was very low.

Providing more detail of how the reviews were conducted requires very little effort, but greatly facilitates verification of what was done – thereby increasing trust and rigour. Various reporting standards exist to help reviewers ensure they report sufficient details; e.g. ROSES reporting standards for systematic evidence syntheses, Haddaway et al. 2018; PRISMA Eco-Evo O'Dea et al. 2020; PRISMA 2020, Page et al. 2021). See also the recommendations chapter, table 3.

Table 4. Summary of project outputs and an assessment of the methods used, briefly describing deviations from accepted gold standards in literature review methodology and highlighting ways in which the outputs could have been made more rigorous by following these gold standards.

Project	Output (policy briefs)	Output (research papers)	Formal evidence synthesis methodology used and aspects that could have been improved using formal evidence synthesis methodology	Topic addressed and suitability for evidence synthesis (argumentation in Annex B, table 3)
BONUS BALTIMARI - Review, evaluation and future of Baltic risk management	• Policy brief 1: Uncertainty in maritime risk management (pdf) • Policy brief 2: Technology readiness levels in maritime risk management (pdf) • Policy brief 3: Shipping-related decision support systems (.pdf) • Policy brief 4: Offshore renewable energy and shipping (pdf) • Policy brief 5: Technology readiness level of maritime risk research in the Baltic (pdf) • Policy brief 6: Research uptake (pdf)	Du L, Goerlandt F, Kujala P (2020) Review and analysis of methods for assessing maritime waterway risk based on non-accident events detected from AIS data. Reliability Engineering and System Safety 200, 106933, https://doi.org/10.1016/j.ress.2020.106933	Somewhat systematic approach to searching, but many potential biases and limited transparency. Screening not well described. Some critical appraisal performed (model validation assessment RQ5a-c), although peer-review used as a proxy for quality, which is not an accepted evidence synthesis critical appraisal criterion. No protocol provided. *Conclusion:* methodology is not a systematic review but includes attempts to limit bias.	**8. Improved maritime risk analysis and mitigation** Clear for all subtopics
		Gil M, Wróbel K, Montewka J, Goerlandt F (2020) A bibliometric analysis and systematic review of shipboard Decision Support Systems for accident prevention, Safety Science, 128, 104717, https://doi.org/10.1016/j.ssci.2020.104717	Combined bibliometric and evidence mapping method used. Limited systematic search of one database followed by limited systematic screening, although few details are provided. Highly specific ranking calculation used to select only 3 papers in each of 9 categories based in part on author list. No protocol provided. *Conclusion*: methodology is a limited form of evidence mapping combined with bibliometrics, but only a fraction of the evidence base is reviewed.	
		Schröder-Hinrichs JU, Hebbar AA, Alamoush AS. Maritime risk research and its uptake in policymaking: A case study of the Baltic Sea Region . J. Mar. Sci. Eng. 2020, 8(10), 742; https://doi.org/10.3390/jmse8100742	No methods provided. *Conclusion:* methodology absent, traditional literature review subject to typical biases.	
		Kaikkonen L, Tuuli Parviainen, Mika Rahikainen, Laura Uusitalo, Annukka Lehikoinen. Bayesian Networks in Environmental Risk Assessment: A Review. Integrated Environmental Assessment and Monitoring Volume17, Issue1 January 2021 Pages 62-78; https://doi.org/10.1002/ieam.4332	Systematic searches performed, although grey literature excluded. Systematic screening performed, including dual screening and consistency checking. Detailed data extraction and coding schema provided, although no consistency checking performed/reported. No protocol provided.	

			Conclusion: methodology approaches systematic mapping, although no *a priori* protocol provided.	
		Kulkarni K, Goerlandt F, Li J, Valdez Banda O, Kujala P (2020) Preventing shipping accidents: Past, present, and future of waterway risk management with Baltic Sea focus. Safety Science, 129, 104798, https://doi.org/10.1016/j.ssci.2020.104798	Narrow systematic-style search performed, followed by a further search focusing only on common authors. Screening may not have been performed – other than author affiliation assessment. Data extraction form used ('rubric') and consistency checking was performed. Authors also performed bibliometric analysis and topic modelling. No protocol provided. *Conclusion:* methodology employed some systematic approaches but combined in a way that does not follow accepted standards and may be subject to bias and limitations in comprehensiveness. Limited transparency is provided.	
BONUS DESTONY - Decision support tool for management of the Baltic Sea cosystem	• Policy Brief: How do virtual tools support the management of the Baltic Sea? (pdf) • Policy Brief: High transparency but lack of confidence assessment in Baltic Sea decision support tools (pdf) • Policy Brief: New tools needed for plans of measures and for new topics (pdf)	Nygård H, van Beest FM, Bergqvist L, Carstensen J, Gustafsson BG, Hasler B, Schumacher J, Schernewski G, Sokolov A, Zandersen M, Fleming V. Decision support tools used in the Baltic Sea area: performance and end-user preferences. Environmental Management 66, 1024–1038 (2020). https://doi.org/10.1007/s00267-020-01356-8	No literature review methodology applied. Some undescribed online searching performed. *Conclusion:* no review methodology performed.	**6. Development of a unified access point for science-based virtual decision support tools for ecosystem-based management in the Baltic Sea and its drainage** Clear apart from subtopic: Suggestions for the most needed decision support tools to be developed by future projects
		Schumacher J, Bergqvist L, Carstensen J, Gustafsson B, Hasler B, Fleming V, Nygård H, Pakalniete K, Sokolov A, van Beest F, Zandersen M, Schernewski G. Bridging the science-policy gap – towards better integration of decision support tools in coastal and marine policy implementation, Front. Mar. Sci., 16 October 2020 | https://doi.org/10.3389/fmars.2020.587500	Systematic-style search of policy database for the policy review, but no description of other methods used. No protocol provided. *Conclusion:* systematic searching methods applied to a policy analysis but no transparency (i.e. no details of methods used to search, screen or appraise records).	
		van Beest F, Nygård H, Fleming V, Carstensen J. On the uncertainty and confidence in decision support tools (DSTs) with insights from the Baltic Sea ecosystem. Ambio 50, 393–399 (2021). https://doi.org/10.1007/s13280-020-01385-x	Literature searches mentioned but not described in any detail. *Conclusion:* no formal evidence synthesis methods used.	

BONUS FUMARI - Future marine assessment and monitoring of the Baltic	• Joint BONUS FUMARI and BONUS SEAM Policy Brief: Identifying Gaps and Opportunities for Future Monitoring of the Baltic Sea (pdf) • Joint BONUS FUMARI and BONUS SEAM Policy Brief: Novel methods advancing Baltic Sea environmental monitoring (pdf) • Joint BONUS FUMARI and BONUS SEAM Policy Brief: Strategic proposals for a revised Baltic Sea monitoring system (pdf)	Mack L, Attila J, Aylagas E, Beermann A, Borja A, Hering D, Kahlert M, Leese F, Lenz R, Lehtiniemi M, Liess A, Lips U, Mattila OP, Meissner K, Pyhälahti T, Setälä O, Strehse JS, Uusitalo L, Willstrand Wranne, Birk S. A synthesis of novel marine monitoring methods with the potential to enhance the status assessment of the BalticSea.Front. Mar. Sci. https://doi.org/10.3389/fmars.2020.552047	No methodology for literature review described. *Conclusion:* no formal evidence synthesis methods used. (call topics not suitable)	**3. Towards improved environmental status assessment and monitoring systems for the Baltic Sea** Unclear or none apparent for all subtopics
		Kahlert M, Eilola K, Mack L, Meissner K, Sandin L, Strömberg H, Uusitalo L, Viktorsson L, Liess A. Gaps in current Baltic Sea environmental monitoring – Science versus management perspectives. Marine Pollution Bulletin Volume 160, November 2020, 111669; https://doi.org/10.1016/j.marpolbul.2020.111669	Very basic search (with a key typographical error in the search string provided manuscript, potential invalidating the search strategy) used on only one database, grey literature excluded. No details of inclusion criteria or how they were applied. No protocol provided. *Conclusion:* systematic-style evidence mapping used with very limited transparency. (call topics not suitable but methodology used anyway)	
		Lehikoinen A, Olsson J, Bergström L, Bergström U, Bryhn A, Fredriksson R, Uusitalo L. (2019) Evaluating complex relationships between ecological indicators and environmental factors in the Baltic Sea: A machine learning approach. Ecological Indicators 101:117-125; https://doi.org/10.1016/j.ecolind.2018.12.053	Not a literature review (call topics not suitable)	
		Koski V.; Kotamäki N.; Hämäläinen H.; Meissner K.; Karvanen J. & Kärkkäinen S. 2020: The value of perfect and imperfect information in lake monitoring and management. Science of the Total Environment https://doi.org/10.1016/j.scitotenv.2020.138396	Not a literature review (call topics not suitable)	
BONUS MARES - Multi-method assessment for resilient	• Policy brief: Ecosystem Services and their Socio-Economic Benefits to Humans (pdf)	Heckwolf MJ, Peterson A, Jänes H, Horne P, Künne J, Liversage K, Sajeva M, Reusch TBH, Kotta J. From ecosystems to socio-economic benefits: a systematic review of coastal ecosystem services	Systematic searching performed on one database, grey literature excluded. Systematic screening including consistency checking. Data extraction form used. No protocol provided.	**7. Non-monetary values of the Baltic Sea ecosystem goods and services provided to human lifestyles and well-being**

ecosystem services and human-nature system integration	• Policy brief: Playing an Eco-GAME to assess the quality of scientific knowledge for evidence-based decision making (pdf) • Policy brief: Transferring knowledge on ecosystems and their benefits in the Baltic Sea region - A Geospatial toolkit to support decision making (pdf)	in the Baltic Sea. Science of The Total Environment Volume 755, Part 2, 10 February 2021, 142565 https://doi.org/10.1016/j.scitotenv.2020.142565	*Conclusion:* methodology is a basic systematic mapping with limited sources of information but conducted and reported to a relatively high standard.	Clear for all subtopics
		Sajeva M, Maidell M, Kotta J, Peterson A. An Eco-GAME meta-evaluation of existing methods for the appreciation of ecosystem services. Sustainability 2020, 12(18), 7805; https://doi.org/10.3390/su12187805	Summary of the paper above (Heckwolf et al. 2021) *Conclusion:* a summary of another literature review.	
		Sajeva M, Maidell M, Kotta J. A participatory geo-spatial toolkit for science integration and knowledge transfer informing SDGs based governance and decision making. September 2020 Sustainability 12(19):8088 DOI: https://doi.org/10.3390/su12198088	Summary of the paper above (Heckwolf et al. 2021) *Conclusion:* a summary of another literature review.	
BONUS ROSEMARIE - Blue health and wealth from the Baltic Sea – a participatory systematic review for smart decisions	• Science-policy brief: Scientific evidence on marine and coastal ecosystem services in the Baltic Sea (pdf) • Science-policy brief: What evidence exists for the impact of Baltic Sea ecosystems on human health and well-being (pdf) • Science-policy brief: Scientific evidence on the use of non-monetary and monetary valuation methods in the Baltic Sea management (pdf)	Kuhn K, Oinonen S, Trentlage J, Riikonen S, Vikström S, Burkhard B: Participatory systematic mapping as a tool to identify gaps in ecosystem services research: insights from a Baltic Sea case study, Ecosystem Services, 48, 2021, 101237, https://doi.org/10.1016/j.ecoser.2020.101237	Systematic searching performed across multiple databases, but grey literature excluded. Systematic screening, including consistency checking. Systematic-style data extraction and coding. No protocol provided. Well documented methods. *Conclusion:* methodology is a full systematic map.	**7. Non-monetary values of the Baltic Sea ecosystem goods and services provided to human lifestyles and well-being** Clear for all subtopics
		Storie J; Suškevis M; Külvik M; Lehtoranta V; Vikström S; Riikonen S; Kuosa H; Kuhn K, Oinonen S. (2020) What evidence exists for the impact of Baltic Sea ecosystems on human health and well-being? A systematic map protocol. Environ Evid 9, 5 (2020). https://doi.org/10.1186/s13750-020-00189-6	Protocol for a systematic map published by Environmental Evidence, the official journal for the Collaboration for Environmental Evidence, which produces guidelines for systematic reviews and maps. *Conclusion:* high quality plan for a systematic map.	
BONUS SEAM - Towards streamlined Baltic Sea environmental assessment and monitoring	• BONUS SEAM Policy brief - Joint approaches in open-sea monitoring of the Baltic Sea (pdf) • BONUS SEAM Policy brief - Strategies for revising	Mack L, Attila J, Aylagas E, Beermann A, Borja A, Hering D, Kahlert M, Leese F, Lenz R, Lehtiniemi M, Liess A, Lips U, Mattila OP, Meissner K, Pyhälahti T, Setälä O, Strehse JS, Uusitalo L, Willstrand Wranne, Birk S. A synthesis of novel marine monitoring methods with the potential to enhance the status assessment of the	*See above for the same paper under FUMARI project* (call topics not suitable)	**3. Towards improved environmental status assessment and monitoring systems for the Baltic Sea** Unclear or none apparent for all subtopics

	monitoring in support of Baltic Sea management (pdf) • Joint BONUS FUMARI and BONUS SEAM Policy Brief: Identifying Gaps and Opportunities for Future Monitoring of the Baltic Sea (pdf) • Joint BONUS FUMARI and BONUS SEAM Policy Brief: Novel methods advancing Baltic Sea environmental monitoring (pdf) • Joint BONUS FUMARI and BONUS SEAM Policy Brief: Strategic proposals for a revised Baltic Sea monitoring system (pdf)	BalticSea.Front. Mar. Sci. https://doi.org/10.3389/fmars.2020.552047		
		Nygård H, Lindegarth M, Darr A, Dinesen GE, Eigaard OR, Lips I Developing benthic monitoring programs to support precise and representative status assessments, a case study from the Baltic Sea, Environ Monit Assess 192, 795 (2020). https://doi.org/10.1007/s10661-020-08764-7	No literature review involved ('review' refers to stakeholder questionnaire). *Conclusion:* no formal evidence synthesis methods used. (call topics not suitable)	
BONUS TOOLS2SEA - Policy tools for Baltic Sea nutrient management	• BONUS TOOLS2SEA Policy brief - Cost-effective reduction of nitrogen and phosphorous emissions to the Baltic Sea (pdf) • BONUS TOOLS2SEA Policy brief - Promise and performance of agricultural nutrient management in the Baltic Sea countries (pdf) • BONUS TOOLS2SEA Policy brief - Protecting the Baltic Sea from agricultural nutrient emissions – strengthening the policy framework (pdf)	Thorsøe, M.H., Andersen, M.S., Brady, M.V. et al. Promise and performance of agricultural nutrient management policy: Lessons from the Baltic Sea. Ambio (2021). https://doi.org/10.1007/s13280-021-01549-3	Searched multiple databases, including grey literature. No further details of the methods are provided. *Conclusion:* no formal evidence synthesis methods used.	**4. Policy instruments and institutions for nutrient abatement** Unclear for 2 subtopics: A knowledge synthesis on designing instruments and institutions for nutrient abatement An outlook on future priority work directions in developing fit-for-purpose nutrient abatement instruments and institutions for the Baltic Sea region
		Brady, M.V., Andersen, M.S., Andersson, A. et al. Strengthening the policy framework to resolve lax implementation of the Baltic Sea Action Plan for agriculture. Ambio (2021). https://doi.org/10.1007/s13280-021-01573-3	Not a literature review	
		Andersson, A., Brady, M.V. & Pohjola, J. How unnecessarily high abatement costs and unresolved distributional issues undermine nutrient reductions to the Baltic Sea. Ambio (2021). https://doi.org/10.1007/s13280-021-01580-4	Extremely restrictive search string used across multiple databases, but complemented by bibliographic checking (backward citation chasing). Systematic screening performed, including consistency checking. Very limited methodological detail provided. *Conclusion:* some systematic mapping principles followed but reported in very poor detail. Cites CEE	

			guidance and provides clear definition of level of detail needed, then provides none.	
		Eglite E, Mohm C, Dierking J. Stable isotope analysis in Baltic Sea food web research: a systematic review and vision for the future. (submitted to AMBIO in January 2021)	Publication not yet available	
BONUS XWEBS - Taking stock of Baltic Sea food webs: synthesis for sustainable use of ecosystem goods and services	• Policy brief: Managing marine resources in a sea of change – Lessons from past trajectories of biological change in the Baltic time machine (pdf) • Policy brief: Use and usefulness of food web knowledge in resource management and marine environmental conservation (pdf) • Policy brief: Using trophic models to solve the food web indicator dilemma – How to match the legislative needs of food web assessments with the structure of food web constituents and associated key ecological processes (pdf) • Policy brief: A perspective for Baltic Sea food web research – How food web knowledge can be integrated in adaptive ecosystem-based management of marine resources (pdf)	Neuenfeldt S, Nordström MC, Dierking J, Uusitalo L, Tomczak M, Haldin J, Opitz S, Bonsdorff E, Ojaveer H. Food web indicators in the 21st century: bridging the gap between scientific advice and resource management needs. Submitted to Ambio (now Korpinen et al.)	Somewhat systematic approaches to searching (multiple searches performed), including grey literature sources in some, but not well described. No further methodological details provided. *Conclusion:* no formal evidence synthesis methods used.	**2. A synthesis of knowledge on the Baltic Sea food webs including an outlook for priority future studies** Nonapparent for 2 subtopics: A feasibility study on application of food web models to validate the use of indicators and associated threshold values under the EU Marine Strategy Framework Directive Descriptor 4 and to implementation of the EU Common Fisheries Policy An analysis of how the knowledge gained in this area has been and could be used more efficiently in order to ensure sustainable use of the Baltic Sea ecosystem services and biological resources
		Nordström MC, Salo T, Eero M, Neuenfeldt S, Blenckner T, Bonsdorff E, Eglite E, Häubner N, Jacob U, Jonsson P, Köster F, Kotta J, Lindegren M, MacKenzie B, Margonski P, Möllmann C, Oesterwind D, Ojaveer H, Otto SA, Reusch T, Sommer U, Temming A, Tomczak M, Tomkiewicz J, Uusitalo L, Winder M, Dierking J. (in prep.) Gap analysis and a future perspective for Baltic Sea food web research. (submitted to AMBIO in Q1 2021)	Publication not yet available	
		Eero, M., Dierking, J., Humborg, C., Undeman, E., MacKenzie, B., Ojaveer, H., Salo, T., Köster, F. (in review) Use and usefulness of food web knowledge in resource management and marine environmental conservation. Submitted to the ICES Journal of Marine Science on September 23 2020	More a *primer* than an evidence synthesis. No methodology described. *Conclusion:* no formal evidence synthesis methods used, but rigour (comprehensiveness, representativeness, reliability, transparency, accuracy, precision) not a key objective.	
		Dierking J, Blenckner T, Bonsdorff E, Salo T, Jonsson P, Rosell EA, Herrmann J-P, Jacob U, Köster F, Kotta J, Kuosa H, Lindegren M, MacKenzie B, Margonski P, Meier M, Müller-Karulis B, Oesterwind D, Ojaveer H, Reusch T, Sommer U, Temming A, Tomczak M, Tomkiewicz J, Winder M, Nordström, M (in preparation) Food	Publication not yet available	

		web complexity and eco-evolutionary dynamics underlie diverse biological trajectories in a sea of change. Submission to Global Change Biology		
		Ojaveer, H., Kotta, J., Outinen, O., Einberg, H., Zaiko, A., Lehtiniemi, M. (in review) Simple is better: assessing ecosystem impacts of marine non-indigenous species based on the absolute effect size. (Journal of Applied Ecology Nov. 29 2020)	Systematic searching used across a small number of databases, but limited methodological detail provided. Other searching not included (citation chasing and grey literature). Systematic screening performed, including consistency checking. No details of consistency checking of data extraction. No critical appraisal performed. *Conclusion:* partially systematic meta-analysis performed subject to bias and limitations.	

Stakeholder engagement
Several projects involved stakeholders from the very beginning (at the application stage), whilst others only involved them later during the project lifetime when input to the methods used was unfeasible (these projects were still by interviewees described to have a strong stakeholder engagement approach). Much of the stakeholder engagement took place through already existing contacts (participation of project partners in working groups, correspondence groups or through already existing communication channels), several indicated difficulty to take in new stakeholders; several indicated that engagement needs particular attention and sufficient allocated resources and time. Several stakeholders from the Swedish Agency for Marine and Water management indicated that it would have been relevant with an increased stakeholder engagement (towards NGOs, policy and management organisations), and additional (digital) workshops and seminars, but they also indicated that their time is very limited.

The interviews confirmed that the complexity and size of the produced material could be challenging for stakeholders. Furthermore, there may be a semantic gap between researchers and end users, meaning that careful planning and meaningful early stage engagement is vital to ensure terminology is understood in the same way and that final reports use a broadly shared understanding of key concepts and definitions.

It has to be noted that the stakeholder landscape greatly differs in different countries, with some structures being accessible with more ready procedures for collaboration, whereas the system in other countries is so fragmented that identification and engagement are challenging.
It was indicated in several interviews with BONUS PIs that stakeholders had completely different values and mindset, and did not always understand methods and results, illustrating the relevance of discussion. Several project leaders indicated challenges understanding management processes, which can be an important prerequisite for producing relevant outputs.

For several of the projects, it was indicated by managers at the Swedish Agency for Marine and Water Management that the projects highlighted areas for new research and confirmed the relevance of ongoing practice and shared views, but did not provide new insight. This was partly because the timing of the projects was in discrepancy with management needs (management cycles in HELCOM or MSFD), because changes in reporting practice or political decisions were not taken into account, partly because the projects missed practical aspects of relevance in relation to their approach (eg including scientific papers where inclusion of grey literature and practice could have been more relevant). A closer link to management could therefore potentially have benefitted project planning and call management.

For tailoring communications, providing exactly the right input is essential, but several interviewees indicated that there are lacking pathways to implementation of new methods in management, that knowledge transfer to management is challenging and slow. Also here, time restraints at the management side can be a critical issue.

Interview questions about funded BONUS Synthesis projects

BONUS project leaders:
Many thanks for agreeing to support us with this interview. Your answers will be used to improve any future calls for syntheses in the potential continuation project for BONUS, BANOS. With this interview, we wanted to ask you some general questions about your experience in applying for, conducting and reporting your syntheses in response to the BONUS synthesis call. If it is OK for you, we'd like to record

this call for note-taking purposes only – only FTH and NRH will have access to the recordings and they are both bound to confidentiality. Your responses will be anonymised and you will not be identifiable from your answers – we will only report on aggregated responses across interviewees and themes that arise. In the event that we would like to use a quotation, we will contact you to verify the quote and ask permission before including it in any report. Do you have any questions at this point?

Methodology and guidance
We'd like to start by asking you about your application to the BONUS synthesis call. Could you tell us a bit about what you understood 'synthesis' to mean?

- Which synthesis methodology was originally planned?
- Were the plans for the synthesis method changed after planning and initiating the project?
- Did you follow any methodological guidance when conducting your synthesis?
- If so, how easy was the guidance to understand/use?
- Did you seek external support/training/mentoring?
- Did you follow guidance for reporting (PRISMA/ROSES)?
- What were the main challenges when conducting your synthesis?
- What would you do differently if you had another chance to apply for the synthesis funding?

Communication with stakeholders
Moving on to think about communicating the findings of your synthesis, who were the project's stakeholders and how were they involved?

- Who were the stakeholders in your project?
- Were the stakeholders involved already from the project planning stage?
- How did you structure the communication with users (policy and management, other?)
- Did you do anything specific to communicate your findings because this was a synthesis?
- Do you think syntheses need to be treated differently when it comes to communication? If yes, how?

Outcome
We are interested to see also the outcome of the projects, and if there are ways we can increase their impact

- What were the concrete outcomes of the project that were taken up in policy and management?
- (if no concrete answer: what are the anticipated concrete outcomes of the project that will be taken up in policy and management?) Do you have any evidence of impact already?
- Were/are there any expected bottlenecks for uptake of project results in policy and management?

BONUS project users/Managers:

- What was/is the (expected) impact of the projects on policy/management?
- Which concrete outcomes of the project led to concrete changes/uses in policy/management?
- Were you involved in the entire project lifetime? How was the interaction?
- How would you define the term 'synthesis' in the context of BONUS?
- Did the projects end up providing results that you understand to be 'syntheses'?
- Could the projects have done anything different to provide more useful results for you (eg. use a different synthesis method)?

References Annex B

AMBIO. (2015). BONUS+ in support of the ecosystem approach to management in the Baltic Sea. Volume 43 issue 1.

BONUS call 2017: synthesis https://www.bonusportal.org/files/5802/BONUS_Briefing_28_Synthesis.pdf

Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z, et al. 2007 External Validation of a Measurement Tool to Assess Systematic Reviews (AMSTAR). PLoS ONE 2(12): e1350. https://doi.org/10.1371/journal.pone.0001350

Woodcock P, Pullin AS, Kaiser MJ 2014, Evaluating and improving the reliability of evidence syntheses in conservation and environmental science: A methodology, Biological Conservation, Volume 176, Pages 54-62