

SYSTEMATIC MAP

Open Access



The nature and extent of evidence on methodologies for monitoring and evaluating marine spatial management measures in the UK and similar coastal waters: a systematic map

Bethan C. O'Leary^{1,2,3*} , Joshua P. Copping¹, Nibedita Mukherjee^{1,4}, Sandra L. Dorning², Bryce D. Stewart², Emma McKinley⁵, Prue F. E. Addison⁶, Chris Williams⁷, Griffin Carpenter⁷, David Righton^{8,9} and Katherine L. Yates¹

Abstract

Background: Anthropogenic degradation of marine ecosystems is widely accepted as a major social-ecological problem. The growing urgency to manage marine ecosystems more effectively has led to increasing application of spatial management measures (marine protected areas [MPAs], sectoral [e.g. fishery] closures and marine spatial planning [marine plans]). Understanding the methodologies used to evaluate the effectiveness of these measures against social, economic, and ecological outcomes is key for designing effective monitoring and evaluation programmes.

Methods: We used a pre-defined and tested search string focusing on intervention and outcome terms to search for relevant studies across four bibliographic databases, Google Scholar, 39 organisational websites, and one specialist data repository. Searches were conducted in English and restricted to the period 2009 to 2019 to align with current UK marine policy contexts. Relevant studies were restricted to UK-relevant coastal countries, as identified by key stakeholders. Search results were screened for relevance against pre-defined eligibility criteria first at title and abstract level, and then at full text. Articles assessed as not relevant at full text were recorded with reasons for exclusion. Two systematic map databases of meta-data and coded data from relevant primary and secondary studies, respectively, were produced.

Review findings: Over 19,500 search results were identified, resulting in 391 relevant primary articles, 33 secondary articles and 49 tertiary reviews. Relevant primary articles evaluated spatial management measures across a total of 22 social, economic and ecological outcomes; only 2.8% considered all three disciplines, with most focused exclusively on ecological (67.8%) or social (13.3%) evaluations. Secondary articles predominately focused on ecological evaluations (75.8%). The majority of the primary and secondary evidence base aimed to evaluate the effectiveness of MPAs (85.7% and 90.9% respectively), followed by fisheries closures (12.5%; 3.0%) with only 1.8% of primary, and 6.1% of secondary, articles focused on marine plans or on MPAs and fisheries closures combined. Most evaluations

*Correspondence: bethan.oleary@york.ac.uk

¹ School of Science, Engineering and Environment, University of Salford, Manchester M5 4WX, UK

Full list of author information is available at the end of the article



© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

reported within primary articles were conducted for a single site (60.4%) or multiple individual sites (32.5%), with few evaluating networks of sites (6.9%). Secondary articles mostly evaluated multiple individual sites (93.9%). Most (70.3%) primary articles conducted principal evaluations, i.e. basic description of effects; 29.4% explored causation; and 0.3% undertook benefit evaluations. Secondary articles predominately explored causation (66.7%) with the remainder conducting principal evaluations. Australia (27.4%), the USA (18.4%) and the UK (11.3%) were most frequently studied by primary articles, with secondary articles reporting mostly global (66.7%) or European (18.2%) syntheses.

Conclusions: The systematic map reveals substantial bodies of evidence relating to methods of evaluating MPAs against ecological outcomes. However, key knowledge gaps include evaluation across social and economic outcomes and of overall merit and/or worth (benefit evaluation), as well as of: marine plans; networks of sites; real-time, temporary or seasonal closures; spatial management within offshore waters, and lagoon or estuary environments. Although the evidence base has grown over the past two decades, information to develop comprehensive evaluation frameworks remains insufficient. Greater understanding on how to evaluate the effectiveness of spatial management measures is required to support improved management of global ocean resources and spaces.

Keywords: Fishery closures, Fishery exclusion zones, No-take zones, Marine management, Marine protected areas, Marine reserves, Marine spatial planning, Maritime planning, MPAs, Policy evaluation

Background

The world's marine resources have substantial environmental, social and economic value [1, 2]. Human uses of the ocean are diverse, ranging from recreational and tourism activities and cultural heritage, through to more extractive uses such as harvesting, dredging, mining, and energy generation. Anthropogenic degradation of marine ecosystems is widely accepted as a major social-ecological problem that undermines the ability of the ocean to provide fundamental ecosystem services, such as food production, protection of shorelines from storms, climate regulation, leisure and recreation value and spiritual enrichment [e.g. 2, 3, 4]. Addressing this, the United Nations have declared the Decades of Ocean Science and Ecosystem Restoration to facilitate improvements in ocean health, and governments around the world have set out a shared vision to sustainably manage, protect and restore marine ecosystems. International commitments, including the Convention on Biological Diversity and the United Nations Sustainable Development Goal 14 [5, 6], include a requirement to designate “effectively and equitably managed, ecologically representative and well-connected systems of protected areas and other effective area-based conservation measures... integrated into the wider... seascapes” [5–7].

Together with the growing urgency to protect the value of marine environments, these commitments have led to the increasing application of spatial management measures in marine areas [e.g. 8–10]. In essence, the aim of spatial management is to incorporate the diversity of human uses, consider the compatibility of different activities, and balance uses with the impacts of these activities on biodiversity and people [11]. Spatial management measures typically comprise of marine protected areas (MPAs) for biodiversity conservation, sectoral (e.g.

fishery) closures to mitigate harmful effects and encourage sustainability, and marine spatial plans (‘marine plans’) to integrate social, economic and environmental considerations into proactive management of marine activities for multiple sector and stakeholders. Many countries have already invested substantially in developing an extensive array of marine spatial management measures. For example, the UK has designated 38% of UK domestic waters as MPAs [12]; has already adopted, or is in the process of developing, a series of national, devolved and regional marine plans [13]; and has implemented several seasonal fishery closures [e.g. 14].

Employed effectively, marine spatial management measures can provide a plethora of ecological, social and economic benefits [15–18], and there has been much work aimed at understanding what effects different spatial management measures have had, to what extent, and the reasons for these outcomes [17, 19, 20]. Such studies can inform the appropriateness of different management options in specific contexts. Yet, initial designation of spatial management is just the first step; achievement of objectives relies upon effective implementation, monitoring, evaluation, and adaptation [21, 22]. Effective monitoring, in particular, is fundamental to document the status of the environment and the activities that occur within it, which in turn informs both the assessment of impacts, including attribution and/or contribution, and the effectiveness of management. The assessment process enables an understanding of the strengths and weaknesses of spatial management, and allows for appropriate adaptation of management measures and policy development. However, despite spatial management being a core component of the global marine management portfolio the multifaceted complexity of the marine

environment, human uses and resultant impacts makes monitoring and assessing the effectiveness of spatial management an ongoing challenge [21].

With the increased application of marine spatial management measures, governments are now asking how they can effectively and efficiently monitor the marine environment to assess the impacts of such spatial management measures [e.g. 23, 24]. Deciding what to monitor and evaluate, how, and how often, is not straightforward, and the choice of approach can have implications for costs, efficacy, replicability, and robustness to challenge [25, 26]. Furthermore, the effectiveness of spatial management measures is affected by decisions made before establishment (e.g. consultation, co-creation with stakeholders, design, and location) and after legal designation (e.g. management changes, continued stakeholder engagement, adequate capacity), as well as by environmental, ecological and social contexts and surrounding uses of marine space [27–30]. Evaluations can therefore be conducted across a plethora of factors that impact effectiveness, including: design (e.g. extent, location, representation of habitats/species etc.); management (e.g. decision processes, capacity etc.); and, outcomes (e.g. what is achieved through management) [27]. Evaluations across each of these aspects are useful for different contexts and to answer different questions. For example, evaluations of spatial management design ask whether sites are best placed to achieve established goals, evaluations of management ask how well processes to support objectives are set up and working, while focusing on outcomes asks what effects the sites are having on the environment and people. Ultimately the entire context, including all these aspects, will collectively determine how effective a spatial management measure is. However, evaluations are often focused on particular aspects, and the choice of what aspect of a spatial management measure to evaluate depends on the priorities of those asking how effective that measure is [31].

The choice of what to monitor, and how, needs to be informed by a defined evaluation process with

established goals. However, there is considerable uncertainty as to what evaluation can, and should be, undertaken. Compounding these challenges is the need to improve understanding as to how seasonality can be captured within monitoring and evaluation programmes and how to assess the effects of real-time closures [23]. The vast array of published literature, coupled with the time and resource limitations facing government organisations and agencies, means that maintaining an up-to-date and comprehensive understanding of monitoring and evaluation options is unfeasible. Thus, understanding what methodologies are available, and how they are being applied, to monitor and evaluate spatial management effectiveness is critical to ensure cost-effective management and identify future research priorities to inform and improve management.

Here, we report on the results of a systematic map designed to inform this evidence need. The map focuses on evaluation approaches and analytical and data collection methodologies employed to conduct evaluations of the effectiveness of marine spatial management measures across ecological, social and economic outcomes [32], rather than design or management processes. The map collates evidence from coastal countries identified by our Stakeholder Group as being relevant to the UK (see “Stakeholder engagement” and “Search for articles” sections and Table 1 for more information). Using this systematic map, we explore what evaluation approaches and analytical and data collection methodologies are available, which are used in different contexts and which, if any, are more commonly applied. This study therefore builds on previous systematic maps and reviews on the effectiveness of marine protected areas [20], protected areas more broadly [33], and systematic conservation planning [19], by collating evidence related specifically to monitoring and evaluation of outcomes from spatial management measures. In doing so, we have sought to develop understanding, rather than assess the monitoring and evaluation approaches themselves. However, while outside our scope, we acknowledge the importance both of effectively

Table 1 Countries of interest for the systematic map

Australia	Greece	Portugal, including the Azores
Albania	Greenland	Slovenia
Belgium	Iceland	Spain, including the Canary and Balearic Islands
Canada	Ireland	Sweden
Croatia	Italy	United Kingdom
Denmark, including the Faroe Islands	Montenegro	United States of America
Finland	Netherlands	
France	New Zealand	
Germany	Norway	
	Poland	

Overseas or dependent territories of each country are excluded from the scope of the map except where specified

and equitably designing accountable monitoring, evaluation and learning systems (e.g. choosing appropriate indicators, outcome measures of interest, context-relevant methods for collecting data, definition of desired outcomes and interpretation of data) and evaluating these systems [26, 31, 34]. Further research into the design, implementation and evaluation of the monitoring, evaluation and learning systems themselves, focused on both short- and long-term outcomes, will therefore continue to be an important research agenda [e.g. 31]. Within this map, however, by explicitly exploring the methodology behind existing studies which aim to document effects and effectiveness of management, we provide a resource for (1) researchers to help in determining priorities for future research and (2) decision-makers to help inform discussions regarding the design of appropriate methodologies to incorporate into future monitoring and evaluation plans for marine spatial management.

Stakeholder engagement

The topic and question for this systematic map were originally proposed by the Review Team and co-developed with a Stakeholder Group, composed of key stakeholders from UK institutions involved in the monitoring and management of the marine environment, including: Marine Scotland Science (MSS), Natural Resources Wales (NRW), Department of Environment, Agriculture and Rural Affairs (DEARA, Northern Ireland), Inshore Fisheries and Conservation Authorities (IFCAs), Department of Environment, Food and Rural Affairs (Defra, England/UK), Centre for Environment, Fisheries and Aquaculture Science (Cefas), Joint Nature Conservation Committee (JNCC), Scottish Natural Heritage (SNH) and Natural England (NE). The Stakeholder Group has a diverse breadth of expertise, covering the array of disciplines needed for this systematic map, and extensive experience regarding evidence gaps facing UK and devolved governments (i.e. subnational bodies to which some powers have been delegated from central government—in the UK's case devolved governments refers to the Scottish Government, Welsh Government and Northern Ireland Executive). Involvement of a broad group of stakeholders enabled diverse perspectives to be represented and ensured utility of the resultant map to policymakers. Discussions were held remotely with stakeholders during protocol development and a face-to-face workshop was held at the University of Salford on 22nd February 2019 with representatives from the majority of the stakeholder organisations and the Review Team. These engagement activities were designed to formulate and agree the primary and secondary review questions, search strategy, eligibility criteria, and meta-data to be recorded. The Stakeholder Group were not involved in

the conduct of the review; however, a second workshop was held at the University of Salford, on 28th January 2020 to present the review findings and to discuss and agree appropriate mechanisms to disseminate findings to end-users more broadly.

Objective of the review

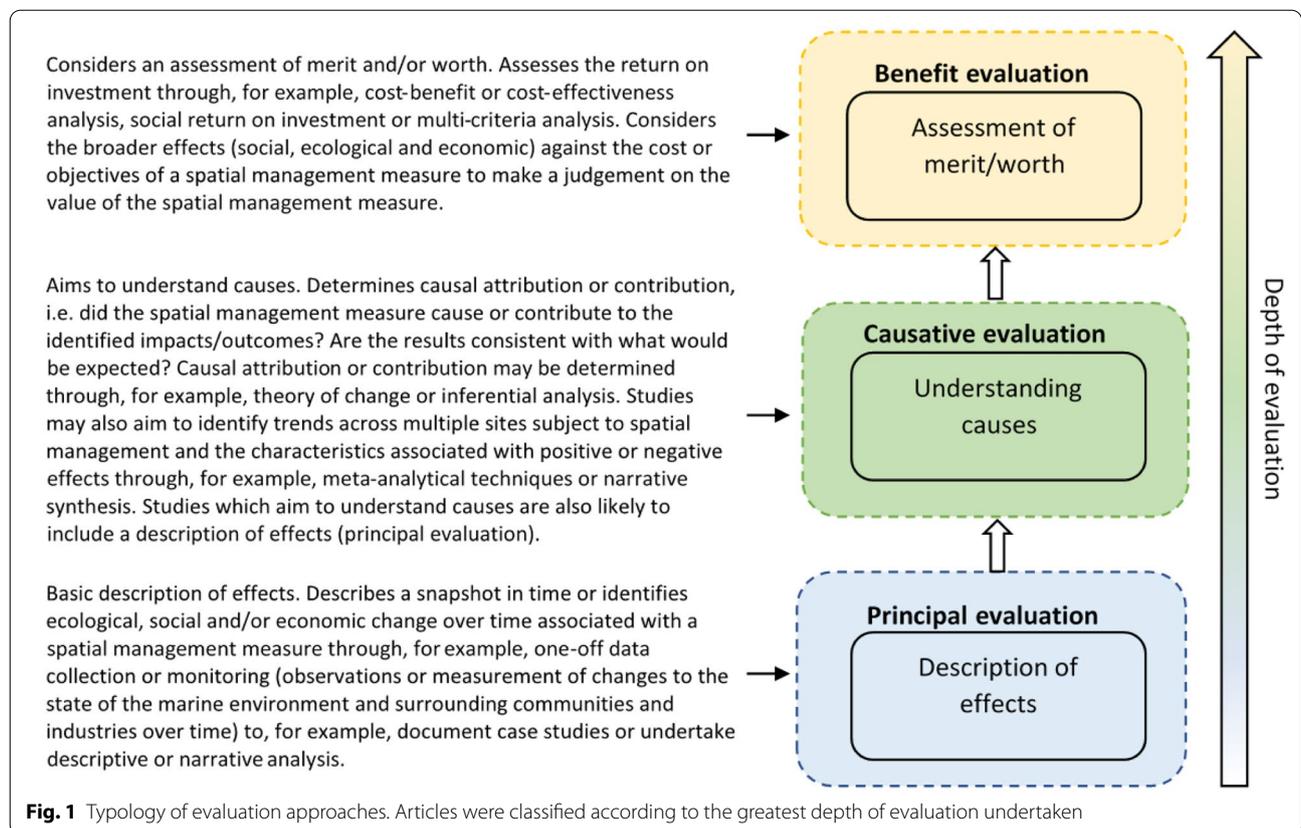
The primary research question for this systematic map was: What is the nature and extent of evidence on methodologies for monitoring and evaluating marine spatial management measures? This question has the following components [35]:

- *Population*: areas under marine spatial management in UK and similar coastal waters
- *Intervention*: monitoring and evaluation methodologies
- *Comparator*: none
- *Outcomes*: ecological, social and/or economic outcome measures of interest.

This review identified and collated retrospective studies that monitored the effects and evaluated the effectiveness of marine spatial management measures across ecological, social and economic outcomes. By 'monitor' we refer to methods applied to observe and measure changes to the state of the marine environment and surrounding communities and industries over time. Monitoring is considered to underpin evaluation. By 'evaluation' we refer to methodologies for collating and analysing data to determine the effects (the change arising from an intervention) or effectiveness (the degree to which something is successful in producing a desired result) of an intervention against its objectives and/or the resources. We define 'evaluation' according to three types, 'principal', 'causative' and 'benefit' which are based on the depth of evaluation undertaken (Fig. 1). Studies were coded to these categories according to the full depth of evaluation undertaken by the article, e.g. an article categorised as a 'causative evaluation' is also likely to include a 'principal evaluation'. Articles that reported monitoring of a site over time after implementation of management and without any evaluation being undertaken (e.g. without direct linking of observed data to effects of management) were considered to be 'monitoring programmes'. Monitoring programmes were excluded from this systematic map given our focus on evaluations of effectiveness and the monitoring methods used to inform these.

We define spatial management as:

- marine protected areas (MPAs)—“a clearly defined geographical space, recognised, dedicated, and



managed [...] to achieve the long-term conservation of nature with associated ecosystem services and cultural values” [36];

- fishery closures—an area within which fishing by one or more methods, or for particular species, is prohibited on a permanent, seasonal or real-time basis for the purpose of delivering fishery benefits [37]; and
- marine spatial planning (‘marine plan’)—an integrated multi-sectoral plan that informs the current and future distribution of activities in space to maintain delivery of ecosystem services in a way that meets ecological, economic and social objectives [38].

The evidence base was categorised using a predefined data coding framework [32] designed to explore the following secondary questions:

- What evaluation approaches and analytical methodologies have been used to evaluate the ecological, social and economic effectiveness of spatial management measures? What types of outcomes are measured? What data collection methods are used to gather these?

- What evaluation approaches and analytical methodologies have been used to understand the effects/ effectiveness of spatial management measures as networks as well as individual sites?
- What evaluation approaches are being applied by coastal countries to assess spatial management?

Methods

This systematic map was conducted according to the peer-reviewed protocol [32] and followed the Collaboration for Environmental Evidence Guidelines and Standards for Evidence Synthesis [35]. The mapping methods conform to the RepORting standards for Systematic Evidence Syntheses (ROSES) for systematic maps [39] (Additional file 1).

Deviations from the protocol

The methods used to conduct this systematic map followed those described in the published protocol [32] revised to reflect updates. In summary, updates comprise:

- Dual screening of articles at title-abstract and full text was undertaken. Disagreements at title-abstract were screened by a third reviewer. Disagreements at full text were discussed and resolved through consensus;
- The statement on procedural independence has been updated to reflect that articles co-authored by two members of the Review Team responsible for screening were selected for inclusion in the subset of articles for title-abstract and full text consistency checking. The protocol for decision-making in these instances is provided together with a description of comments made on one of these articles.
- Dual coding of included articles. Disagreements were discussed by the Review Team;
- An additional organisational website, the Great Barrier Reef Marine Park Authority, was searched;
- Text describing the content of each coding category was revised to improve clarity and ensure consistent application by the Review Team;
- As specified in the protocol, iterative coding of outcome measures of interest and data collection methods was conducted. This resulted in the addition of: two ecological (behaviour; nutrient capacity) and three social (compliance; displacement; conflict) outcome measures; and one social (direct non-extractive sampling) and two economic (experimental fishing; participant observation) primary data collection methods;
- Coding options for six columns were revised to improve clarity and/or reflect included studies (type of population; seasonality of spatial management measure; duration of spatial management measure (years); primary data collection duration; reference spatial management measure(s); evaluation data timeframe);
- Countries of interest were clarified to exclude overseas or dependent territories, except where specified, to reflect discussions during the first stakeholder workshop as to coastal countries they considered to be of UK-relevance in a social, political or ecological context (see Table 1);
- Meta-data relating to management information was recorded for all studies regardless of the number of sites studied;
- Given the limited evidence base identified for many of the outcomes recorded, cut-off points to identify boundaries (number of studies) at which a topic will be considered as either lacking evidence and therefore being poorly studied, or as having sufficient studies to allow for more meaningful exploration of the monitoring and evaluation methodologies they employ, were not utilised. Instead, a percentile colour gradient reflecting all evidence was used.

Search for articles

This systematic map was based on literature searches conducted in June 2019 using four bibliographic databases accessed using institution subscriptions from the University of Salford (UoS) or the University of York (UoY): (1) Web of Science Core Collections (UoS subscription consisting of the following indices: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, and ESCI); (2) Scopus (UoS); (3) Aquatic Sciences and Fisheries Abstracts (UoY); and (4) Directory of Open Access Journals (UoY). See Additional file 2 for full search strings used on each website and the date of visit. We searched one search engine, Google Scholar, and extracted the first 200 search results as citations using Publish or Perish software [32]. To test the comprehensiveness of the search strategy a scoping search was carried out with results from iterations of the search string compared against an a-priori defined test library with 15 articles of known relevance during protocol development—all 15 articles were located using the final search string [32]. Articles retrieved from bibliographic databases and Google Scholar were combined into a single Endnote library. Duplicates were removed prior to screening.

Searches were also performed between June and October 2019 across 39 relevant organisational websites and one data repository to capture grey literature. In addition, bibliographic searches of all identified relevant tertiary review articles were undertaken. Authors of articles that could not be located were contacted directly to request a copy.

Databases and websites were searched using English language search terms (Additional file 2). Searches were restricted to articles published between 2009 and 2019 to increase relevance to the UK marine policy landscape [40] and to reflect the recent increase in application of marine spatial management measures (particularly MPAs and marine plans) [9, 10].

Article screening and study eligibility criteria

Screening process

Articles were assessed for inclusion according to a hierarchical assessment of relevance: screening articles first at title and abstract concurrently, followed by the full text of potentially relevant articles. Only articles published in English were considered; however, all returned non-English articles whose titles and abstracts were available in English that passed title-abstract screening were retained for potential use in future studies (Additional file 3).

Retrieved literature from websites, supplementary searches and the Stakeholder Group was screened separately to those retrieved from bibliographic databases and Google Scholar; articles deemed relevant at full text

were combined with other records prior to compilation of the systematic map.

In deviation from the published protocol, but in line with best practice described in the CEE guidelines [35], the decision was taken to independently dual screen all articles at title-abstract and full text level due to additional resources from new members joining the Review Team. Nonetheless, prior to reviewers' commencing screening, consistency checking was performed using a random subset of 10% of articles at each stage ($n=1282$ and $n=81$ respectively) and, where the level of agreement was below 0.6 according to a kappa test, further consistency checking was conducted on an additional set of articles. Three reviewers initially undertook consistency checking at title-abstract level however a new member then joined the Review Team and a second round of consistency checks was undertaken with four reviewers. Only three of these had sufficient agreement and resources to participate in title-abstract screening, which then began. However, following this another reviewer became available to participate in title-abstract screening and, prior to doing so, undertook consistency checking using the same second sample as the other reviewers. Four reviewers concurrently undertook consistency checking at full text level. All disagreements during consistency checking were discussed in detail. Disagreements at title-abstract level were discussed amongst the three reviewers during initial consistency checking, amongst the three reviewers with sufficient levels of agreement during the second round of checks because the fourth was unable to attend the team meeting or participate further in title-abstract screening due to resource constraints, and between one of these reviewers and the last reviewer later. Members of the Review Team that had authored, or co-authored, articles identified as potentially relevant referred these to another reviewer for assessment during screening. However, articles co-authored by two members of the Review Team responsible for screening were randomly selected for inclusion in the subset of articles for title-abstract ($n=2$) and full text ($n=2$) consistency checking. In all instances, decisions taken by the other, non-author, reviewers were applied and the articles were removed from kappa tests. Of these, only one of the articles included in the title-abstract sample was subject to disagreement and therefore discussed amongst the Review Team. In this instance, the Review Team agreed to include the article through to full text screening and, while the author abstained from the assessment, they did comment during discussions that they agreed with its exclusion. Following consistency checking, disagreements at title-abstract level were screened by a third reviewer with articles considered unclear taken forwards

($n=587$, 4.6% of articles); disagreements at full text were discussed and resolved through consensus.

Eligibility criteria

Articles were screened according to the following criteria:

Relevant population(s): Areas under implemented marine spatial management (fishery closures, MPAs, marine plans) restricted to the identified geographical locations (Table 1). Proposed spatial management measures were not considered. Large areas (regions, provinces or exclusive economic zones) where broader legislation protects certain species were excluded from the definition of MPA or fishery closure. Studies with their primary focus on freshwater and/or terrestrial environments were excluded.

Relevant intervention(s): Monitoring and evaluation methodologies employed to assess effectiveness (Fig. 1). Articles assessed as being 'monitoring programmes' were excluded (see "Objective of the review" section).

Relevant comparator interventions: None. Studies were not required to have a comparator intervention for inclusion.

Relevant study designs: Ecological studies were required to contain multiple reference sites or a time-series of data to warrant inclusion in the systematic map (as opposed to snapshot studies only inside a managed area). Social and economic studies were not required to have a specific study design. Elements relating to study design (e.g. time-series of data and details of reference sites) were, however, recorded across ecological, social and economic studies to enable further understanding of evaluation methodologies across different fields of study. Theoretical studies (including predictive modelling studies) and commentary articles or opinion pieces were excluded.

Relevant outcome(s): Any ecological, social and/or economic outcome(s) reported by studies. As the focus of the systematic map was on outcomes, studies related to governance or designation process (e.g. administrative, political, legal, planning or design activities) were excluded. Studies focusing on environmental parameters (e.g. water quality, sediment, etc.) were excluded from the definition of ecological outcomes.

Our aim was to provide a resource for decision-makers, while describing the evidence base. Therefore, we included both primary (i.e. generation of new data from either field or existing data [e.g. 41]) and secondary (i.e. literature that consists of analytical interpretations and evaluations that are derived from primary source literature [e.g. 17]) literature; however, these were placed into separate databases for coding given the different level of meta-data that could be coded from these articles and are reported separately. Studies which report large-scale

regional or global evaluations of relevant spatial management measures, that included countries of interest, were included in the systematic map. Tertiary literature (i.e. broader literature reviews that consist of a distillation and collection of primary and secondary sources but contain no new analysis [e.g. 42]) were recorded separately to act as a resource for end-users. Given the lack of new analytical interpretations or evaluations of effectiveness in tertiary reviews, and the inclusion of relevant primary and secondary literature identified from bibliographic screening of tertiary reviews in the two systematic maps, these were not coded.

A list of articles excluded at full text with reasons for exclusion is provided in Additional file 3.

Study validity assessment

Given the broad scope and size of this systematic map, the validity of articles was not assessed. However, elements of study design that might relate to validity (e.g. presence of a reference site, evaluation data timeframe) were coded to provide a basic overview of study methodology. Studies were recorded as using data before or after implementation of the spatial management measure being evaluated (either designation or change in regulations), however the data studies used could be accessed through primary and/or secondary sources and be either single data points or a time-series. Whether a study used a control site (site outside of spatial management) or another area under spatial management (either a different designation or zones of different regulation within the same designated area) as a reference site to evaluate against was also recorded. However, while some studies will have directly compared their data and evaluation to a reference site using standardised sampling strategies, others will have used data from different sources. Consequently, such studies are not necessarily true before-after-control-impact studies and so while this coded information provides some indication of study design we did not consider it as part of the methodology or research design employed by the evaluation. No studies were excluded from the systematic map database based on these extracted data.

Data coding strategy

Meta-data, information describing each study, was extracted from each article considered to be relevant at full text review and recorded using the systematic map spreadsheet as a standardised coding tool (Additional file 4). Data coding was conducted concurrently with full text screening. Reviewers coded relevant articles in separate versions of this spreadsheet which were

then combined during consistency checking (see below). Missing or unclear information was recorded as such. All coding was documented in a systematic map database, with each line representing one study outcome measure of interest (i.e. each independent outcome measure considered by each study). Multiple studies reported within one article were, therefore, entered as independent lines in the database. Distinct primary articles that report the same study outcome measure of interest based on the same dataset as a study published in an earlier article (including those where the dataset had been expanded) were linked in the database, where identified.

The following main categories of data were extracted:

- Bibliographic information (i.e. unique Article ID, full citation, publication location [e.g. journal article, book chapter etc., publication type [grey or scientific], unique study ID, publication year]).
- General information about the spatial management measure studied (i.e. type [MPA, etc.], seasonality [year-round, temporary etc.], duration [<3 years, 3–10 years, etc.], scale [single site, network etc.], size [km²], number, name(s) of spatial management measures; general description of location, country of study, continent, latitude and longitude; any information on management relating to fishing and other activities, study environment [inshore, offshore etc.]; focus of article evaluation [ecological, social, economic or a combination]; focus of study outcome measure [ecological, social or economic]).
- Monitoring and evaluation methodologies, including outcomes evaluated (i.e. source and type of data used in evaluation, duration of primary data collection duration; evaluation typology [principal, causative or benefit], primary evaluation method used [e.g. narrative analysis, descriptive analysis etc.], type of reference spatial management measure used [e.g. control site outside spatial management, zones within MPA etc.], evaluation data timeframe [after/before-after implementation]; outcome measure(s) of interest, primary data collection method(s) employed [e.g. indirect/direct user surveys, direct non-extractive sampling etc.], secondary data source(s) used).

Full details of all coding categories are provided in Additional file 4. Meta-data extraction was performed by three reviewers independently, such that each article was coded by two reviewers. One of these reviewers was then responsible for combining coded spreadsheets and consistency checking across the whole database. Before full data coding commenced, consistency checking was

undertaken for coding of a subset of 100 studies. All disagreements were discussed, and coding categories refined, prior to coding the remaining full texts (see “[Deviations from the protocol](#)” section). Following this, any uncertainties and issues that arose during the data extraction process were flagged by the reviewer and discussed and resolved by the Review Team in regular meetings.

Data mapping method

The evidence base identified within this systematic map was described narratively using descriptive statistics and within the systematic map database, a searchable spreadsheet of studies and related coding results (Additional file 4). Framework based synthesis, a matrix-based approach that supports construction of thematic categories into which data can be coded and analysed [43], was used to identify knowledge clusters and gaps. These were identified by cross-tabulating key variables and quantifying the number of articles and/or studies as a proxy for extent of evidence. Studies from relevant primary and secondary literature were reported separately.

Review findings

Review descriptive statistics

In total, 19,515 results were retrieved from searches across bibliographic databases and Google Scholar, including 6708 duplicates (Fig. 2). Most (12,006) articles were excluded at title and abstract screening due to irrelevance. 801 articles were screened at full text, of which 353 primary articles and 32 secondary articles were included (see Additional file 3 for exclusion reasons). 24 (3.1%) articles could not be found or accessed (Additional file 3). A total of 39 articles were included from stakeholders, organisational websites and bibliographic searches of relevant reviews (Additional file 2). Ultimately, 391 primary articles were included in the final map (full bibliography in Additional file 3) which generated 858 studies reporting monitoring and evaluation methods relevant to the review (Additional file 4). A further 33 relevant secondary articles were included in the second map (full bibliography in Additional file 3) which generated 63 studies. In total, 49 relevant tertiary review articles were identified, 36 of which were retrieved from bibliographic databases and Google Scholar, with the remainder from supplementary searches (Additional file 3).

Of the 424 included primary and secondary articles, 89.4% were articles published in scientific peer-reviewed journals, with the remainder being reports (7.3%) and theses (3.3%). The volume of articles published over time was variable across both primary and secondary literature (Fig. 3).

Mapping the quantity and quality of studies relevant to the question

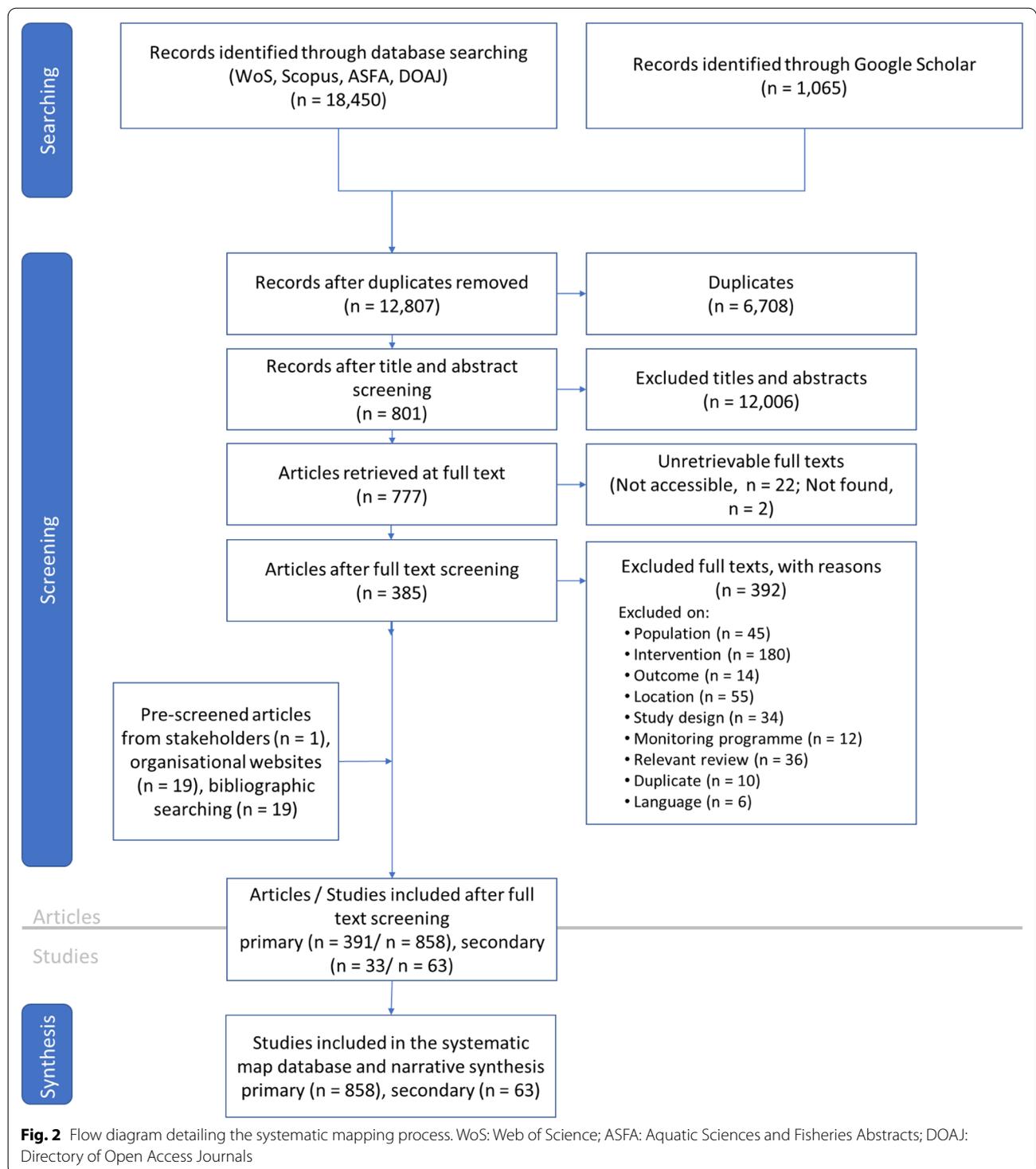
Spatial management measures evaluated (population)

Of the 391 relevant primary articles identified, the majority (85.7%) focused on marine protected areas with only 12.5% of articles on fisheries closures, 0.5% on marine plans and 1.3% on both MPAs and fishery closures. Most primary articles evaluated single sites (60.4%) or multiple individual sites (32.5%) with only 6.9% considering networks and 0.3% marine plans, as described by study authors (Table 2). Secondary articles ($n=33$) predominantly focused on multiple individual sites (93.9%) with the remainder considering single sites (Table 2). Most articles (76.7% and 84.4% primary and secondary articles respectively) did not explicitly state the seasonality of the spatial management measure they considered (Table 2), or provide details of regulations/restrictions for activities other than fishing within the managed area(s) (80.8% and 100% primary and secondary articles respectively). Although the latter may be because activities other than fishing were not regulated within the spatial management measure being studied rather than a reporting bias, these studies provided no information to this effect making it impossible to establish this. Regulations and/or restrictions in place for fisheries were mentioned in 80.8% and 72.7% of primary and secondary articles, respectively. The majority (41.9%) of primary articles evaluated sites older than 10 years at the time of their assessment, while for secondary articles most articles (45.6%) did not specify the age of sites (Table 2).

Primary articles mainly evaluated spatial management measures in inshore waters (<12 nautical miles from shore), while secondary articles did not generally report study environment (Table 2). The majority of evaluations were conducted in Europe (43.2%, $n=169$) or Oceania (31.7%, $n=124$), followed by North America (22.8%, $n=89$). The dominant countries of study were Australia (27.4%, $n=107$), USA (18.4% $n=72$), and UK (11.3%, $n=44$) (Fig. 4). Secondary articles were mostly global (66.7%, $n=22$) or European (18.2%, $n=6$) in extent with 84.8% ($n=28$) of articles evaluating spatial management measures from multiple countries of interest.

Evaluation typologies and methodologies employed (intervention)

The majority (70.3%, $n=275/391$) of primary articles conducted principal evaluations, i.e. basic description of effects, with 29.4% ($n=115$) of articles exploring causation. Only one article undertook a benefit evaluation. Conversely, most secondary articles (66.7%, $n=22/33$) conducted causative evaluations with the remainder undertaking principal evaluations. No trend



over time was found in the depth of evaluations (evaluation typology) or the foci of article evaluation being published. Most primary articles focused on conducting principal and causative evaluations for ecological outcomes only (45.3% [n = 177/391] and 22.5% [n = 88]

respectively), with the evidence base for other foci of evaluations (social, economic and combinations of these) being much more limited across all evaluation typologies (Fig. 5). Similarly, over three-quarters (75.8%, n = 25/33) of secondary articles focused on

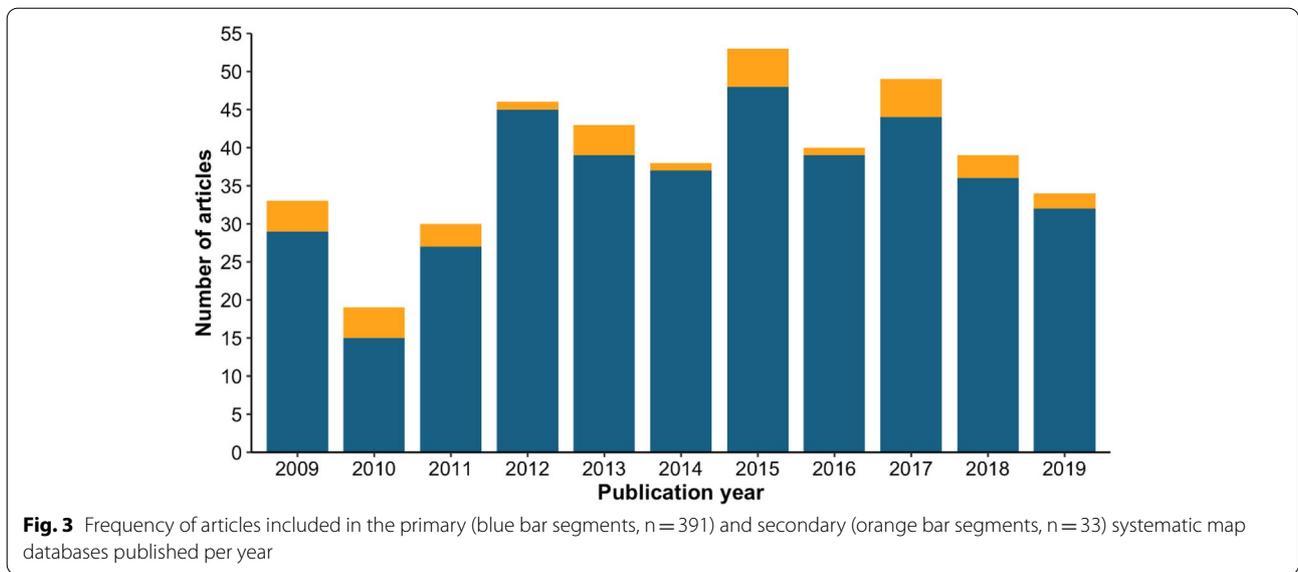
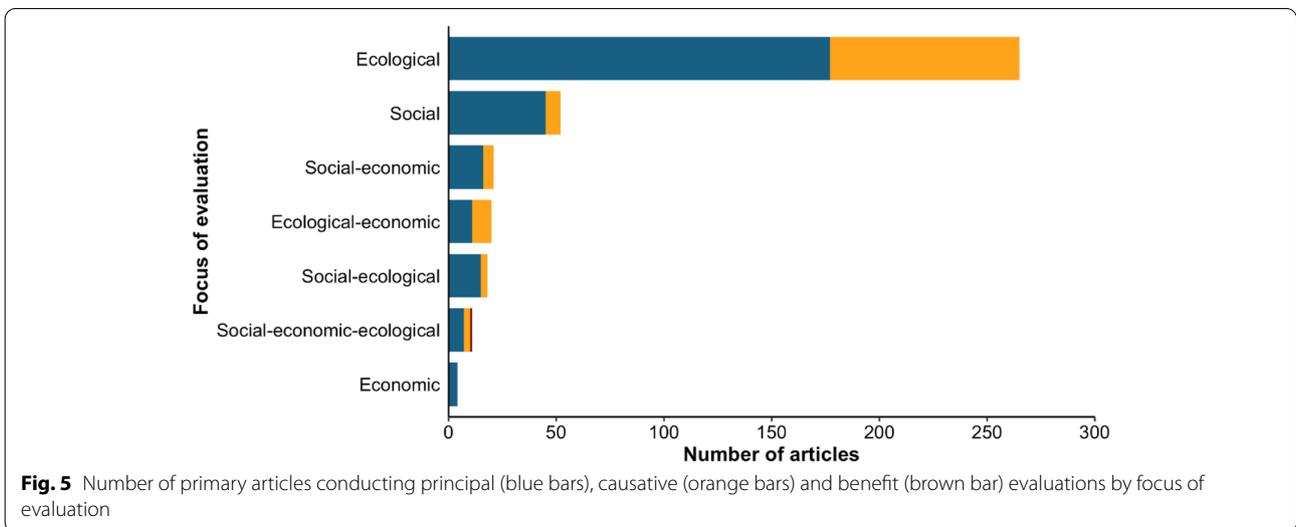
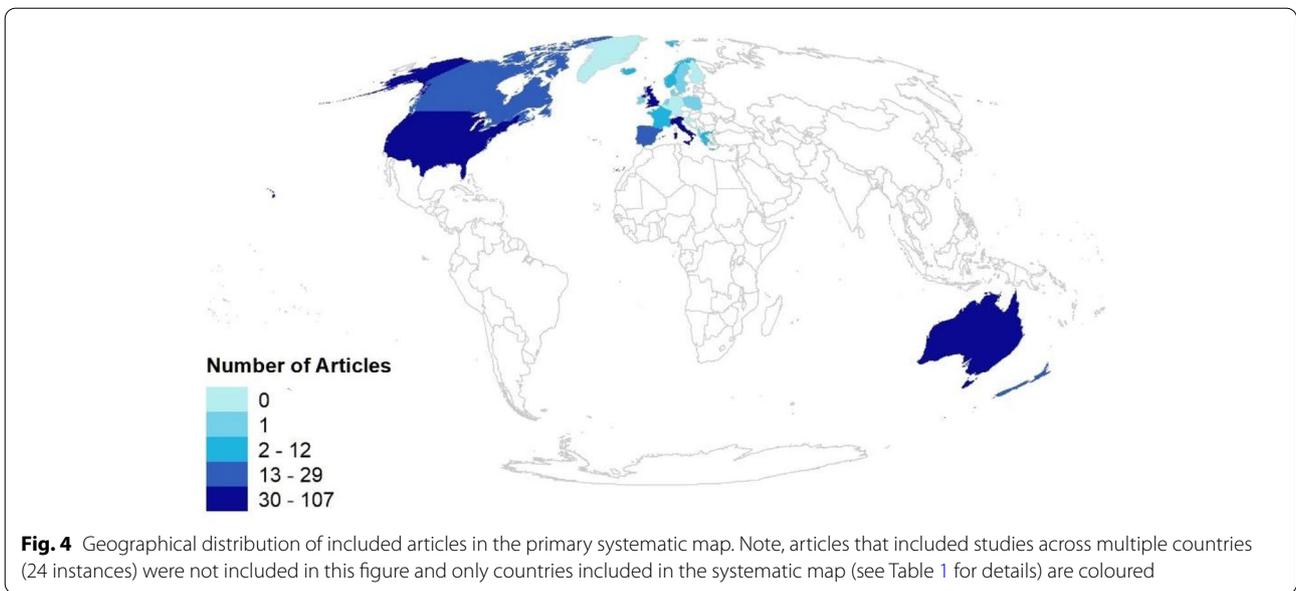


Table 2 Descriptive article statistics for the primary and secondary systematic maps

	Primary systematic map	Secondary systematic map
Scale		
Single site	236 (60.4%)	2 (6.1%)
Multiple individual sites	127 (32.5%)	31 (93.9%)
Network	27 (6.9%)	0 (0.0%)
Marine plan	1 (0.3%)	0 (0.0%)
Seasonality		
Not specified	236 (60.4%)	2 (6.1%)
Year-round	127 (32.5%)	31 (93.9%)
Mixed	27 (6.9%)	0 (0.0%)
Seasonal	1 (0.3%)	0 (0.0%)
Real-time	2 (0.5%)	0 (0.0%)
Temporary	0 (0.0%)	0 (0.0%)
Duration (years)		
< 3	6 (1.5%)	0 (0.0%)
3–10	102 (26.1%)	0 (0.0%)
> 10	164 (41.9%)	5 (15.2%)
Not specified	75 (19.2%)	15 (45.5%)
Mixed	44 (11.3%)	13 (39.4%)
Study environment		
Inshore	268 (68.5%)	2 (6.1%)
Offshore	22 (5.6%)	1 (3.0%)
Inshore & offshore	14 (3.6%)	2 (6.1%)
Estuary	11 (2.8%)	0 (0.0%)
Mixed	3 (0.8%)	1 (3.0%)
Lagoon	2 (0.5%)	0 (0.0%)
Not specified	71 (18.2%)	27 (81.8%)

evaluating spatial management measures according to ecological outcomes only.

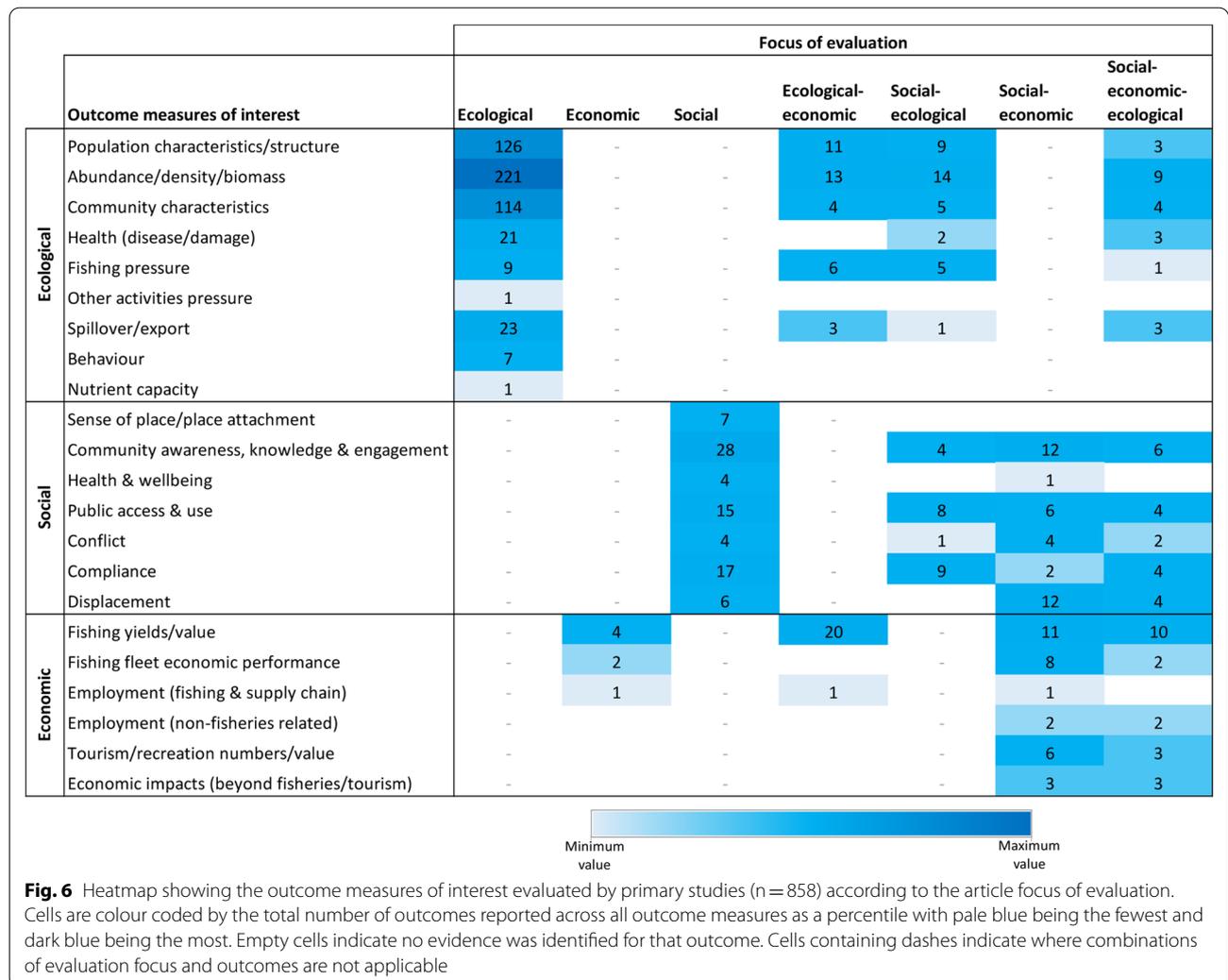
Analytical methodologies employed by articles across both primary and secondary literature were limited. Most principal evaluations (88.7% [n=244/275] primary and 63.6% [n=7/11] secondary) conducted descriptive analysis (statistically describing, aggregating, and presenting the constructs of interest or associations between these constructs). For primary articles, most causative evaluations used inferential statistics (98.3% [n=113/115], statistical testing of hypotheses/explanatory modelling) while most causative evaluations undertaken in secondary articles used meta-analytical techniques (81.8%, n=18/22). For ecological studies that reported collecting primary data (n=618), the most common data collection methods were direct non-extractive sampling (e.g. diver surveys, towed/drop-down video, observations from boats/shore, aerial photos: 63.3%) followed by extractive sampling (e.g. traps, towed fishing gear, grab samples: 22.3%). Social studies that collected primary data (n=154) mostly used primary data collected via direct user surveys (e.g. structured/semi-structured/unstructured interviews, focus groups, workshops: 65.6%), participant observation (e.g. remote observation of activities through, for example, cameras or field surveys: 18.2%) or indirect user surveys (e.g. online/postal questionnaires/surveys: 13.0%). Economic studies that collected primary data (n=63) did so mainly through direct or indirect user surveys (57.1% and 17.5% respectively) and experimental fishing (15.9%).



Measured outcomes

In the primary systematic map, articles (n = 391) evaluated spatial management measures across a total of 22 social, economic and ecological outcomes (Fig. 6); only 2.8% (n = 11) conducted evaluations across all three disciplines, with most focused exclusively on ecological evaluations (67.8%, n = 265) and fewer on social (13.3%, n = 52) or economic (1.0%, n = 4) evaluations. The most frequent ecological outcome measures used by studies to evaluate the effectiveness of spatial management measures in the primary literature were abundance/density/biomass (n = 221/858) followed by population

characteristics/structure (n = 126), and community characteristics (n = 114) (Fig. 6). These three outcome measures accounted for 62.1% of all (n = 858) studies included in the primary systematic map, and 86.1% (n = 533/619) of all studies evaluating against ecological outcome measures. Social outcome measures of interest were more evenly distributed across studies with most focusing on community awareness, knowledge and management (n = 50/160) followed by public access and use (n = 33) and compliance (n = 32). Economic outcomes were used the least to evaluate the effectiveness of spatial management measures with a total of 79

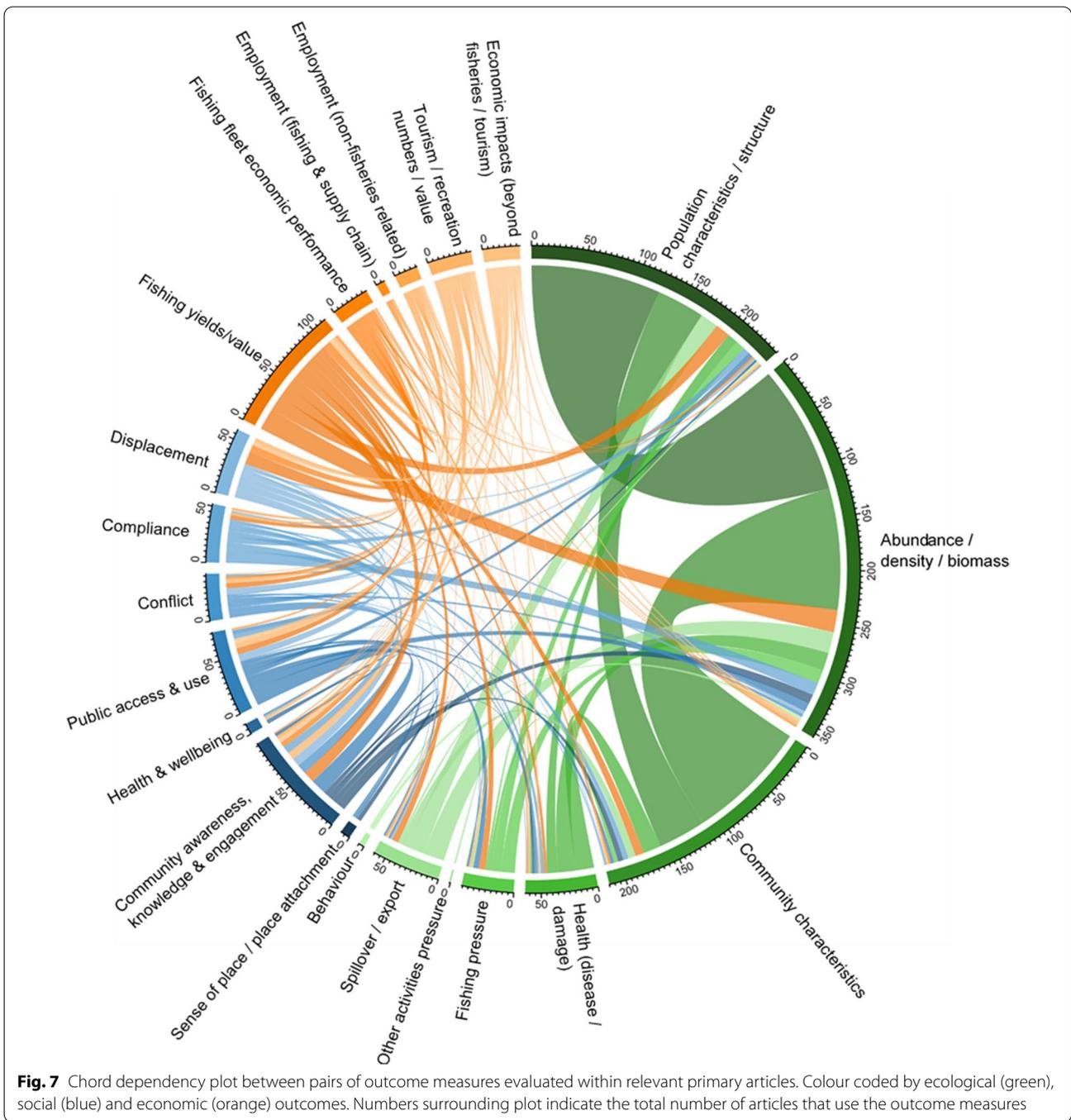


studies recorded. Fishing yields/value was the most frequently evaluated economic outcome, accounting for 57.0% (n = 45) of all economic studies.

On average, 2.2 outcome measures of interest (range 1–9) were evaluated in primary articles. The most commonly occurring pairs of outcomes evaluated across all articles were abundance/density/biomass with either population characteristics/structure (n = 122/858 studies) or community characteristics (n = 115), followed by population characteristics/structure and community characteristics (n = 45) (Fig. 7). Other outcomes were much less commonly paired. Articles evaluating social outcomes most commonly paired community awareness, knowledge and engagement with public access and use (n = 15), and those evaluating economic outcomes paired fishing fleet economic performance with fishing yields/value (n = 8). For articles that evaluated outcome measures across multiple foci, the most commonly occurring pairs of outcome measures studied

were fishing yields/value with abundance/density/biomass and population characteristics (n = 21 and n = 14 studies respectively).

In the secondary systematic map, articles evaluated spatial management measures across 16 social, economic and ecological outcomes; only two conducted evaluations across all three disciplines. As with primary articles, secondary articles evaluating the effectiveness of spatial management measures were dominated by ecological outcomes with abundance/density/biomass and community characteristics accounting for 57.1% of all (n = 36/63) studies and 75.0% of all (n = 22/48) ecological outcome measures studied (Fig. 8). Social and economic outcomes studied were each only represented by one or two studies. On average, secondary articles evaluated spatial management measures against 1.9 outcome measures of interest (range 1–6). As with primary articles, the most frequent outcome measures used together to evaluate the effectiveness of spatial

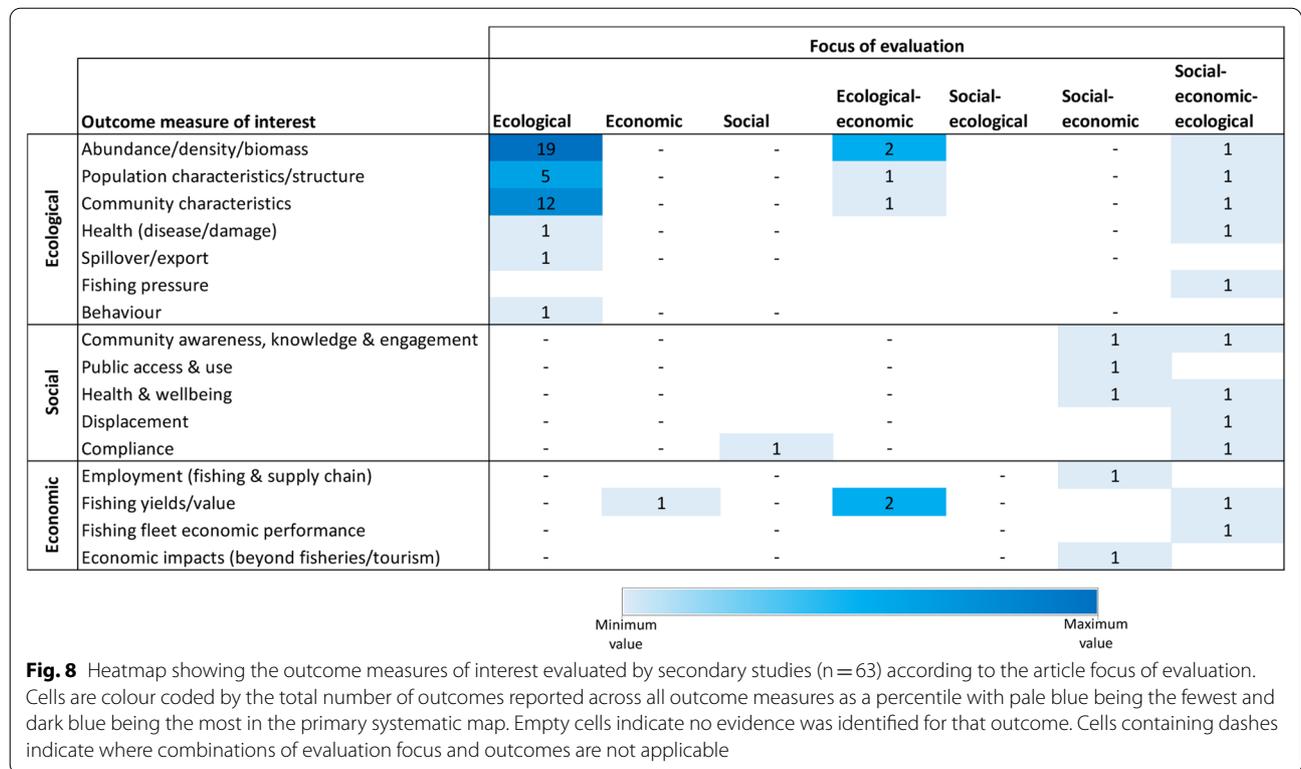


management measures were abundance/density/biomass with either community characteristics ($n = 10$) or population characteristics ($n = 7$), and population characteristics and community characteristics ($n = 5$).

Linking population with intervention and outcomes

Review findings in earlier sections have been structured by article type (primary and secondary) to reflect

the distinction between the two systematic maps and by population, intervention and outcome to reflect the components of our primary research question (see "Objective of the review" section). However, this map aims to provide an overview of the evidence base for evaluation approaches, analytical methods and data collection methods applied to inform evaluations of the effectiveness of different types of spatial management measures,



across a broad range of ecological, economic and social outcomes. Consequently, this section summarises the review findings by the type of spatial management measure (population) focusing on the primary systematic map.

Descriptive statistics by spatial management measure are presented in Table 3. As stated previously, evaluations of single or multiple individual MPAs in place for three or more years in inshore waters dominate the evidence base. The evidence base of articles evaluating networks of sites under spatial management is limited to 27 articles—most of which focus on MPAs (n = 21).

Marine protected areas Primary articles focused on evaluating the effectiveness of MPAs (n = 335) predominantly conducted principal (71.0%, n = 238), followed by causative (28.7%, n = 96), evaluations; only one (0.3%) conducted a benefit evaluation. Most articles focused on conducting evaluations for ecological outcomes (70.4%, n = 159 principal evaluations, n = 77 causative evaluations) or social outcomes (13.4%, n = 41 principal evaluations, n = 4 causative evaluations) only. Principal evaluations relied on descriptive analysis (89.1%, n = 212); causative evaluations on inferential analysis (97.9%, n = 94). A cost–benefit analysis was used by the article conducting a benefit evaluation.

In total, articles conducting principal evaluations of MPA effectiveness generated 517 studies, causative evaluations 208 studies, and benefit evaluations six studies.

Of these, abundance/density/biomass (28.6% of principal evaluation studies, 35.6% of causative evaluation studies, 11.1% of benefit evaluation studies), population characteristics/structure (16.8% and 18.8% of principal and causative evaluation studies, respectively) and community characteristics (15.1% and 16.3% of principal and causative evaluation studies, respectively) were the most frequently considered across all evaluation typologies. Direct non-extractive sampling was the most frequent method of data collection across all but one measured ecological outcomes; spillover/export was most frequently explored through movement/recapture studies (Fig. 9). The second most dominant method of data collection for exploring ecological outcomes was extractive sampling. Social outcomes were predominantly measured using direct user surveys, followed by participant observation and indirect user surveys. Economic outcomes were similarly explored using direct and indirect user surveys, although fishing yields/value was also often explored through experimental fishing. Evaluations were mainly conducted in Europe (44.2%, n = 148) and Oceania (33.7%, n = 113) followed by North America (20.0%, n = 67).

The evidence base of articles that evaluate networks of MPAs (n = 21) is more limited than that for multiple individual sites (n = 100) or single sites (n = 214). Mirroring the general patterns of the whole map (Tables 2, 3), of

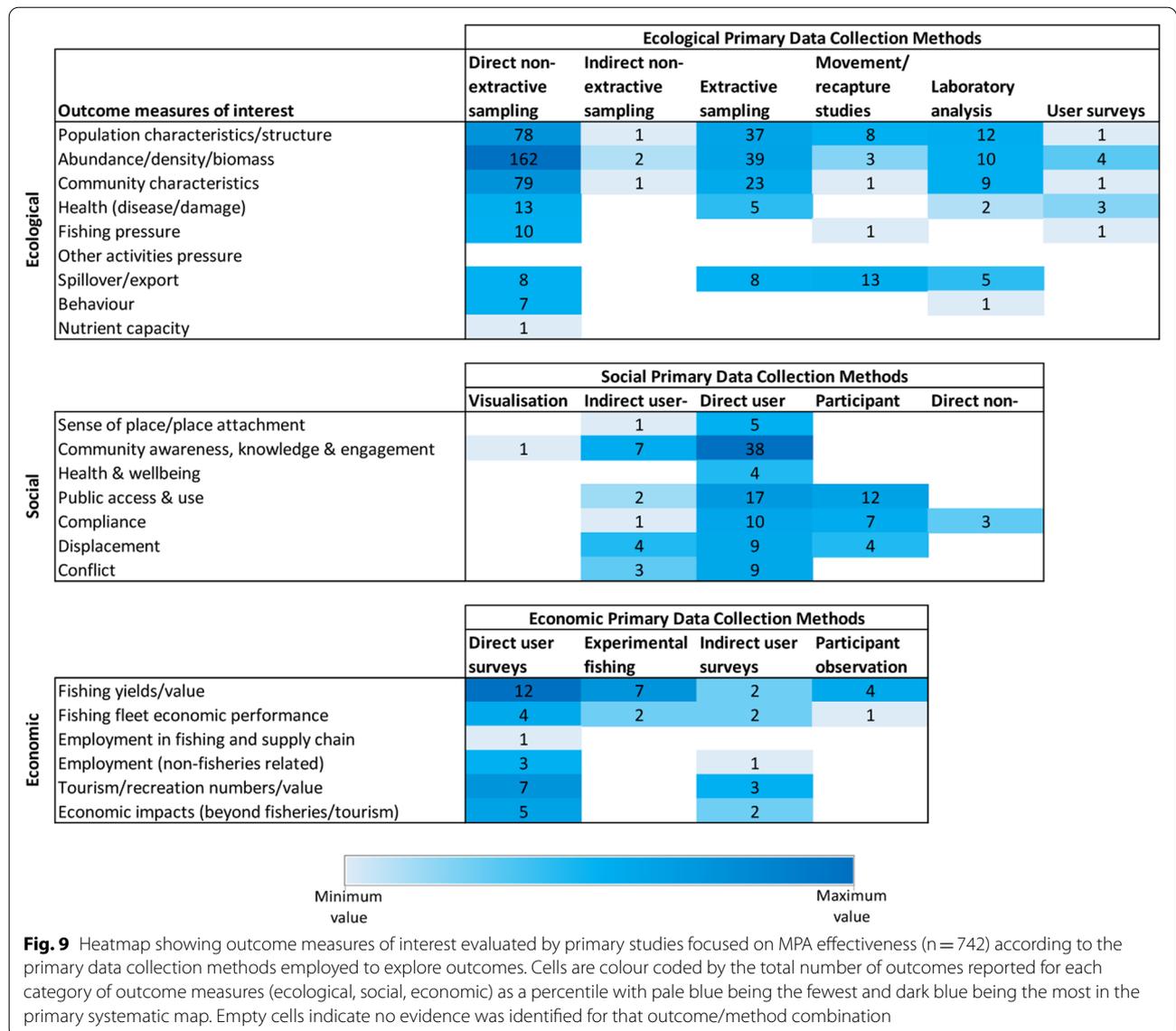
Table 3 Descriptive article statistics for the primary systematic map by spatial management measure (population)

	MPA	Fishery closure	Marine plan	MPA + Fishery closure
Scale				
Single site	214 (63.9%)	21 (42.9%)	1 (50.0%)	0 (0.0%)
Multiple individual sites	100 (29.9%)	23 (46.9%)	0 (0.0%)	4 (80.0%)
Network	21 (6.3%)	5 (10.2%)	0 (0.0%)	1 (20.0%)
Marine plan	0 (0.0%)	0 (0.0%)	1 (50.0%)	0 (0.0%)
Seasonality				
Not specified	272 (81.2%)	21 (42.9%)	2 (100.0%)	5 (100.0%)
Year-round	50 (14.9%)	15 (30.6%)	0 (0.0%)	0 (0.0%)
Mixed	10 (3.0%)	4 (8.2%)	0 (0.0%)	0 (0.0%)
Seasonal	3 (0.9%)	7 (14.3%)	0 (0.0%)	0 (0.0%)
Real-time	0 (0.0%)	2 (4.1%)	0 (0.0%)	0 (0.0%)
Temporary	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Duration (years)				
< 3	4 (1.2%)	2 (4.1%)	0 (0.0%)	0 (0.0%)
3–10	83 (24.8%)	17 (34.7%)	1 (50.0%)	1 (20.0%)
> 10	141 (42.1%)	22 (44.9%)	0 (0.0%)	1 (20.0%)
Not specified	71 (21.2%)	4 (8.2%)	0 (0.0%)	0 (0.0%)
Mixed	36 (10.7%)	4 (8.2%)	1 (50.0%)	3 (60.0%)
Study environment				
Inshore	236 (70.4%)	29 (59.2%)	1 (50.0%)	2 (40.0%)
Offshore	10 (3.0%)	12 (24.5%)	0 (0.0%)	0 (0.0%)
Inshore & offshore	10 (3.0%)	4 (8.2%)	0 (0.0%)	0 (0.0%)
Estuary	11 (3.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Mixed	3 (0.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Lagoon	2 (0.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Not specified	63 (18.8%)	4 (8.2%)	1 (50.0%)	3 (60.0%)

those articles that consider MPAs as part of a network, most ($n = 14$) conduct principal evaluation using descriptive analysis while the remainder ($n = 7$) use inferential analysis within a causative evaluation. MPA network evaluations predominantly evaluate MPAs established for between 3 and 10 years ($n = 11$) or for more than 10 years ($n = 6$), and almost all are conducted in inshore environments ($n = 16$). None specified the seasonality of MPAs. Evaluations were mainly conducted in North America ($n = 15$) with the remainder in Oceania ($n = 4$) or Europe ($n = 2$).

Fishery closures 49 primary articles evaluated the effectiveness of fishery closures, all of which conducted either a principal ($n = 31$) or causative ($n = 18$) evaluation. Foci of evaluations were predominately ecological only (51.0%, $n = 15$ principal evaluations, $n = 10$ causative evaluations), with the remainder of articles spread across other foci (social, economic or combinations of all). Principal evaluations mostly used descriptive analysis (90.3%, $n = 28$) while all causative evaluations applied inferential analysis.

Articles conducting principal evaluations of fishery closures generated 66 studies and focused on abundance/density/biomass (27.3%, $n = 18$), population characteristics/structure (21.2%, $n = 14$), fishing yields/value (15.2%, $n = 10$), with the remaining 24 studies split across ten other outcome measures. Articles that undertook a causative evaluation generated 40 studies with almost half of these focused on abundance/density/biomass (27.3%, $n = 18$) or population characteristics/structure (21.2%, $n = 14$). Direct non-extractive and extractive sampling were the most frequently applied methods of data collection across all but one measured ecological outcomes; spillover/export was only explored through movement/recapture studies (Fig. 10). Social outcomes were predominantly measured using direct user surveys or participant observation. Only one economic outcome—fishery yields/value—was considered using direct user surveys, experimental fishing and participant observation. Fishery closure evaluations were equally distributed across Europe and North America (each 40.8%, $n = 20$), followed by Oceania (18.4%, $n = 9$).

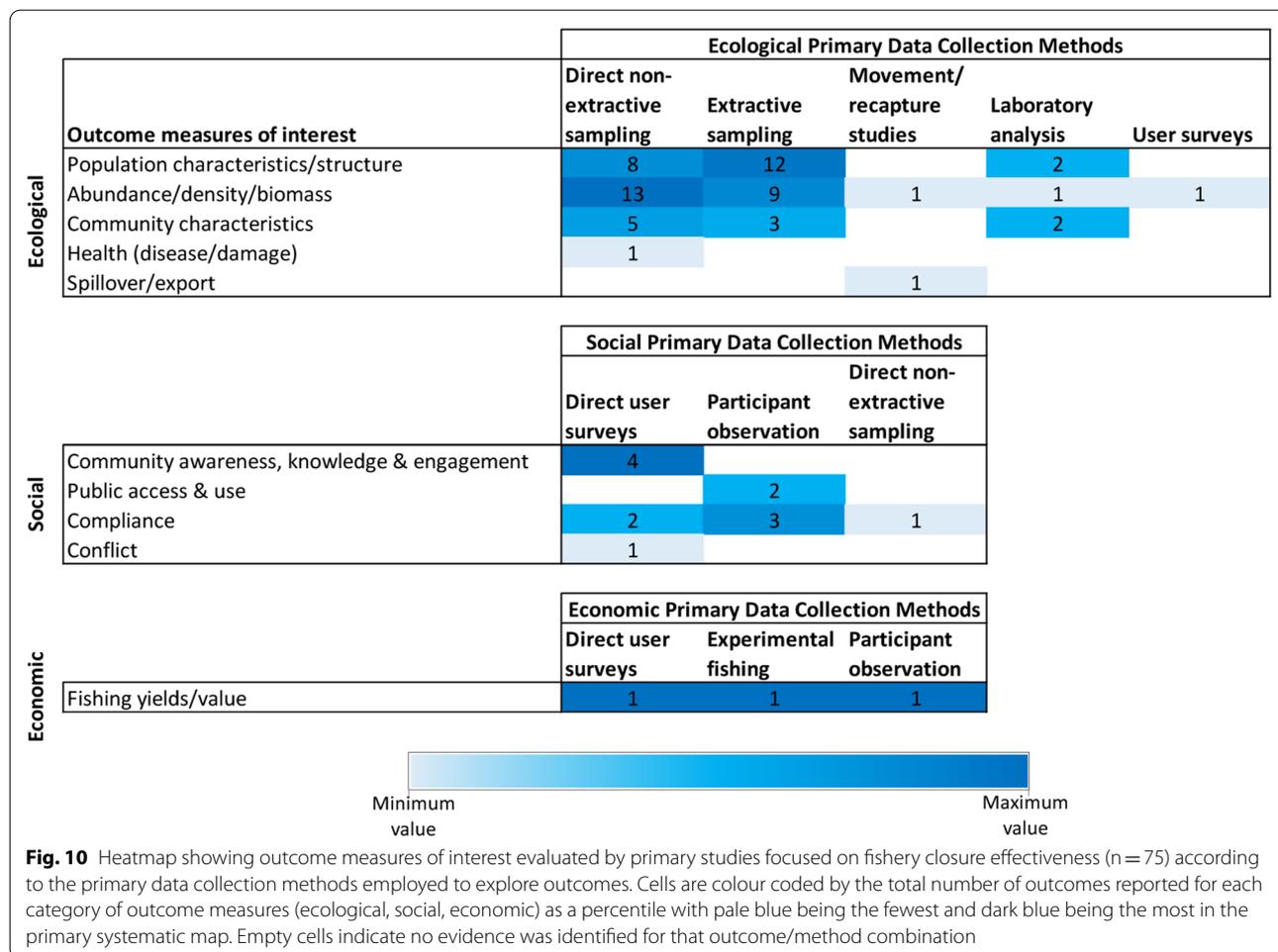


Five articles evaluated a network of fishery closures rather than multiple individual sites (n=23) or a single site (n=21). For those that considered networks of sites, articles either conducted principal evaluation using descriptive analysis (n=2) or causative evaluation using inferential analysis (n=3). These evaluations focused on fishery closures established in North America (n=5) between 3 and 10 years (n=2) ago, more than 10 years ago (n=1) or with sites of mixed ages. Almost all are conducted in inshore environments (n=4) and focus on year-round restrictions (n=3).

Marine plans The evidence base for evaluating the effectiveness of marine plans was limited with only two articles identified, both of which conducted a principal evaluation. One article took a case study approach

to compare the effectiveness of five marine plans from around the world against social and economic outcomes (community awareness, knowledge and engagement; economic impacts (beyond fisheries/tourism); tourism/recreation numbers/value; fishing yields/value) using direct user surveys to collect data. The other article evaluated ecological outcomes (community characteristics; abundance/density/biomass) using descriptive analysis and employing direct non-extractive sampling in data collection for one marine plan located in Oceania.

MPAs and fishery closures Five articles focused on MPAs and fishery closures combined. These undertook principal evaluation through descriptive (n=3) or narrative (n=1) analysis, or causative evaluation using



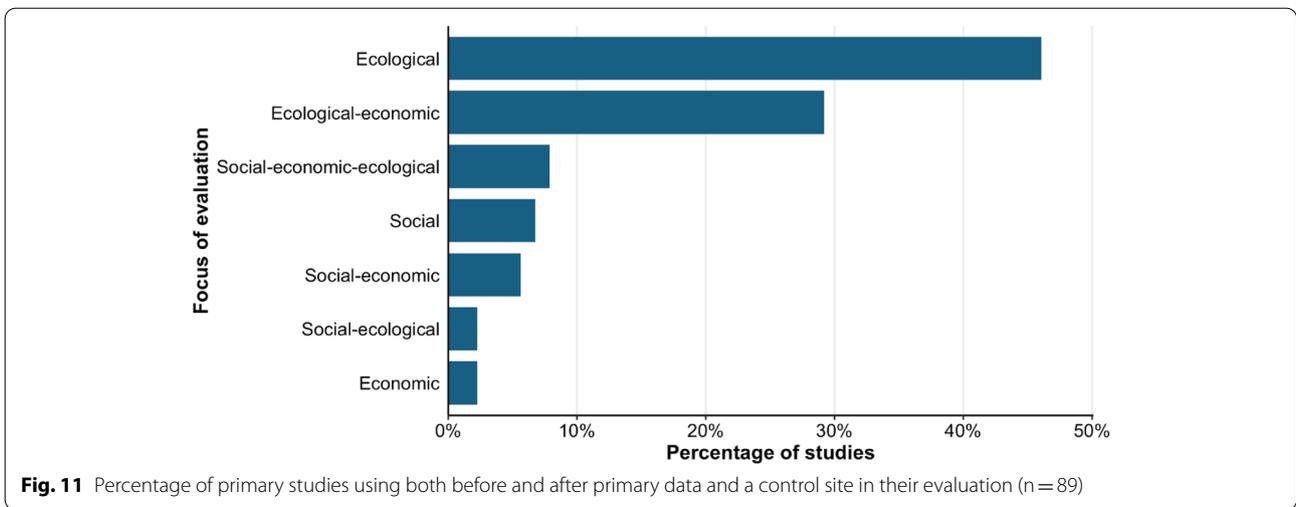
inferential analysis (n=1). Most principal evaluations focused on ecological outcomes only (n=2) with one article focused solely on social outcomes and one on social, economic and ecological outcomes. The article that undertook a causative evaluation only focused on ecological outcomes. A total of ten different outcome measures of interest were reported from the five articles with the ecological outcomes of abundance/density/biomass (n=4) and community characteristics (n=3) most frequently considered. Data were collected for ecological outcomes through a variety of methods: user surveys, extractive sampling and direct non-extractive sampling. Data for social and economic outcomes were collected through indirect and direct user surveys. Sites evaluated were located across North America (n=2), Europe (n=1) and Oceania (n=1) with one article evaluating a range of sites from around the world.

Only one article undertook a principal evaluation for a network of sites in Oceania with the remainder focusing on multiple individual sites. To evaluate the network, the articles used descriptive analysis focused on the

ecological outcomes of community characteristics and abundance/density/biomass and collected primary data through direct non-extractive sampling.

Study design

In the primary systematic map, most studies evaluated the effectiveness of spatial management measures against a control site (site outside of spatial management) or another area under spatial management using data collected after designation or regulations were put in place (61.8%, n=530/858). 144 studies (16.8%) did not use a control site or a spatial reference site, instead evaluating against a temporal timeframe: 81 of these used only data after the management designation or regulations were established, 60 used data before and after, and three were unclear in their evaluation timeframe. Half (50.0%, n=80/160) of all studies evaluating social outcomes did not use a control site or another area under spatial management in their evaluation, compared with 38.0% (n=30/79) of economic studies and 9.2% (n=57/619) of ecological studies. Only 89 studies (10.4%) used primary

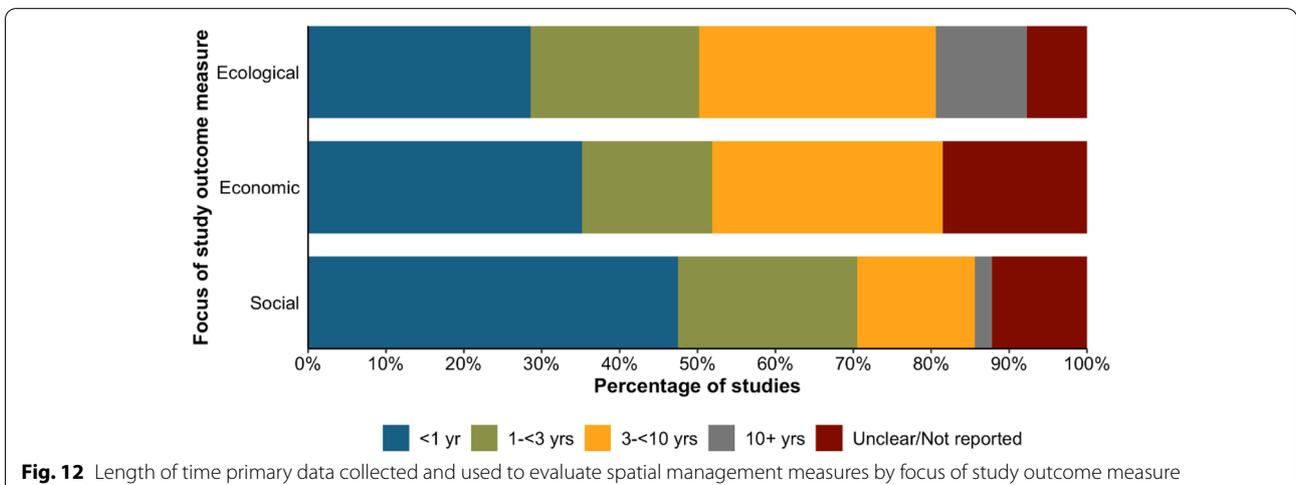


data from a control site and before and after designation/regulation to evaluate the effectiveness of a spatial management measure: of these most had an ecological (46.1%, n=41) or ecological-economic (29.2%, n=26) focus of evaluation (Fig. 11). The depth of evaluation undertaken made little difference to the use of both before-after data and a control site with 8.5% (n=51/599) of principal, 15.2% (n=38/250) of causative and no benefit evaluations using this approach. No trend over time was identified in the combined use of before-after data with a control site.

For those studies that reported using primary data to evaluate outcomes against (n=734), study duration (i.e. years for which primary data were collected and outcome measures were evaluated) was relatively evenly distributed across duration categories up to 10 years: 32.2% of studies collected data over a period of less than 1 year; 21.7%, one to less than 3 years; and

27.7% across three to less than 10 years. Only 9.1% of studies reported collecting data for ten or more years, and 9.4% of studies either did not report the length of time they collected data for or were unclear. For those studies where primary data were collected, the majority of social (46.3%, n=63/136) and economic (34.0%, n=18/53) outcomes were evaluated using primary data collected in less than 1 year. Social outcomes were more commonly evaluated using primary data collected in fewer than 3 years (69.9%, n=95/136) compared with economic outcomes (50.9%, n=27/53) and ecological outcomes (50.1%, n=273/545) (Fig. 12). Almost all (95.5%, n=64/68) studies that collected primary data for ten or more years evaluated spatial management measure(s) against ecological outcomes with the remainder considering social outcomes.

In the secondary systematic map, all studies evaluated the effectiveness of spatial management measures using



a control site or another area under spatial management ($n=63$), however unlike primary studies the majority of these (60.4%) were unclear as to the evaluation data timeframe used ($n=19$) or reported using data from different timeframes for different sites ($n=19$). Evaluations against ecological outcomes were predominately conducted using a control site (60.4%, $n=29/48$), while evaluations against social and economic outcomes generally used another site under spatial management (75.0% [$n=6/8$] and 71.4% [$n=5/7$] respectively). Only six studies (9.5%) used both before and after primary data, and a control site, against which to evaluate the effectiveness of a spatial management measure: two with an ecological and four with a social-economic-ecological focus of evaluation; four of these conducted a principal evaluation (16.0%) with the remaining two undertaking a causative evaluation (5.3%).

Limitations of the map

Limitations due to the search strategy

The search strategy employed to generate this map was designed to capture the breadth of relevant topics; however, it was not exhaustive. We recognise that a substantial volume of relevant literature likely exists in other languages, from other countries and from articles published prior to 2009. Moreover, there is a risk that our search terms were too narrow, and therefore that some studies using less common synonyms may have been missed. This risk was inevitable, as this systematic map spans ecological, economic and social disciplines, although we tried to mitigate this through the broad expertise of our Review Team and Stakeholder Group. While our search strategy attempted to capture the diversity of terminology used by these fields through piloting and testing with an interdisciplinary Review Team and Stakeholder Group, and by keeping the search string broad focusing only on Population and Intervention terms, we recognise that we may have omitted some terms in our search that may have resulted in missed literature areas. For example, MPAs are referred to around the world in a multitude of ways, not all of which will have been captured in our search string—less commonly used terms include 'sites of community importance' or 'refuge areas' [44]. Furthermore, while we undertook extensive bibliographic searching by screening the reference lists of 49 relevant tertiary reviews, we did not conduct forward and backward citation screening of included primary or secondary literature given available resources.

Limitations due to bias in pool of articles found

Meta-data coding within this map was intended to capture general characteristics of articles and the need to categorise studies means that some degree of subjectivity

is inevitable. While we attempted to reduce this through dual coding with one member of the Review Team being responsible for consistency checking across the whole database and discussions within the Review Team regarding uncertainties, some level of subjectivity is likely to have remained. Furthermore, given the broad scope and size of this systematic map, no critical appraisal of internal validity was undertaken. Instead, meta-data on elements of study design that might relate to validity were extracted to provide a basic overview of the robustness of evidence. This highlighted that many studies relied on evaluating the effectiveness of a spatial management measure according to a temporal relationship, often without a 'before' time period, or without the use of a spatial reference site. This limits the ability of evaluations to attribute changes in an outcome to the spatial management measure. Nonetheless, without conducting in-depth critical appraisal of the included studies, it is not possible to provide a clear indication of the overall reliability of the evidence base. Finally, differences in both use of terminology across different authors and regional application was encountered and meta-data was extracted as reported by each study. For example, spatial management measures can have multiple designations meaning several spatial management measures were reported as both a fishery closure and as an MPA across different primary studies, and sites that may have been considered an example of a marine plan, in another study or by another author, reported themselves as MPAs. The latter in particular will have contributed to the dominance of evaluations for MPAs and the lack of evaluations for marine plans.

Conclusions

This systematic map provides an overview of existing evidence on methodologies for monitoring and evaluating marine spatial management measures in countries of relevance to the UK. We identified a total of 391 primary articles and 33 secondary articles, describing 858 and 63 studies respectively, which revealed a number of knowledge gaps and biases in the current evidence base. In particular, we found clear concentrations of research efforts on marine protected areas over other forms of spatial management measures, and on ecological, rather than social or economic, outcomes of the spatial management measures included in this study. The implications of these gaps and biases are explored below.

Implications for policy/management

Investment in developing marine spatial management has grown significantly in recent decades and, whilst historically management has focused on ecological aspects, social and economic aspects are increasingly considered

[22]. This change reflects the growing recognition that long-term sustainability requires multidisciplinary, integrated management to balance the social and economic implications of marine management with ecological sustainability [28, 45]. Appropriate monitoring and evaluating of management actions is essential to improve understanding of what constitutes an effective action and how to achieve goals. However, as our systematic map shows, the monitoring and evaluation that has taken place within our countries of interest (Table 1) over the last decade (2009–2019) remains predominantly ecological, with substantial knowledge gaps around social and economic monitoring and evaluation of spatial management measures. Insufficient social and economic monitoring and evaluation of spatial management measures limits the capacity of policy makers and managers to assess and respond to the social and economic implications of these management tools and further exacerbates existing challenges of incorporating social and/or economic considerations into policy and management [46]. Spatial management measures are social constructs and their success often depends on social and economic factors [28, 47–49]. Failure to incorporate social and economic considerations, and indeed a reasonable range of outcomes within the social and economic categories, alongside ecological outcomes, risks undertaking incomplete evaluations that do not truly represent the implications of a spatial management measure, which could affect its long term sustainability. The paucity (2.8%, $n=11$) of evaluations identified in this systematic map that integrate ecological, social and economic aspects highlights persistent challenges around achieving, multidisciplinary integrated management of the marine environment [22]. Nevertheless, those integrated evaluations that have been identified here could provide a useful resource for policy makers looking to develop more multidisciplinary approaches.

The differences in monitoring and evaluation approaches identified for the different social, economic, and ecological outcomes also has policy implications. The vast majority of both principal (64.4%, $n=177/275$) and causative (76.5%, $n=88/115$) evaluations focus solely on ecological outcomes with comparatively few considering social or economic outcomes (Fig. 5). Study design also varied between outcome groups; in particular, the relative lack of social and economic studies that included data from before implementation of a given spatial management measure, and evaluated the site against control data (Fig. 11), limits the capacity to assign causation to social and economic outcomes. The challenges of assigning causation to spatial management measure outcomes based on existing social-economic monitoring programmes are widespread [31] and the resultant lack

of clarity inherently makes policy and management decisions around social-economic outcomes more difficult.

The outputs from this systematic map (i.e. the map database) provide a resource that will help improve understanding of current approaches to evaluation. Decision-makers responsible for implementing, managing and evaluating spatial management measures may therefore find this map useful to (1) provide an indication of the extent of the current evidence base and help in deciding how to evaluate spatial management against particular outcomes, and (2) guide future scope of evaluation according to objectives and desired outcomes of spatial management measures.

Implications for research

The findings of this systematic map show knowledge clusters around the evaluation of the effects of marine protected areas on ecological outcomes, particularly abundance/density/biomass, population characteristics/structure and community characteristics. However, this systematic map has also highlighted several absolute knowledge gaps where no evidence exists for specific outcome measures of interest that we pre-defined in the protocol across all foci of evaluation: (1) social outcomes—historic/cultural heritage assets and character of seascapes; and (2) economic outcomes—natural capital value. Key knowledge gaps include evaluation across social and economic outcomes, or combinations of these, and of the overall merit and/or worth (benefit evaluation) of spatial management measures. Other knowledge gaps relate to evaluation of: marine plans; networks of sites; real-time, temporary or seasonal closures; spatial management within offshore waters, and lagoon or estuary environments.

The lack of evaluation studies for these knowledge gaps mean that there is insufficient evidence to support informed decision-making in these areas and further research is required. Additional research into the social and economic outcomes of spatial management measures also needs to include development of data collection programmes that allow for more causative evaluations.

Finally, this systematic map suggests there is a lack of long-term (> 10 years) studies amongst those that collect primary data to evaluate outcomes against, as well as few that use primary data from before the designation/regulation was put in place and a control site. With ecological effects from spatial management expected to develop over decades [50] and social and economic effects also likely to change over time [51], there is a need for evaluation over longer timescales to fully identify the effects spatial management can have. Although

the evidence base has grown over the past two decades, this systematic map suggests that information to develop comprehensive evaluation frameworks remains insufficient. Greater understanding on how to evaluate the effectiveness of spatial management measures is required to support improved management of global ocean resources and spaces.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13750-021-00227-x>.

Additional file 1. ROSES form for systematic maps.

Additional file 2. List of peer-reviewed and grey literature sources searched.

Additional file 3. List of articles included, unobtainable and excluded at full text screening with reasons for exclusion.

Additional file 4. Full coded meta-data for primary and secondary systematic maps.

Acknowledgements

We would like to extend our thanks to our Stakeholder Group for their participation in our two workshops and development of the original protocol. We would also like to thank Kieran Curran for help with website searches and pdf retrieval. We would also like to thank our three reviewers for their constructive comments that helped us greatly improve our manuscript.

Authors' contributions

The review process was coordinated by BO'L and KLY. BO'L, JPC, NM, SLD and KLY conducted screening. BO'L, JPC, NM, SLD and KLY refined the data coding tool. BO'L, JPC, NM and KLY conducted data extraction. BO'L and JPC analysed the data. BO'L drafted the first version of the manuscript. All authors edited and commented on the manuscript. All authors read and approved the final manuscript.

Funding

This work is funded by the UK Natural Environment Research Council (NERC, Innovation grant reference: NE/S016015/1). SLD was supported by funding from the Marshall Scholarship. DR's time was supported by funding from the European Maritime and Fisheries Fund (EMFF) through grant ENG 2162 (MAR-MATE: Marine management tool for evaluating conservation evidence). KLY is supported by a NERC Knowledge Exchange Fellowship (NE/P00668X/1).

Availability of data and materials

All data generated or analysed during this study are included in this published article [and its additional files].

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹School of Science, Engineering and Environment, University of Salford, Manchester M5 4WX, UK. ²Department of Environment and Geography, University of York, York YO10 5NG, UK. ³Centre for Ecology & Conservation, College of Life and Environmental Sciences, University of Exeter, Penryn Campus, Penryn, UK. ⁴CBASS, Brunel University London, Kingston Lane, Uxbridge UB8

3PH, UK. ⁵School of Earth and Ocean Sciences, Cardiff University, Park Place, Cardiff CF10 3AT, Wales, UK. ⁶Interdisciplinary Centre for Conservation Science, Department of Zoology, University of Oxford, Oxford OX1 3SZ, UK. ⁷New Economics Foundation, 10 Salamanca Place, London SE1 7HB, UK. ⁸Centre for Environment, Fisheries and Aquaculture Science (CEFAS), Lowestoft, UK. ⁹School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

Received: 14 September 2020 Accepted: 29 May 2021

Published online: 11 June 2021

References

- Costanza R. The ecological, economic, and social importance of the oceans. *Ecol Econ*. 1999;31:199–213.
- Costanza R, de Groot R, Sutton P, der Ploeg S, Anderson SJ, Kubiszewski I, et al. Changes in the global value of ecosystem services. *Glob Environ Chang*. 2014;26:152–8.
- Gamfeldt L, Lefcheck JS, Byrnes JEK, Cardinale BJ, Duffy JE, Griffin JN. Marine biodiversity and ecosystem functioning: what's known and what's next? *Oikos*. 2015;124:252–65.
- Worm B, Barbier EB, Beaumont N, Duffy JE, Folke C, Halpern BS, et al. Impacts of biodiversity loss on ocean ecosystem services. *Science*. 2006;314:787–90.
- United Nations. Sustainable Development Goal 14: conserve and sustainable use the oceans, seas, and marine resources for sustainable development. 2015; <https://sustainabledevelopment.un.org/sdg14>. Accessed 7 Nov 2018.
- Convention on Biological Diversity. COP Decision X/2. Strategic plan for biodiversity 2011–2020. 2010; <https://www.cbd.int/decision/cop/?id=12268>. Accessed 25 Sep 2018.
- OSPAR Convention for the Protection of the Marine Environment of the North-East Atlantic, OSPAR Recommendation 2003/3; Record of the Joint Ministerial Meeting of Helsinki & OSPAR Commissions 2003. 2003: Bremen.
- Chaniotis P, Cioffi B, Farmer R, Cornthawite A, Flavell B, Carr H. Developing an ecologically-coherent and well-managed Marine Protected Area network in the United Kingdom: 10 years of reflection from the Joint Nature Conservation Committee. *Biodiversity*. 2018;19:140–7.
- Ehler CN. Marine spatial planning: an idea whose time has come. In: Yates KL, Bradshaw CJA, editors. *Offshore energy and marine spatial planning*. London: Routledge; 2018. p. 6–17.
- Lubchenco J, Grorud-Colvert K. Making waves: the science and politics of ocean protection. *Science*. 2015;350:382–3.
- Yates KL, Schoeman DS, Klein CJ. Ocean zoning for conservation, fisheries and marine renewable energy: assessing trade-offs and co-location opportunities. *J Environ Manage*. 2015;152:201–9.
- JNCC. UK Marine Protected Area network statistics. 2020; <https://jncc.gov.uk/our-work/uk-marine-protected-area-network-statistics/>. Accessed 25 Feb 2021.
- Houses of Parliament. Marine planning POSTNOTE Number 388. 2011; https://www.parliament.uk/documents/post/postpn388_Marine-Planning.pdf. Accessed 31 Jan 2019.
- Clarke J, Bailey DM, Wright PJ. Evaluating the effectiveness of a seasonal spawning area closure. *ICES J Mar Sci*. 2015;72:627–2637.
- Russi D, Pantzar M, Kettunen M, Gitti G, Mutafoglu K, Kotulak M, et al. *Socio-Economic Benefits of the EU Marine Protected Areas*. 2016: Report prepared by the Institute for European Environmental Policy (IEEP) for DG Environment.
- Giakoumi S, Scianna C, Plass-Johnson J, Micheli F, Grorud-Colvert K, Thiriet P, et al. Ecological effects of full and partial protection in the crowded Mediterranean Sea: a regional meta-analysis. *Sci Rep*. 2017;7:8940.
- Gill DA, Mascia MB, Ahmadi GN, Glew L, Lester SE, Barnes M, et al. Capacity shortfalls hinder the performance of marine protected areas globally. *Nature*. 2017;543:665–9.
- Zupan M, Fragkopoulou E, Claudet J, Erzini K, Horta e Costa B, Gonçalves EJ. Marine partially protected areas: drivers of ecological effectiveness. *Front Ecol Environ*. 2018;16:381–7.

19. McIntosh EJ, Chapman S, Kearney SG, Williams B, Althor G, Thorn JPR, et al. Absence of evidence for the conservation outcomes of systematic conservation planning around the globe: a systematic map. *Environ Evid*. 2018;7:22.
20. Sciberras M, Jenkins SR, Mant R, Kaiser MJ, Hawkins SJ, Pullin AS. Evaluating the relative conservation value of fully and partially protected marine areas. *Fish Fish*. 2015;16:58–77.
21. Day J. The need and practice of monitoring, evaluating and adapting marine planning and management—lessons from the Great Barrier Reef. *Mar Policy*. 2008;32:823–31.
22. Addison PFE, Collins DJ, Trebilco R, Howe S, Bax N, GHedge P, et al. A new wave of marine evidence-based management: emerging challenges and solutions to transform monitoring, evaluating, and reporting. *ICES J Marine Sci*. 2017;75:941–52.
23. Marine Management Organisation, Evidence requirements for the Marine Management Organisation (MMO). 2018.
24. Department for Environment Food and Rural Affairs, Marine Programme Evidence Plan Policy portfolio: Marine and Fisheries Policy area within portfolio: Marine environment, marine biodiversity, marine and freshwater fisheries. PB13919 2013.
25. Rees SE, Sheehan EV, Stewart BD, Clark R, Appleby T, Attrill MJ, et al. Emerging themes to support ambitious UK marine biodiversity conservation. *Marine Policy*. 2020;117:103864.
26. Ribas LGS, Pressey RL, Loyola R, Bini LM. A global comparative analysis of impact evaluation methods in estimating the effectiveness of protected areas. *Biol Conserv*. 2020;246:108595.
27. Rodrigues ASL, Cazalis V. The multifaceted challenge of evaluating protected area effectiveness. *Nat Commun*. 2020;11:5147.
28. Yates KL, Clarke B, Thurstan RH. Purpose vs performance: what does marine protected area success look like? *Environ Sci Policy*. 2019;92:76–86.
29. Yates KL, Payo Payo A, Schoeman DS. International, regional and national commitments meet local implementation: a case study of marine conservation in Northern Ireland. *Mar Policy*. 2013;38:140–50.
30. Pendleton LH, Ahmadi GN, Browman HI, Thurstan RH, Kaplan DM, Bartolino V. Debating the effectiveness of marine protected areas. *ICES J Mar Sci*. 2018;75:1156–9.
31. Stelzenmuller V, Cormier R, Gee K, Shucksmith R, Gubbins M, Yates KL, et al. Evaluation of marine spatial planning requires fit for purpose monitoring strategies. *J Environ Manage*. 2021;278:111545.
32. O'Leary BC, Stewart BD, McKinley E, Addison PFE, Williams C, Carpenter G, et al. What is the nature and extent of evidence on methodologies for monitoring and evaluating marine spatial management measures in UK and similar coastal waters? A systematic map protocol. *Environ Evid*. 2019;8:1–9.
33. Pullin AS, Bangpan M, Dalrymple S, Dickson K, Haddaway NR, Healey JR, et al. Human well-being impacts of terrestrial protected areas. *Environ Evid*. 2013;2:19.
34. Roberts KE, Valkan RS, Cook CN. Measuring progress in marine protection: a new set of metrics to evaluate the strength of marine protected area networks. *Biol Cons*. 2018;219:20–7.
35. Collaboration for Environmental Evidence, Guidelines and Standards for Evidence Synthesis in Environmental Management. Version 5.0, A.S. Pullin, et al., Editors. 2018.
36. IUCN WCPA, Applying IUCN's Global Conservation Standards to Marine Protected Areas (MPA). Delivering effective conservation action through MPAs, to secure ocean health & sustainable development. 2018: Version 1.0. Gland, Switzerland. 4pp.
37. Elvidge CD, Ghosh T, Baugh K, Zhizhin M, Hsu F-C, Katada NS, et al. Rating the effectiveness of fishery closures with visible infrared imaging radiometer suite boat detection data. *Front Mar Sci*. 2018;5:132.
38. Foley MM, Halpern BS, Micheli F, Armsby MH, Caldwell MR, Crain CM, et al. Guiding ecological principles for marine spatial planning. *Mar Policy*. 2010;34:955–66.
39. Haddaway NR, Macura B, Whaley P, Pullin AS. 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. *Environ Evid*. 2018;7:7.
40. UK Government. Marine and Coastal Access Act 2009. 2009; <https://www.legislation.gov.uk/ukpga/2009/23/contents>. Accessed 31 Jan 2019.
41. Howarth LM, Roberts CM, Hawkins JP, Steadman DJ, Beukers-Stewart BD. Effects of ecosystem protection on scallop populations within a community-led temperate marine reserve. *Mar Biol*. 2015;162:823–40.
42. Great Barrier Reef Marine Park Authority, Great Barrier Reef Outlook Report 2019. 2019, GBRMP, Townsville.
43. Dixon-Woods M. Using framework-based synthesis for conducting reviews of qualitative studies. *BMC Med*. 2011;9:39.
44. Mangano MC, O'Leary BC, Mirto S, Mazzola A, Sará G. The comparative biological effects of spatial management measures in protecting marine biodiversity: a systematic review protocol. *Environ Evid*. 2015;4:21.
45. Thébaud O, Link JS, Kohler B, Kraan M, López R, Poos JJ, et al. Managing marine socio-ecological systems: picturing the future. *ICES J Mar Sci*. 2017;74:1965–80.
46. McKinley E, Acott T, Yates KL. Marine social sciences: Looking towards a sustainable future. *Environ Sci Policy*. 2020;108:85–92.
47. Jupiter SD, Epstein G, Ban NC, Mangubhai S, Fox M, Cox M. A social-ecological systems approach to assessing conservation and fisheries outcomes in Fijian locally managed marine areas. *Soc Nat Resour*. 2017;30:1096–111.
48. Twichell J, Pollnac R, Christie P. Lessons from Philippines MPA management: social ecological interactions, participation, and MPA performance. *Environ Manage*. 2018;61:916–27.
49. Yates KL. Meaningful stakeholder participation in marine spatial planning with offshore energy. In: Yates KL, Bradshaw CJA, editors. *Offshore energy and marine spatial planning*. London: Routledge; 2018. p. 169–88.
50. Edgar GJ, Stuart-Smith RD, Willis TJ, Kininmonth S, Baker SC, Banks S, et al. Global conservation outcomes depend on marine protected areas with five key features. *Nature*. 2014;506:216–20.
51. Ban NC, Gurney GG, Marshall NA, Whitney CK, Mills M, Gelcich S, et al. Well-being outcomes of marine protected areas. *Nat Sustain*. 2019;2:524–32.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

