
Article:
Woodcock, Paul, O'Leary, Bethan Christine orcid.org/0000-0001-6595-6634, Kaiser, Michel et al. (1 more author) (2017) Your evidence or mine?:Systematic evaluation of reviews of marine protected area effectiveness. Fish and fisheries. pp. 668-681. ISSN: 1467-2960

https://doi.org/10.1111/faf.12196

Your evidence or mine? Systematic evaluation of reviews of marine protected area effectiveness

Journal:	*Fish and Fisheries*
Manuscript ID	FaF-16-May-OA-095.R1
Wiley - Manuscript type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Woodcock, Paul; Bangor University, Centre for Evidence-Based Conservation, School of the Environment, Natural Resources and Geography O'Leary, Bethan C; Bangor University, Centre for Evidence-Based Conservation, School of the Environment, Natural Resources and Geography; Kaiser, Michel; Bangor Universtiy, School of Ocean Sciences Pullin, Andrew; Bangor University, Centre for Evidence-Based Conservation, School of the Environment, Natural Resources and Geography
Key terms:	biodiversity conservation, CEESAT, evidence review, evidence synthesis, evidence-base, review evaluation
Abstract:	Marine Protected Areas (MPAs) are a key strategy for mitigating the impacts of fisheries, but their designation can be controversial, and there is uncertainty surrounding when and where MPAs are most effective. Evidence synthesis that collates primary research on MPA effectiveness can provide a crucial bridge between research, policy, and practice. However, reviews vary in scope and rigour, meaning decision-makers face the challenge of identifying appropriate reviews. Documenting differences amongst reviews can therefore support non-specialists in locating the most relevant and rigorous reviews, and can also assist researchers in targeting evidence gaps. We addressed these priorities by systematically searching for reviews examining effectiveness of MPAs for biodiversity, critically appraising methods used, and categorising review scope. The 27 reviews assessed overlapped in scope (suggesting some redundancy) and differed substantially in reliability. Key strengths related to the effects of MPAs on fish abundance and the influence of MPA size and age on effectiveness. However, several gaps were noted, with some questions not addressed and others lacking highly reliable syntheses – importantly, the latter may create the perception that particular questions have been adequately addressed, potentially deterring new syntheses. Our findings indicate key aspects of review conduct that could be improved (e.g. documenting critical appraisal of primary research, evaluating potential publication bias), and can facilitate evidence-based policy by guiding non-specialists to the most reliable and relevant reviews. Lastly, we suggest that future reviews with broader taxonomic coverage and considering the influence of a wider

	range of MPA characteristics on effectiveness would be beneficial.

Title (option 1): Your evidence or mine? Systematic evaluation of reviews of marine protected area effectiveness

Title (option 2): Your evidence or mine? Systematic evaluation of the scope and reliability of reviews of marine protected area effectiveness

Running Title: Evaluating marine protected area reviews

Paul Woodcock[1], Bethan C. O'Leary[1], Michel J. Kaiser[2], Andrew S. Pullin[1]*,

[1] Centre for Evidence-Based Conservation, School of Environment, Natural Resources and Geography, Bangor University, Gwynedd LL57 2UW, UK

[2] School of Ocean Sciences, Bangor University, Menai Bridge, Anglesey, UK, LL59 5AB

* Corresponding author email: a.s.pullin@bangor.ac.uk, fax: 01248 354997, tel: 01248 382444

Abstract

Marine Protected Areas (MPAs) are a key strategy for mitigating the impacts of fisheries, but their designation can be controversial, and there is uncertainty surrounding when and where MPAs are most effective. Evidence synthesis that collates primary research on MPA effectiveness can provide a crucial bridge between research, policy, and practice. However, reviews vary in scope and rigour, meaning decision-makers face the challenge of identifying appropriate reviews. Documenting differences amongst reviews can therefore support non-specialists in locating the most relevant and rigorous reviews, and can also assist researchers in targeting evidence gaps. We addressed these priorities by systematically searching for reviews examining effectiveness of MPAs for biodiversity, critically appraising methods used, and categorising review scope. The 27 reviews assessed overlapped in scope (suggesting some redundancy) and differed substantially in reliability. Key strengths related to the effects of MPAs on fish abundance and the influence of MPA size and age on effectiveness. However, several gaps were noted, with some questions not addressed and others lacking highly reliable syntheses – importantly, the latter may create the perception that particular questions have been adequately addressed, potentially deterring new syntheses. Our findings indicate key aspects of review conduct that could be improved (e.g. documenting critical appraisal of primary research, evaluating potential publication bias), and can facilitate evidence-based policy by guiding non-specialists to the most reliable and relevant reviews. Lastly, we suggest that future reviews with broader taxonomic coverage and considering the influence of a wider range of MPA characteristics on effectiveness would be beneficial.

Keywords: biodiversity conservation, CEESAT, evidence review, evidence synthesis, evidence-base, review evaluation.

Table of contents:

Introduction

Fisheries exert one the most widespread anthropogenic impacts on marine ecosystems, and can threaten the populations and processes that underpin vital ecosystem services (Butchart et al., 2010, Ramirez-Llodra et al., 2011, Halpern et al., 2012). Establishing marine protected areas (MPAs), in which fishing is restricted to varying degrees, is one of the principal tools for mitigating these impacts (Gaines et al., 2010, Halpern et al., 2010, OSPAR, 2010, Lascelles et al., 2012). Accordingly, the extent of the marine environment with some level of protection from fisheries (and other human activities) has increased steadily from around 0.9% in 2000 to an estimated 3.5% in 2015 (Thomas et al. 2014; Lubchenco & Grorud-Colvert 2015), and is set to rise further in line with the Convention on Biological Diversity target of 10% coverage by 2020 (CBD, 2010).

There are a range of options for expanding the MPA network in terms of design, placement, and management. Different taxa may also benefit from protection to varying degrees, and across different timescales (Fox et al., 2012, Hays and Scott, 2013). Given the importance that strategies to mitigate fisheries impacts place on MPAs, the increasing promotion of MPAs as a fisheries management tool, and the potential socio-economic and political challenges associated with establishing new reserves, it is essential that scientific evidence is used to identify and communicate the factors that influence effectiveness – thereby allowing new MPAs to be optimally designed and the predicted benefits to be understood. While primary research forms the basis of this evidence, increasing publication rates (Pautasso, 2012, Larsen and von Ins, 2010, Li and Zhao, 2015) and the variable quality of primary studies (Willis et al., 2003, Caveen et al., 2012) creates problems for decision-makers in: (1) keeping up-to-date with emerging research; (2) evaluating the appropriateness of methods, data analysis and interpretation in each study; and (3) obtaining an accurate representation of the overall evidence base on MPA effectiveness.

Evidence syntheses can assist decision-makers by summarising primary literature on MPAs, with reviews providing a crucial bridge linking primary research with policy and practice. The number of reviews examining MPAs is increasing rapidly (Caveen et al., 2012). However, reviews that do not follow rigorous methods to maximise objectivity and comprehensiveness in searching for, appraising, and synthesising primary research may unintentionally misinform or misrepresent the evidence base. For example,

Huntington (2011) argued that the majority of meta-analyses that examined the effectiveness of MPAs did not address possible publication bias (the tendency to publish positive or hypothesis-affirming results rather than null or controversial findings; Møller and Jennions, 2001) and so may have provided an incomplete picture of the available primary research. Decision-makers and other non-specialists may lack the resources or expertise to systematically collate and appraise all reviews prior to use, and are therefore faced with a similar challenge as for primary literature: identifying the most relevant and rigorous reviews and appreciating the strengths and limitations of the reviews used. Furthermore, where reviews overlap in scope, apparently conflicting interpretations of evidence can reflect variation in review reliability or subtle differences in emphasis amongst reviews. This leads to a perception amongst policymakers that the science is inconclusive, resulting in no decisions being made, unnecessary delays, or selective use of evidence. The existence of a review on a particular topic could also give the impression that the topic has already been investigated and so does not require further exploration, even if the review is potentially less reliable. Future high quality syntheses might thus be deterred, resulting in what could be termed 'cryptic' evidence gaps.

To address the above issues, we evaluated the scope and the methods used by reviews that examine the effectiveness of MPAs as a tool for mitigating the impacts of fisheries on biodiversity. We carried out a systematic search for relevant reviews and categorised the scope of each review according to: (i) the geographic region(s) explored (global, temperate, tropical, polar), (ii) the taxa considered (fish, invertebrates, algae, birds, mammals, reptiles), (iii) the characteristics of MPAs investigated (size, age, level of protection, size of buffer zone, connectivity), and (iv) the measures used to evaluate MPA effectiveness (abundance, biomass, species richness, size distribution of individuals within or amongst species). We then assessed the reliability (objectivity, transparency and comprehensiveness) of each review using a standardised, published protocol (Woodcock et al., 2014), and identified general strengths and weaknesses in the review literature. Finally, we combined the categorisation of review scope with the assessment of review rigour to describe the review landscape on MPA effectiveness.

The principal objectives of our study are therefore to:

1) Assist decision-makers in quickly identifying the most relevant and rigorous reviews on topics of interest, and any limitations in the evidence used.

2) Assist decision-makers and researchers in targeting gaps in the review literature and avoiding duplication of previous reviews.

3) Identify strengths and weaknesses in the methods used by reviews, to assist researchers in maintaining and improving the rigour of future evidence syntheses.

We focused on reviews that synthesised empirical research on MPA effectiveness. Empirical data represent a large and growing volume of evidence, and reviews of this research have clear potential to support decision-making if results are provided on the outcomes of implementing MPAs, or on the characteristics that influence MPA effectiveness (e.g. Lester et al. 2009; Sciberras et al. 2013). We stress however, that the findings from such reviews should be considered in conjunction with insights from the extensive body of theoretical work on MPA effectiveness (e.g. Gaines et al. 2003; White et al. 2011), as well as site-specific considerations relating to stakeholder priorities and the objectives of individual MPAs.

Materials and methods

Review searching and screening

We compiled a database of review articles that examined MPA effectiveness through searches of peer-reviewed and grey literature using multiple databases (Web of Science, Scopus, Aquatic Sciences and Fisheries Abstracts, ScienceDirect, Centre for Agriculture and Bioscience International, Directory of Open Access Journals, and Index to Theses online), www.googlescholar.com, and websites of a range of organisations (Table A1). We used search terms adapted from a recent systematic review that evaluated the effectiveness of fully and partially protected MPAs (Sciberras et al., 2013, Sciberras et al., 2015). Search strings were modified according to the database used, but included the terms 'marine protected area', 'marine reserve', 'marine sanctuary', 'no-take area', 'partially protected area', 'fishery reserve', 'marine area closure', 'gear restriction zone' and 'buffer zone' to identify research related to MPAs. To narrow the focus to review articles we combined these terms with 'review, 'meta-analysis' and 'synthesis' For example, the search string used for locating studies in Web of Science, AFSA, CABI, ScienceDirect and Scopus was:

("marine reserve"* OR *"marine sanctuary"* OR *(marine* AND *"no-take zone")* OR *(marine* AND *harvest refug*)* OR *(marine* AND *"buffer zone")* OR *(marine* AND *partial** AND *protect*)* OR *(marine* AND

138 *"closed area")* OR *(marine* AND *"area closure")* OR *(fisher** AND *(reserve* OR *closure))* OR *("fishing*
139 *gear restriction*")* OR *("recreational fishing"* AND *protection)* OR *"marine protected area*")*

140 AND

141 (*review* OR *"meta-analy*"* OR *synthes**)

142 We only considered reviews published in the year 2000 or later to restrict our assessment to recent
143 literature. Searches took place from 14-21 May 2014 and so our study encompasses the period from 2000
144 up to this point. Full details of the search strategy including search strings with Boolean operators and
145 search dates are given in Table A1.

146 All studies found by the search were assessed for relevance and retained if the following inclusion criteria
147 were met: Type of Article: Relevant reviews should be focused on synthesising primary research that
148 collects field data to compare MPAs (fully or partially protected) with unprotected areas. This excludes
149 articles clearly marked as opinions, perspectives, technical reports/management documents that are not
150 explicitly presented as syntheses, modelling studies in which parameters are estimated through literature
151 review, and studies that analyse long-term survey data (including articles that apply meta-analytical
152 techniques – e.g. Ojeda-Martinez et al., 2007). Reviews primarily focused on synthesising the results from
153 models, or on methodological aspects of MPA monitoring and evaluation were also excluded, as were
154 reviews that focused on the ecological principles of MPA design (rather than synthesising empirical
155 research on MPA effectiveness). Whilst each of these pieces of evidence are potentially valuable, it would
156 not be appropriate to evaluate such studies using a tool designed for assessing reviews of primary research.
157 For example, analyses of long-term survey data would not necessarily be expected to follow all of the
158 methods required to produce a rigorous review of primary research (e.g. comprehensive and transparent
159 search for relevant literature, assessment of publication bias etc.). Population: Reviews can consider any
160 taxa, in any region. Intervention: Reviews must primarily examine the effects of fully and/or partially
161 protected MPAs. Outcome: Reviews must clearly examine the effectiveness of MPAs with respect to at
162 least one of: abundance, species richness, biomass, organism size. Because our emphasis was primarily on
163 the direct implications of MPAs for biodiversity conservation and mitigating the impacts of fisheries,
164 reviews that focused principally on ecosystem properties (e.g. nutrient cycling) or ecological processes

(e.g. competition, trophic interactions) were not considered. Questions relating only to socio-economic effects also fall outside the scope of our study.

We screened all articles returned by the search for relevance, first based on the title with retained articles then assessed based on the abstract. Decisions on article inclusion can be subjective and so 10% of articles screened at the abstract stage were also independently evaluated for relevance by a second person. Following conventional practice for systematic reviews (CEE 2013), kappa values were used to evaluate agreement on article relevance (Cohen 1960, Landis & Koch 1977). Kappa values account for the agreement expected by chance, and are calculated as:

κ = (observed agreement-expected agreement) / (1-expected agreement)

'Observed agreement' is the proportion of decisions in which there is agreement (i.e. both assessors regard an article as relevant, or both assessors regard an article as non-relevant). 'Expected agreement' is calculated as: [(proportion of articles accepted as relevant by A1 * proportion of articles accepted by A2) + (proportion of articles rejected by A1 * proportion of articles rejected by A2)], where A1 and A2 are the two assessors. Kappa scores of 0.6-0.8 tend to be regarded as indicating good agreement: we obtained a kappa score of 0.75, indicating that decisions over article relevance were sufficiently repeatable (CEE, 2013). Where there was disagreement on relevance during Abstract screening, articles were retained. All articles retained after the abstract screening stage were then read in full and assessed for relevance. Articles in which the relevance was uncertain at a particular stage were retained for the subsequent stage. Lastly, bibliographies of all reviews retained after full-text screening were searched for additional references – this approach increases the comprehensiveness of our search by capturing relevant reviews that may have omitted our search terms from the Abstract. Any potentially relevant studies located in this way were screened using the same title>abstract>full-text process.

Assessing review scope

We compiled 153 questions related to the effectiveness of MPA for biodiversity conservation and mitigating the impacts of fisheries. The parameters of these questions are provided in Table 1 and consider region (e.g. global, tropical etc.), taxa (fish, invertebrates etc.), MPA characteristics (e.g. size, age etc.) and outcome measures (e.g. effects on abundance, biomass etc.). Questions therefore take the broad form:

192 'What are the effects of MPAs on [*fish*]?', 'How does MPA [*size*] influence effectiveness?' etc. At this
193 level, there are 19 distinct questions, representing each element in Table 1. We then considered each
194 possible two-way combination of Taxa, Region, MPA Characteristic, and Outcome Measure to assess
195 specific questions, e.g. 'What are the effects of MPA [*size*] on [*fish*]?', What are the effects of [*tropical*]
196 MPAs on [*species richness*]?' At this level, there are 134 distinct questions. Note that constructing
197 questions by systematically combining terms in this way results in some questions that are likely to be
198 more relevant than others. However, our intention is for the evaluation of review scope and rigour to be
199 valuable to policymakers and researchers with a diverse range of priorities. In the case of the MPA
200 literature, much of the research focus has been on harvested species, whereas policy questions are
201 increasingly broad, addressing a wide range of taxa (e.g. EU Birds Directive, EU Habitats Directive). For
202 this type of exercise, we therefore view a systematic approach as preferable to identifying questions in a
203 more *ad hoc* manner based on perceived importance.

204 Reviews were categorised according to the question(s) addressed and the type of synthesis undertaken
205 (narrative synthesis [reviews that use prose to summarise and draw conclusions from primary research] or
206 meta-analysis). For the purposes of this study, we did not focus on any specific element of MPA
207 connectivity – reviews examining how any aspect of MPA connectivity influences effectiveness were
208 therefore considered to address this question. Defining the questions addressed by narrative syntheses
209 proved challenging because such syntheses are often broad-ranging with no clear boundaries to objectively
210 decide whether or not a particular question has been addressed in sufficient detail. As such, narrative
211 syntheses were categorised according to the broad focus (biodiversity conservation or fisheries) and the
212 region, type of protection (highly protection MPA [no-take] or all forms of protection [MPA]) and MPA
213 characteristic(s) explored. Each meta-analysis was categorised according to all 153 questions outlined in
214 the preceding paragraph. We categorised a meta-analytical review as addressing a particular question if
215 effect sizes were quoted directly (e.g. response ratio comparing fish density inside vs outside MPA),
216 presented graphically or used in statistical tests of relationships (e.g. relationship between effect size and
217 MPA size).

218 In calculating effect sizes for one property (e.g. the influence of MPA size), meta-analyses could include
219 other properties as potential confounding variables (e.g. MPA age), without directly calculating effect sizes

for these confounding variables. From a policy perspective, it would therefore not be possible to use such an analysis to fully understand the relationship between MPA age and effectiveness. Reviews that included relevant terms (from Table 1) as potentially confounding variables without directly reporting effect sizes for these terms were therefore noted (Supplementary Information) but not considered to directly address questions relating to the confounding variables. Finally, there may be instances in which meta-analyses are based on a small number of primary studies and so the generality of findings would be less certain. To be included as addressing a particular question, we set an arbitrary minimum threshold that meta-analyses should contain at least 10 primary research studies addressing that question. Where meta-analyses addressed a particular question but contained less than 10 studies, we noted this intended focus, as well as instances in which reviews indicated an intention to investigate a question but expressly stated that insufficient studies were available. If recent, such reviews might suggest the need for further primary research rather than additional reviews.

Critical appraisal of review reliability

We used a standardised protocol designed to assess the reliability of environmental evidence reviews (the Collaboration for Environmental Evidence Synthesis Assessment Tool [CEESAT], Woodcock et al., 2014) to critically appraise the methods of each relevant review. CEESAT assesses reviews based on 13 criteria (Table 2) for which a review can receive 3 points, 1 point or 0 points (maximum 39). The higher the score, the greater the confidence that the review is robust. Whilst CEESAT does have important limitations (e.g. does not account for methodological or interpretation errors or fraud, or include a detailed evaluation of the appropriateness of any statistical techniques used) it considers each key step of the review process and so provides a good overall picture of the likelihood that the review uses transparent methods to produce an objective, rigorous, and comprehensive synthesis of all available primary research.

All reviews were independently appraised by two assessors using CEESAT. Disagreements in scoring were then discussed and scores were amended if relevant information was overlooked by an assessor. When disagreements reflected uncertainty between assessors over whether or not a criterion was met, the midpoint score of the two assessors was used. We evaluated agreement in scoring by using a Spearman's rank test to compare the overall scores for each review between assessors, and by examining repeatability in scoring for individual criteria using (i) % agreement and (ii) kappa test as described above, but extended

to the three scoring categories of 0, 1, or 3. However, larger disagreements in the scores awarded for a criteria are more important than smaller disagreements (e.g. if one scorer awards a 0 for a given criterion, it would be more concerning if the second scorer awarded a 3 than a 1). As such, we also assessed agreement in scoring for each criterion using weighted kappa (Cohen 1968; Landis & Koch 1977; Viera & Garrett 2005; Shea et al. 2007). Matrices of the observed scores awarded by the two assessors were produced for each individual criterion, giving 13 separate matrices, each containing nine cells indicating the number of reviews awarded 0, 1, or 3 points by each assessor. Similar matrices of expected scores for each criterion were calculated as for a chi-squared test. A matrix containing nine cells representing the magnitude of disagreement between assessors was then constructed, e.g. a 1-0 disagreement is ranked as magnitude 1, whereas a 3-0 disagreement is ranked as magnitude 3 (Viera and Garrett, 2005, Shea et al., 2007). For an individual criterion, each cell in the observed matrix is then multiplied by the corresponding weight (e.g. cells where there is a 3-1 disagreement or a 1-3 disagreement are multiplied by 2). The observed weighted disagreement for that criterion is the sum of these values, with the expected weighted disagreement calculated in the same manner. The weighted kappa score for a criterion (which reflects *agreement*, and is interpreted in the same way as the unweighted kappa) is then:

κ = 1-(observed weighted disagreement/expected weighted disagreement)

Lastly, we divided the total CEESAT scores into three categories: 0-13, 13.5-26 and 26.5+ (reflecting an average score across the 13 criteria of 0-1, 1-2 and 2-3) to represent low, intermediate/moderate and high reliability (although see Woodcock et al., 2014 for further discussion regarding the interpretation of scores). Each review was assigned to one of these reliability categories based on the overall CEESAT score.

Evaluating the review landscape

Using our critical appraisal and assessment of review scope we then visually represented reviews examining the effectiveness of MPAs for biodiversity conservation and mitigating the impacts of fisheries in two matrices, one covering meta-analytical reviews and one summarising narrative syntheses. These matrices were designed to guide decision-makers to the most relevant and reliable reviews, and to enable easy visualisation of gaps and redundancy (multiple reviews on closely related topics) to target future

reviews. Detailed information indicating which reviews address each specific question is given in a series of supporting tables. Strengths of MPA reviews and aspects of review methods that could be improved were explored and evidence gaps and redundancy were identified.

Results

Searching and screening

Searches (Table A1) returned 2,485 results; these were refined to 287 after screening at title stage, 98 after abstract screening, and finally reduced to 24 relevant reviews following full-text examination. The bibliographies of relevant reviews were then hand-searched for additional references, giving a final total of 27 included reviews. To maintain transparency, a complete list of all included and excluded articles (at full text) together with reasons for exclusion is provided in Table A2-A3.

Review scores

Review scores ranged from 0-34 (mean = 12.3 $\pm$ 1.8 standard error [SE]), median = 13.5, Fig. 1a): note that because scores are the average across the two assessors, non-integer values are possible. Although no review achieved the maximum score of 39, the maximum possible points (3) were awarded for each criterion at least once. The majority of reviews (93%) achieved low ($\leq$13, N=13) or intermediate (13-26, N=12) scores. Criteria 3.1, 5.1 and 6.1 represented particular strengths (see Table 2 for explanation of criteria), whilst criteria 1, 3.2, 3.3 and 6.3 were consistent weaknesses in review conduct (Fig. 1b). Of the 27 reviews assessed, 18 contained meta-analysis and nine conducted a narrative synthesis. As would be expected, the mean score for meta-analyses was higher than for all reviews combined (mean = 17.3 $\pm$ 1.6 SE, median = 16), although a substantial range of scores was still evident (9.5-34).

Repeatability of scoring

The total scores awarded to each review were highly correlated between assessors (Spearman's ρ=0.96, $p<0.001$) and the mean absolute difference in scores was small (1.7 $\pm$ 0.3). Scoring for individual criteria was also generally consistent: for 11 of 13 criteria, agreement was greater than 70% and weighted kappa scores were around 0.7 or higher (Table 3; substantial agreement, Landis and Koch, 1977). The latter indicates that most disagreements were relatively minor (e.g. 0 vs 1, rather than 0 vs 3).

Scope of meta-analytical reviews

Most of the broad question elements in Table 1 were examined to some degree by meta-analyses but a clear focus was apparent with respect to the taxa (fish), MPA characteristics (size and age), and outcome measures (abundance) considered (Fig. 2, Table A4). For example, the vast majority of meta-analyses examined if/how MPA size influences the effectiveness of the protected area, with fewer investigations into the importance of other characteristics such as the level of protection (N=3), buffer zone size (N=2) or connectivity (N=1). No meta-analyses were detected that examined the effectiveness of MPAs in polar regions, or the effects of MPAs on birds, mammals or reptiles. At least one high-scoring review (≥26 points) was available for 11 of the broad questions, although these questions were also the subject of low-moderate scoring reviews. Two broad questions ('effects of MPAs on algae' and 'influence of connectivity on MPA effectiveness') were each only addressed by one moderate scoring review, representing cryptic evidence gaps that might not be readily apparent.

Gaps across review questions became more pronounced when specific questions were considered (Fig. 2, Tables A5-A10). Several more specific questions were the subject of multiple reviews with at least one high scoring review (e.g. 'fish' and 'abundance') demonstrating that there is some duplication in the review literature. However there are also examples of cryptic evidence gaps in which reviews were present but none were high-scoring (e.g. 'tropics' and 'species richness'). Furthermore, in addition to an absence of reviews considering polar regions, birds, mammals and reptiles, no reviews were identified for a further 15 specific questions, and an additional seven questions were either not addressed due to the low number of primary studies available, or were only addressed through a synthesis of <10 studies.

Scope of narrative syntheses

Narrative syntheses were generally of broader scope than meta-analyses (Fig. 3, Tables A11-A17). The majority (7 out of 9 narrative reviews) discussed the effects of MPAs globally rather than focusing on specific regions. The focus was split evenly between MPA effects on biodiversity and fisheries and most reviews considered MPAs as a whole rather than just highly protected (no-take) MPAs. No moderate- or high-scoring narrative reviews were identified (range in narrative review scores 0-12). Gaps and/or redundancy were noted in the majority of questions.

328 **Discussion**

329 The increasing importance of MPAs in global conservation strategies has stimulated extensive primary
330 research examining the effectiveness of MPAs for mitigating the impacts of fisheries (Lester et al., 2009,
331 Caveen et al., 2012). Reliably incorporating this research into policy requires syntheses that use
332 systematic, with objective methodologies to address key questions. However, our findings highlight
333 substantial variation in scope and rigour amongst reviews that examine the effectiveness of MPAs for
334 biodiversity conservation (Figure 1). This variation illustrates the need to ensure non-specialists can locate
335 the most rigorous reviews on questions of interest, and parallels that found in other fields of ecology and
336 environmental management (Philibert et al. 2012; Koricheva et al. 2014; O'Leary et al. 2016).
337 Furthermore, we found that reviews strongly favoured particular questions – e.g. exploring if the size and
338 age of an MPA influenced the effects on fish. Whilst these questions are vital for both biodiversity
339 conservation and fisheries management, there is a danger that findings from such syntheses could be
340 extrapolated to other taxonomic groups (e.g. birds, mammals) and that rigorous syntheses exploring the
341 importance of other MPA design characteristics (e.g. connectivity) will not be undertaken. The summary
342 of review scope and rigour provided here can assist future reviews in ensuring that the intended questions
343 complement rather than duplicate the existing review literature. Note that our study encompasses the time
344 period from 2000-2014. As with reviews of primary literature, the value of this information as a means to
345 explore review rigour and scope will thus be maximised if updates are conducted after a suitable time
346 period: by providing detailed methods and transparent descriptions we hope to facilitate such updates.

347 *General strengths and weaknesses in the conduct of reviews: implications for policy and research*

348 We found reviews to be of variable reliability with often overlapping scope (Figs. 2 and 3, Tables A4-
349 A17). Reviews regularly applied several approaches (e.g. meta-analytical techniques and transparent
350 reporting of inclusion criteria) that are important for rigorous synthesis (Fig. 1a and b). However, certain
351 aspects of MPA review conduct could be improved, such as ensuring that decisions over which articles are
352 relevant to include in the review are repeatable and transparent (by conducting kappa tests and listing all
353 articles read at full-text but excluded from the synthesis) and that critical appraisal of the methods of
354 included studies is undertaken and clearly reported (Fig. 1b). Narrative reviews were all assessed as being
355 of low reliability (N=9), partly reflecting the lack of quantitative synthesis. Nonetheless, there is no

inherent reason that narrative reviews cannot, for example, provide clear information on search strategies and scope, and document the extracted data. Given that narrative reviews can still contain valuable insights (and do influence policy) we argue that such reviews could benefit considerably from adopting such practices – indeed, one narrative review (Peppin et al. 2011) assessed during the initial evaluation of CEESAT achieved a score of 20 (Woodcock et al. 2014), which is similar to many of the meta-analyses considered here.

Scoring cannot distinguish between reviews undertaken using less rigorous methods and those that do not document rigorous methods where used. Just as with primary research, transparent reporting of review methods is vital, because it allows the review to be verified and updated. We therefore highlight the importance of effective reporting, and suggest that this represents a relatively straightforward means by which many reviews (narrative and meta-analyses) could be improved. More generally, we stress that in our view, limitations in reviews in environmental science reflect a lack of awareness of relevant systematic review methods, rather than a deliberate intention to mislead.

From a policy perspective, the large number of reviews with low-intermediate scores represents a potential problem. In low-intermediate scoring reviews, steps that are important for producing a comprehensive, objective, and transparent evidence synthesis are either absent or incomplete. Such limitations reduce the likelihood that the review provides an accurate picture of all available primary research. Although the effects of omitting certain steps on review reliability and findings are context-specific, in the absence of clear mechanisms to communicate the rigour of review methods to non-specialists, there is a risk that decision-makers will not take into account potential limitations in the conduct of the review(s) consulted.

Redundancy in the review literature

We identified substantial redundancy in the review literature (multiple reviews asking the same question) which could create difficulties for decision-makers looking to base decisions on the most robust synthesis available. In some instances, redundancy is a consequence of reviews providing effect sizes for broader questions and then exploring a range of more specific questions, or updating a particular question. Although such analyses are valuable for completeness and comparisons, decision-makers often lack the resources to locate and evaluate all relevant reviews. These situations therefore risk leading to policy and

practice that is not based on the most rigorous available evidence. As such, we hope that the results from studies such as ours can assist decision-makers in rapidly locating the reviews most likely to accurately synthesise all relevant evidence on the specific questions of interest. These outputs may also inform future research direction. For example, questions relating to fish abundance and MPA size have been the subject of reliable meta-analyses and so in the absence of substantial new research, attention might be better focused on synthesising evidence on other questions. Such investigations might include consideration of a broader range of taxa and MPA characteristics, as well as more specific factors that influence the effects of MPAs on fish abundance in order to inform on the degree to which findings are generalisable.

Gaps in the review literature

Gaps in the review literature are to be expected to an extent as a result of differences in public interest, policy relevance, availability of primary research (potentially influenced by e.g. logistical constraints in sampling fauna or flora), and question validity. However, some evidence gaps are in areas of high policy relevance. For example, the protection of seabirds and marine mammals is an important driver of MPA designation under European Directives, and MPAs are globally important tools in the conservation of a range of taxa (Hooker & Gerber 2004; Christianen et al. 2014). Furthermore, designation (or non-designation) can be controversial, and greater confidence in decisions would likely arise if robust evidence syntheses on the effectiveness of MPAs for multiple taxa were available. Our study also suggests some differences in the availability of reviews on tropical versus temperate MPAs. Relatively few meta-analyses quantify the effectiveness of MPAs in the former, particularly for less well-studied taxa and certain aspects of MPA design (note that global-scale reviews incorporating primary research from the tropics do not necessarily specifically evaluate the effectiveness of tropical MPAs). Decision-makers in the tropics would therefore be reliant on moderately reliable syntheses from this region and/or global syntheses that combine data from temperate and tropical MPAs. This contrasts with temperate regions, for which the effects of MPA size and age are specifically quantified by several reviews (Figure 2-3 and Supporting Information). There are also some instances in which reviews have been conducted but a highly rigorous synthesis is lacking (Figure 2-3). These could represent cryptic evidence gaps, in which the presence of existing reviews may create the perception that the question has been considered, and potentially deter new synthesis or primary research for several years.

Identification of gaps in the review literature highlights the need for new, more reliable syntheses (or primary research) to be conducted, providing a more solid basis for policy. Importantly, gaps become more frequent as questions become more specific, indicating that users should consider how applicable more general reviews are to particular contexts. It is important to stress that our primary emphasis was on properties relevant to the effectiveness of MPAs as a conservation tool for mitigating the impacts of fisheries on biodiversity. Valuable extensions of our study could therefore more specifically consider the review literature examining the extent to which MPAs provide fisheries benefits, as well as possible gaps in terms of the effects of MPAs on ecosystem functioning (e.g. productivity, nutrient cycling, food web structure), more sophisticated outcome metrics relating to conservation effectiveness (e.g. IUCN threat status), and the socioeconomic consequences of MPAs.

Conclusions

MPAs are a key component of global conservation strategies, but there is considerable uncertainty surrounding when and where reserves are most effective. Evidence reviews examining the effectiveness of MPAs are therefore likely to directly influence decision-making and future research. However, the overlapping scope and variation in reliability we identified amongst reviews presents a potentially important problem from the perspective of decision-makers seeking to make evidence-informed decisions. Our evaluation of reviews is intended to support decision-making by guiding non-specialists to the most reliable and relevant reviews. Findings from such reviews should be considered alongside other key pieces of evidence, in particular the extensive body of theoretical work on MPA effectiveness (e.g. Gaines et al. 2003; White et al. 2011) and more context-specific information relating to individual MPAs. Our findings can also assist researchers in identifying and targeting key knowledge gaps for review or new data collection including (but not limited to) ensuring broader taxonomic coverage, consideration of a wider range of MPA characteristics and examination of more specific questions for which we have identified evidence gaps.

Acknowledgements

We thank our potential end-users Ally Dingwall (Sainsbury's), Tom Pickerell (Seafish, Seafood Watch), Jon Harman (Seafish), Mike Mitchell, David Parker (Young's Seafood), David Jarrad (Shellfish

Association of Great Britain) for their contribution to discussions regarding review reliability. This project was supported by a UK Natural Environmental Research Council Knowledge Exchange Grant NE/J006386/1.

References

Butchart, S.H.M., Walpole, M., Collen, B., *et al.* (2010) Global Biodiversity: Indicators of Recent Declines. *Science* **328,** 1164-1168.

Caveen, A.J., Sweeting, C.J., Willis, T.J., Polunin, N.V.C. (2012) Are the scientific foundations of temperate marine reserves too warm and hard? *Environmental Conservation* **39,** 199-203.

CBD (2010) COP Decision X/2. Strategic plan for biodiversity 2011–2020. Available at: http://www.cbd.int/decision/cop/?id=12268. Accessed 6 April 2015.

CEE (2013) Guidelines for systematic review and evidence synthesis in environmental management.

Christianen, M.J.A., Merman, P.M.J., Bouma, T.J., *et al*. (2014) Habitat collapse due to overgrazing threatens turtle conservation in marine protected areas. *Proceedings of the Royal Society B* **281** 20132890, doi:10.1098/rspb.2013.2890.

Cohen, J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* **20**, 37-46

Cohen, J. (1968) Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin* **70** 213-220

Fox, H.E., Soltanoff, C.S., Mascia, M.B., *et al.* (2012) Explaining global patterns and trends in marine protected area (MPA) development. *Marine Policy* **36,** 1131-1138.

Gaines, S.D., Gaylord, B., Largier, J.L. (2003) Avoiding current oversights in marine reserve design. *Ecological Applications*, **13**, S32-46.

Gaines, S.D., Lester, S.E., Grorud-Colvert, K. et al. (2010) Evolving science of marine reserves: New developments and emerging frontiers. *Proceedings of the National Academy of Sciences* **107**, 18251-18255.

Grorud-Colvert, K., Claudet, J., Tissot, B.N., *et al.* (2014) Marine Protected Area networks: assessing whether the whole is greater than the sum of its parts. *PLoS ONE* **9,** e102298.

Halpern, B.S., Lester, S.E., McLeod, K.L. (2010) Placing marine protected areas onto the ecosystem-based management seascape. *Proceedings of the National Academy of Sciences* **107**, 18312-18317.

Halpern, B.S., Longo, C., Hardy, D.*, et al.* (2012) An index to assess the health and benefits of the global ocean. *Nature* **488,** 615-620.

Hays, G.C., Scott, R. (2013) Global patterns for upper ceilings on migration distance in sea turtles and comparisons with fish, birds and mammals. *Functional Ecology* **27,** 748-756.

Hooker, S.K., Gerber, L.R. (2004) Marine reserves as a tool for ecosystem-based management: the potential importance of megafauna. *BioScience* **54**, 27-39.

Huntington, B.E. (2011) Confronting publication bias in marine reserve meta-analyses. *Frontiers in Ecology and Environment* **9,** 375-376.

Koricheva, J. & Gurevitch (2014) Uses and misuese of meta-analysis in plant ecology. *Journal of Ecology* **102**, 828-844.

Landis, J.R., Koch, G.G. (1977) The measurement of observer agreement for categorical data. *Biometrics* **33,** 159-174.

Larsen, P.O., von Ins, M. (2010) The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. *Scientometrics* **84,** 575-603.

Lascelles, B.G., Langham, G.M., Ronconi, R.A., Reid, J.B. (2012) From hotspots to site protection: identifying marine protected areas for seabirds around the globe. *Biological Conservation* **156,** 5-14.

Lester, S.E., Halpern, B.S., Grorud-Colvert, K.*, et al.* (2009) Biological effects within no-take marine reserves: a global synthesis. *Marine Ecology Progress Series* **384,** 33-46.

Lubchenco, J., Grorud-Colvert, K. (2015) Making waves: The science and politics of ocean protection. *Science* **23**, 382-383.

Li, W., Zhao, Y. (2015) Bibliometric analysis of global environmental assessment research in a 20-year period. *Environmental Impact Assessment Review* **50,** 158-166.

Marine Conservation Institute (2015) MPAtlas. Available from: www.mpatlas.org. Accessed 27/10/2015.

Marinesque, S., Kaplan, D.M., Rodwell, L.D. (2012) Global implementation of marine protected areas: is the developing world being left behind? *Marine Policy* **36,** 727-737.

493 Møller, A.P., Jennions, M.D. (2001) Testing and adjusting for publication bias. *Trends in Ecology &*
494 *Evolution* **16,** 580-586.
495 Ojeda-Martinez, C., Bayle-Sempere, J.T., Sanchez-Jerez, P., Forcada, A., Valle, C. (2007) Detecting
496 conservation benefits in spatially protected fish populations with meta-analysis of long term
497 monitoring data. *Marine Biology* **151,** 1153-1161.
498 O'Leary, B.C., Kvist, K., Bayliss, H.R., Derroire, G., Healey, J.R., Hughes, K., Kleinschroth, F.,
499 Sciberras, M., Woodcock, P., Pullin AS (2016) The reliability of evidence review methodology in
500 environmental science and conservation. *Environmental Science and Policy* **64**, 75-82.
501 OSPAR (2010) The North-East Atlantic environment strategy of the OSPAR Commission for the
502 protection of the marine environment of the North-East Atlantic 2010-2020 (OSPAR Agreement
503 2010-2013).
504 Pautasso, M. (2012) Publication growth in biological sub-fields: patterns, predictability and sustainability.
505 *Sustainability* **4,** 3234-3247.
506 Peppin, D., Fule, P., Beyers, J. et al. (2011) Does seeding after severe forest firest in western USA mitigate
507 impacts on soils and plant communities. CEE review 08-023 (SR60). Collaboration for
508 Environmental Evidence. www.environmentalevidence.org/SR60.html.
509
510 Philibert, A., Loyce, C., Makowski, D. (2012) Assessment of the quality of meta-analyses in agronomy.
511 *Agriculture, Ecosystems & Environment* **148**, 72-82.
512 Ramirez-Llodra, E., Tyler, P.A., Baker, M.C*., et al.* (2011) Man and the last great wilderness: human
513 impact on the deep sea. *PLoS ONE* **6,** e:22588.
514 Sciberras, M., Jenkins, S.R., Kaiser, M.J., Hawkins, S.J., Pullin, A.S. (2013) Evaluating the biological
515 effectiveness of fully and partially protected marine areas. *Environmental Evidence* **2:4**.
516 Sciberras, M., Jenkins, S.R., Mant, R., Kaiser, M.J., Hawkins, S.J., Pullin, A.S. (2015) Evaluating the
517 relative conservation value of fully and partially protected marine areas. *Fish and Fisheries* **16,**
518 58-77.
519 Shea, B.J., Bouter, L.M., Peterson, J*., et al.* (2007) External validation of a measurement tool to assess
520 systematic reviews. *PLoS ONE* **2,** e:1350.

Thomas, H.L., Macsharry, B., Morgan, L. et al. (2014) Evaluating official marine protected area coverage for Aichi Target 11: appraising the data and methods that define our progress. *Aquatic Conservation: Marine and Freshwater Ecosystems* **24**, 8-23.

Viera, A.J., Garrett, J.M. (2005) Understanding interobserver agreement: the kappa statistic. *Family Medicine* **37,** 360-363.

White, J.W., Botsford, L.W., Baskett, M.L., *et al.* (2011) Linking models with monitoring data for assessing performance of no-take marine reserves. *Frontiers in Ecology and the Environment*, **9**, 390-399.

Willis, T.J., Millar, R.B., Babcock, R.C., Tolimiera, N. (2003) Burdens of evidencs and the benefits of marine reserves: putting Descartes bes des horse? *Environmental Conservation* **30,** 97-103.

Woodcock, P., Pullin, A.S., Kaiser, M.J. (2014) Evaluating and improving the reliability of evidence syntheses in conservation and environmental science: a methodology. *Biological Conservation* **176,** 54-62.

Supporting Information

Additional Supporting Information may be found in the online version of this article: **Tables A1-A17**.

537 **Table 1**: Key aspects of Marine Protected Areas (MPAs) that may influence effectiveness (geographic
538 region, taxon of interest, design characteristic), and outcome measures frequently used to assess MPA
539 effectiveness.

Region	Taxa	MPA Characteristic	Outcome Measure
Global	Fish	MPA Size	Abundance
Tropical	Invertebrate	MPA Age	Biomass
Temperate	Algae	MPA Connectivity	Species Richness
Polar	Mammal	MPA Buffer Zone Size	Organism Size
	Bird	MPA Protection Level	
	Reptile		

540

541 **Table 2:** Criteria and brief rationale for the Collaboration for Environmental Evidence Synthesis
542 Assessment Tool (CEESAT). See Woodcock et al. (2014) for details.

Criteria	Rationale
1.1 Was an *a-priori* protocol available for comment before the synthesis was conducted?	Increases robustness of review against *post hoc* changes in methods and scope.
2.1 Does the search for literature use a comprehensive range of resources?	Increases likelihood that all potentially relevant articles are captured by search.
2.2 Are the search strings clearly defined?	Allows search to be repeated and evaluated. Avoids open-ended searches.
3.1 Does the review apply clearly documented inclusion criteria to all potentially relevant studies found during the search?	Increases transparency. Reduces risk of subjective decisions influencing the studies included in the review.
3.2 Does the review demonstrate that inclusion decisions are repeatable?	Demonstrates repeatability of review, and that subjective decisions have not overly influenced the articles included.
3.3 Are inclusion/exclusion decisions transparent?	Ensures that the process of including and excluding studies can be externally verified.
4.1 Does the review report critical appraisals of the methods of each study?	Makes quality of the evidence-base for the synthesis clear.
4.2 Are studies objectively weighted according to methodological quality?	Gives greater emphasis to more robust studies.
5.1 Is data extraction documented, repeatable and consistent?	Reduces potential for bias in the extraction of metrics from individual studies.
5.2 Are the extracted data reported for each study?	Ensures that the extracted data can be verified and analysed by readers.
6.1 Is a quantitative synthesis conducted?	Reduces potential for subjectivity to influence data synthesis.
6.2 Is heterogeneity in the impact of the intervention investigated statistically?	Indicates the degree to which results are generalisable and the appropriateness of combining studies.
6.3 Does the review consider possible publication bias?	Reduces potential for bias arising from non-publication of non-significant or controversial results.

543

544 **Table 3: Agreement in scoring between reviewers.** Data shown for each criterion are % of reviews for
545 which the two reviewers awarded the same score, kappa test, and kappa test weighted by magnitude of
546 disagreement. Kappa score of 1 = perfect agreement, kappa score of 0 = agreement no different from that
547 expected by chance).

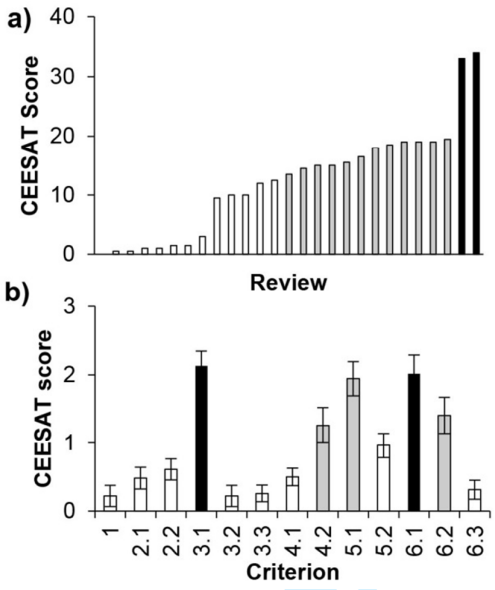
	Criteria	Agreement (%)	Kappa	Weighted Kappa
1.1	Protocol	100	1.00	1.00
2.1	Search resources	85	0.70	0.80
2.2	Search string stated	41	0.15	0.36
3.1	Documented inclusion criteria	74	0.51	0.72
3.2	Evidence that inclusion decisions repeatable	100	1.00	1.00
3.3	Documented exclusion decisions	85	0.53	0.67
4.1	Critical appraisal of methods	81	0.65	0.71
4.2	Objective weighting	78	0.62	0.70
5.1	Data extraction documented	78	0.59	0.70
5.2	Extracted data reported	59	0.35	0.49
6.1	Quantitative synthesis	96	0.92	0.97
6.2	Heterogeneity investigated	81	0.65	0.66
6.3	Publication bias considered	93	0.78	0.79

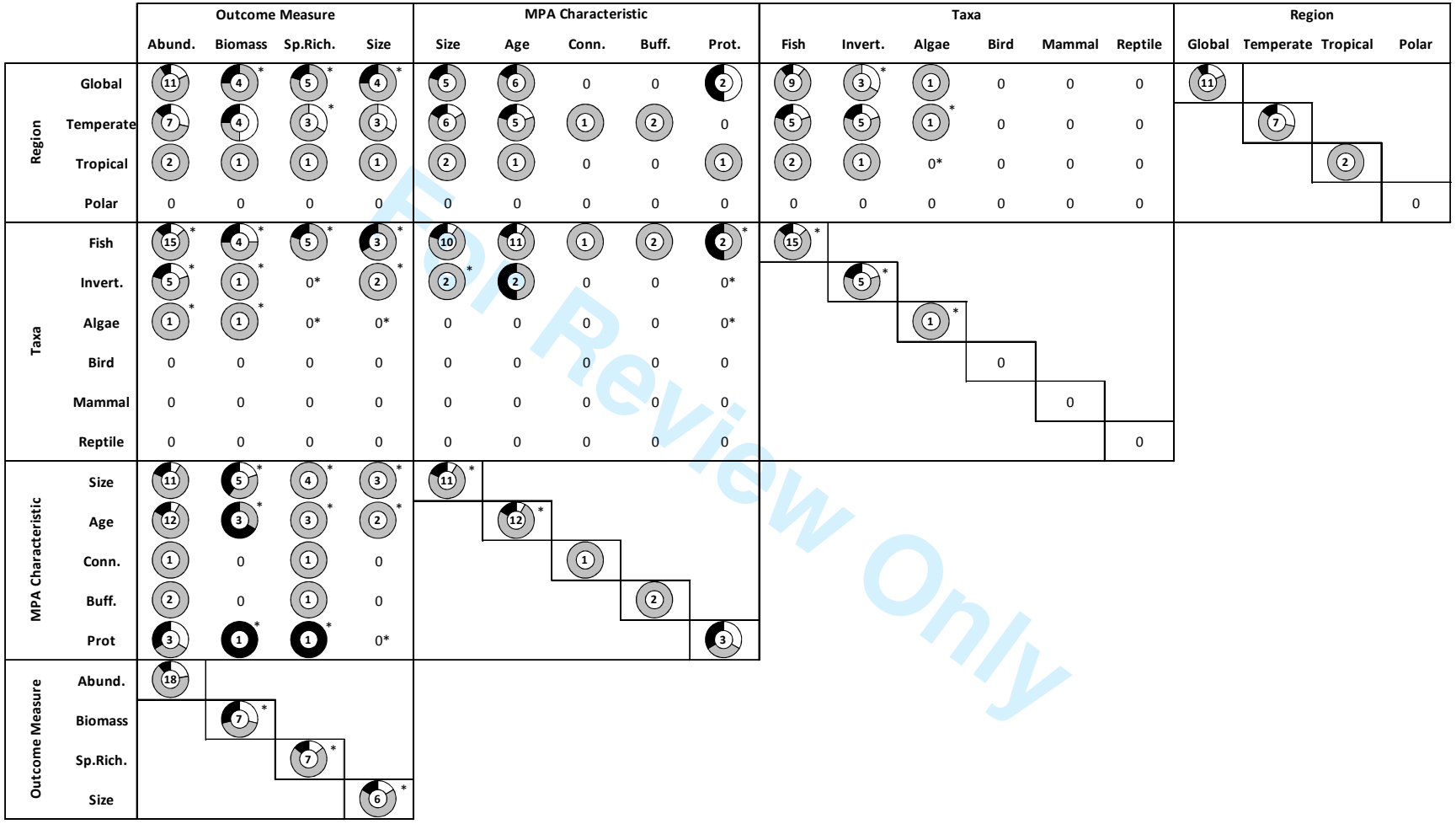
548

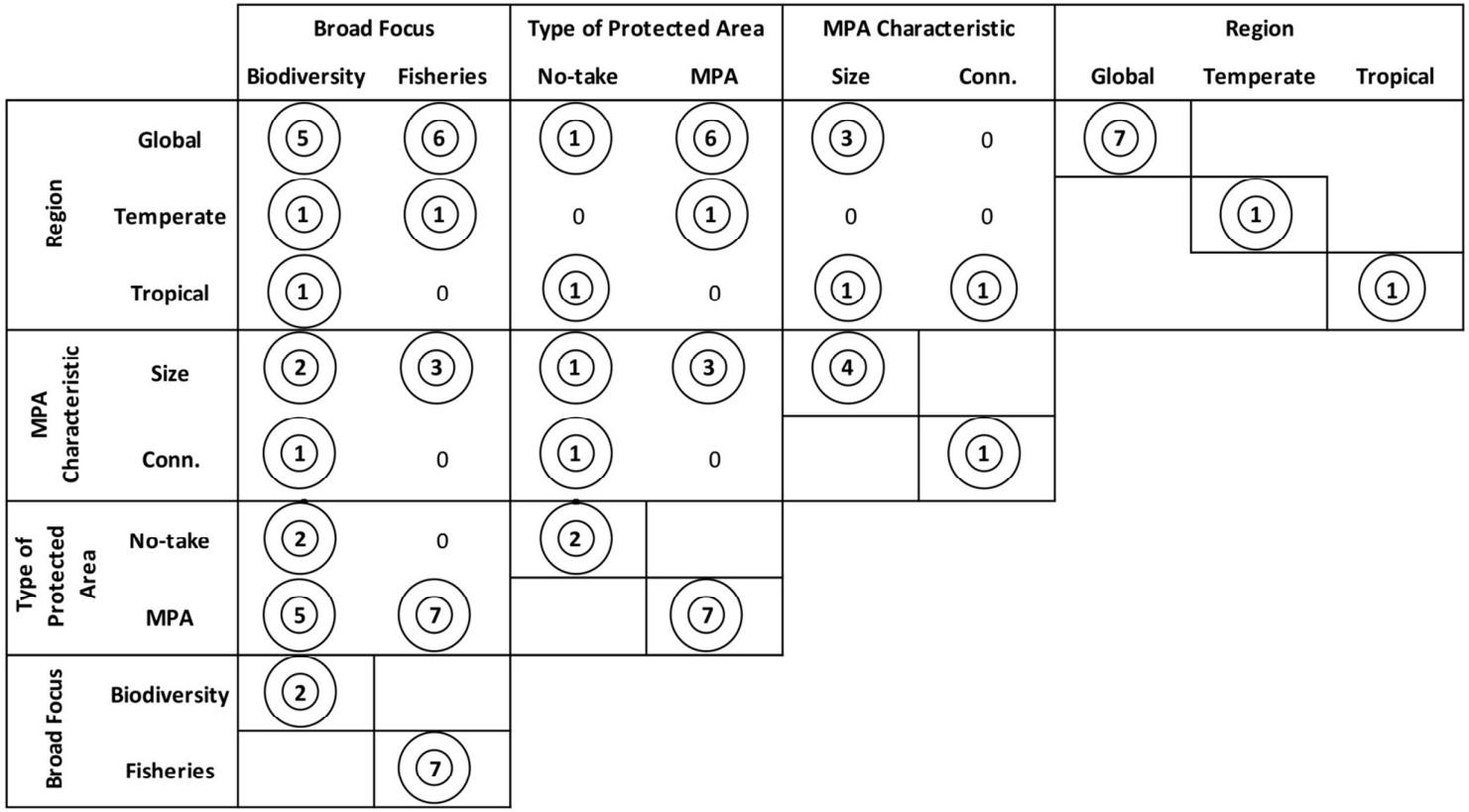
549 **Fig. 1**. **CEESAT scores for reviews examining the effectiveness of MPAs. (a)** total review scores, and
550 **(b)** mean score ± S.E. for each criterion. Scores are white (mean score per criterion of <1), grey (mean
551 score from 1-2), and black (mean score per criterion of >2). Higher scores indicate that the review
552 demonstrates greater objectivity, transparency, and comprehensiveness, and is therefore more likely to
553 provide an accurate reflection of the primary literature.

554 **Fig. 2: Matrix summarising the reliability and scope of meta-analytical reviews that examine MPA**
555 **effectiveness for biodiversity conservation.** Matrix overview of the 19 broad and 134 specific questions
556 we considered in our evaluation. Doughnut pie charts indicate the proportion of review achieving low (0-
557 13; white), moderate (13-26; grey), or high (>26; black) CEESAT scores. Total number of reviews
558 considering each question is in the centre of each chart. The matrix should be read using combinations
559 from the top and left headings to form the question of interest; relevant reviews can then be found in
560 Tables A4-10. For example, to explore the effect of MPA size on fish, locate MPA size under MPA
561 Characteristics in the top set of headings and read down to fish under Taxa on the left; consult Table A6
562 for details of reviews. Stars indicate reviews that considered the question but with <10 primary studies, or
563 stated that the question could not be investigated due to low number of primary studies. White areas
564 indicate questions that are not applicable, e.g. Global/Temperate question combinations. Abbreviations in
565 headings refer to: Outcome Measures - Abund=abundance and Sp.Rich=species richness; MPA
566 Characteristics - Conn=connectivity, Buff=buffer zone size, Prot=level of protection; and Taxa –
567 Invert=invertebrates.

568 **Fig. 3: Matrix summarising the reliability and scope of narrative syntheses that examine MPA**
569 **effectiveness for biodiversity conservation.** Matrix should be read using combinations from the top and
570 left headings to form the question of interest; full details of reviews can then be found in Tables A11-17.
571 For consistency, shading of doughnut pie charts are as for Figure 2. In practice, all narrative reviews we
572 assessed scored from 0-13, and so are coloured white. Blank areas indicate questions that are not
573 applicable, e.g. Global/Temperate question combinations. Abbreviation 'Conn.' in MPA Characteristics
574 refers to connectivity.






Broad Focus
Type of Protected Area
MPA Characteristic
Region
Biodiversity
Fisheries
No-take
MPA
Size
Conn.
Global
Temperate
Tropical
Region
Global
Temperate
Tropical
MPA Characteristic
Size
Conn.
Type of Protected Area
No-take
MPA
Broad Focus
Biodiversity
Fisheries
5 6 1 6 3 0 7
1 1 0 1 0 0 1
1 0 1 0 1 1 1
2 3 1 3 4
1 0 1 0 1
2 0 2
5 7 7
2
7

Supporting Information: Woodcock, P., O'Leary, B.C., Kaiser, M.J. and Pullin, A.S. *Fish & Fisheries*

This document contains:

Table A15: Scope of narrative syntheses that examine the effectiveness of marine protected areas: MPA Characteristic x Focus

Table A16 Scope of narrative syntheses that examine the effectiveness of marine protected areas: MPA Characteristic x Type of Protected Area

Table A17 Scope of narrative syntheses that examine the effectiveness of marine protected areas: Type of Protected Area x Focus

Table A1: Search strategy for locating reviews that examine the effectiveness of marine protected areas (MPAs). Resources and search strings are adapted from those presented in a recent systematic review of MPA effectiveness (Sciberras et al. 2013). All searches were limited to articles published after 2000 and in English.

Resource and Search Fields	Date	Search String
Web of Knowledge (topic)	14/5/2014	*("marine reserve*"* OR *"marine sanctuary"* OR *(marine* AND *"no-take zone")* OR *(marine* AND *harvest refug*)* OR *(marine* AND *"buffer zone")* OR *(marine* AND *partial** AND *protect*)* OR *(marine* AND *"closed area")* OR *(marine* AND *"area closure")* OR *(fisher** AND *(reserve* OR *closure))* OR *("fishing gear restriction*")* OR *("recreational fishing"* AND *protection)* OR *"marine protected area*")*
Aquatic Sciences and Fisheries Abstracts (advanced search, anywhere except full text)	14/5/2014	
CABI (all fields)	14/5/2014	
ScienceDirect (expert search, title-abstract-keywords)	14/5/2014	
Scopus (title, abstract, keywords)	15/5/2014	*AND* (*review* OR *"meta-analy*"* OR *synthes**)
Directory of Open Access Journals (advanced search, search all)	20/5/2014	*"marine reserve"* *"marine reserves"* *"marine sanctuary"* *marine* AND *"no take zone"* *marine* AND *"buffer zone"* *marine* AND *"harvest refuge"* *marine* AND *"harvest refugium"* *marine* AND *"closed area"* *marine* AND *"area closure"* *fisheries* AND *reserve* *fisheries* AND *closure* *"fishing gear restriction"* *"marine protected area"*
Index to Theses Online	20/5/2014	*"marine reserve"* *"marine sanctuary"* *"marine national park"* *"marine protected area"*
www.googlescholar.com (advanced search, with the exact phrase, anywhere in the article. First 50 hits examined in each search)	20/5/2014	*"marine reserve"* *"marine protected area"* *"marine sanctuary"* *(marine* AND *"harvest refuge")* The above search strings were combined individually with: AND (*review* OR *synthes** OR *"meta-analy*"*)

Resource and Search Fields	Date	Search String
JNCC: www.jncc.defra.gov.uk (site search) DEFRA: www.gov.uk (Publications > DEFRA, all publication types) CEFAS: http://www.cefas.defra.gov.uk/ (site search) NCCOS: http://ccma.nos.noaa.gov/ (search Publications) Department of Conservation, New Zealand: www.doc.govt.nz/	21/5/2014 21/5/2014 21/5/2014 21/5/2014 21/5/2014	*"marine protected area"* *"marine reserve"*
National MPA Center: http://marineprotectedareas.noaa.gov/ Resources > MPA Center Publications	21/5/2014	Hand-searched all Factsheets, and Reports, including Archives
Department of Sustainability, Environment, Water: http://www.environment.gov.au/ Publications > Marine > Marine Reserves (searched all)	21/5/2014	Hand-searched all articles

Table A2 Reviews assessed as relevant for inclusion, with review ID and CEESAT score. Meta-analyses are assigned number IDs and narrative syntheses letters.

Review ID	Reference	Score
a	Almany, G.R., S.R. Connolly, D.D. Heath, J.D. Hogan, G.P. Jones, L.J. McCook *et al.* (2009) Connectivity, biodiversity conservation and the design of marine reserve networks for coral reefs. *Coral Reefs*, **28**, 339-351.	**0.5**
b	Botsford, L.W., F. Micheli, A.M. Parma (2007) Biological and ecological considerations in the design, implementation and success of MPAs. *FAO Fisheries Report Background Paper 2,* 109-148	**1.5**
c	Burfeind, D.D., K.A. Pitt, R.M. Connolly & J.E. Byers (2013) Performance of non-native species within marine reserves. *Biological Invasions*, **15**, 17-28	**12**
[1]	Claudet, J. C.W. Osenberg, L. Benedetti-Cecchi, P. Domenici, J.-A. Garcia-Charton, A. Perez-Ruzafa, F. Badalmenti, J. Bayle-Sempere, A. Brito, F. Bulleri et al. (2008) Marine reserves: size and age do matter. *Ecology Letters*, **11**, 481-489	**16.5**
[2]	Claudet, J., C.W. Osenberg, P. Domenici, F. Badalmenti, M. Milazzo, J.M. Falcón, I. Bertocci, L. Benedetti-Cecchi, J.-A. Garcia-Charton, R. Goni *et al.* (2010) Marine reserves: fish life history and ecological traits matter. *Ecological Applications*, **20**, 830-839	**18.5**
[3]	Côté, I.M., I.M. Mosqueira & J.D.R. Reynolds (2001) Effects of marine reserve characteristics on the protection of fish populations: a meta-analysis. *Journal of Fish Biology*, **59**, 178-189.	**19.5**
[4]	Fenberg, P.B., J.E. Caselle, J. Claudet, M. Clemence, S.D. Gaines, J.A. García-Charton, E.J. Gonçalves, K. Grorud-Colvert, P. Guidetti, S.J. Jenkins *et al.* (2012) The science of European marine reserves: status, efficacy, and future needs. *Marine Policy*, **36**, 1012-1021.	**10**
d	García-Charton, J.A., A. Pérez-Ruzafa, C. Marcos, J. Claudet, F. Badalmenti, L. Benedetti-Cecchi, J.M. Falcón, M. Milazzo, P.J. Schembri, B. Stobart *et al.* (2008) Effectiveness of European Atlanto-Mediterranean MPAs: Do they accomplish the expected effects on populations, communities and ecosystems? *Journal for Nature Conservation*, **16**, 193-221	**1**
e	Gell, F.R. & C.M. Roberts (2002) The Fishery Effects of Marine Reserves and Fishery Closures. WWF-US. 1250 24th Street, NW, Washington, DC 20037, USA.	**0.5**
f	Gubbay, S. (2006) Marine protected areas. A review of their use for delivering marine biodiversity benefits. *English Nature Research Reports*, **688**.	**0**
[5]	Guidetti, P. & E. Sala (2007) Community-wide effects of marine reserves in the Mediterranean Sea. *Marine Ecology Progress Series*, **335**, 43-56.	**12.5**
[6]	Halpern, B.S. (2003) The impact of marine reserves: do reserves work and does reserve size matter. *Ecological Applications*, **13**, 117-137	**14.5**
[7]	Halpern, B.S. & R.R. Warner (2002) Marine reserves have rapid and lasting effects. *Ecology Letters*, **5**, 361-366.	**13.5**

Review ID	Reference	Score
[8]	Lester, S.E., B.S. Halpern, K. Grorud-Colvert, J. Lubchenco, B.I. Ruttenberg, S.D. Gaines *et al.* (2009) Biological effects within no-take marine reserves: a global synthesis. *Marine Ecology Progress Series*, **384**, 33-46.	**15.5**
[9]	Lester, S.E. & B.S. Halpern (2008) Biological responses in marine no-take reserves versus partially protected areas. *Marine Ecology Progress Series*, **367**, 49-56.	**10**
[10]	Maliao, R.J., A.T. White, A.P. Maypa & R.G. Turingan (2009) Trajectories and magnitude of change in coral reef fish populations in Philippine marine reserves: a meta-analysis. *Coral Reefs*, **28**, 809-822.	**18**
g	Mesnildrev, L., D. Gascuel & O. Le Pape (2013) Integrating marine protected areas in fisheries management systems: some criteria for ecological efficiency. *Aquatic Living Resources*, **26**, 159-170	**1.5**
[11]	Micheli, F., B.S. Halpern, L.W. Botsford & R.R. Warner (2004) Trajectories and correlates of community change in no-take marine reserves. *Ecological Applications*, **14**, 1709-1723.	**15**
[12]	Molloy, P.P., J.D. Reynolds, M.J.G. Gage, I. Mosqueira & I. M. Côté (2008) Links between sex change and fish densities in marine protected areas. *Biological Conservation*, **141**, 187-197.	**19**
[13]	Molloy, P.P., I.B. McLean & I.M. Côté (2009) Effects of marine reserve age on fish populations: a global meta-analysis. *Journal of Applied Ecology*, **46**, 743-751.	**19**
[14]	Mosqueira, I., I.M. Côté, S. Jennings & J.D. Reynolds (2000) Conservation benefits of marine reserves for fish populations. *Animal Conservation*, **4**, 321-332	**19**
h	Palumbi, S.R. (2001) The ecology of marine protected areas. In: Marine Community Ecology (eds: M. Bertness, S. Gaines & M. Hay). Sinauer Press.	**1**
[15]	Pande, A., A.B. MacDiarmid, P.J. Smith, R.J. Davidson, R.G. Cole, D. Freeman *et al.* (2008) Marine reserves increase the abundance and size of blue cod and rock lobster. *Marine Ecology Progress Series*, **366**, 147-158.	**15**
[16]	Ruttenberg, B., K. Grorud-Colvert, P. Guarderas, S. Lester & J. Lubchenco (2007) The science of marine reserves: a synthetic analysis of global effects and a new communication tool for Latin America and the Caribbean. *Proceedings of the 60th Gulf and Caribbean Fisheries Institute*	**9.5**
[17]	Sciberras, M., S.R. Jenkins, M.J. Kaiser, S.J. Hawkins & A.S. Pullin (2013) Evaluating the biological effectiveness of fully and partially protected marine areas. *Environmental Evidence*, **2:4**, http://environmentalevidencejournal.org/content/2/14	**34**
[18]	Stewart, G.B., I.M. Côté, M.J. Kaiser, B.S. Halpern, S.E. Lester, H.R. Bayliss *et al.* (2008) Are marine protected areas effective tools for sustainable fisheries management? Biodiversity impact of marine reserves in temperate zones. CEE review 06-002 (SR23). Collaboration for Environmental Evidence: www.environmentalevidence.org/SR23.html	**33**
i	Ward, T.J., D. Heinemann & N. Evans (2001) The Role of Marine Reserves as Fisheries Management Tools: A Review of Concepts, Evidence and International Experience. Bureau of Rural Sciences, Canberra, Australia	**3**

Table A3: Studies read at full-text but subsequently excluded. Note that some articles would have been excluded for several reasons in addition to those stated in the Table. There was ambiguity over inclusion/exclusion for some reviews – these were excluded but marked as 'Borderline'.

Review	Reason for Exclusion
Appeldoorn, R.S. & K.C. Lindeman (2003) A Caribbean-wide survey of marine reserves: spatial coverage and attributes of effectiveness. *Gulf and Caribbean Research*, **14**, 139-154	*Primarily an assessment of reserve characteristics (number, size etc.) rather than reserve effectiveness*
Baine, M. & J. Side (2003) Habitat modification and manipulation as a management tool. *Reviews in Fish Biology & Fisheries*, **13**, 187-199.	*Focus on artificial reefs, not MPAs*
Bergen, L.K. & M.H. Carr (2003) Marine reserves: how can science best inform policy? *Environment*, **45**, 8-19.	*Presented as a broad introductory overview, not synthesis of primary research*
Blount, B.G. & A. Pitchon (2007) An anthropological research protocol for marine protected areas. *Human Organisation*, **66**, 103-111.	*Focus on socioeconomic aspects of MPAs, and implications for anthropological research*
Botsford, L.W. (2005) Potential contributions of marine reserves to sustainable fisheries: recent modelling results. *Bulletin of Marine Science*, **76**, 245-259.	*Reviews the results from models, not field data collection.*
Botsford, L.W., F. Micheli & A. Hastings (2003) Principles for the design of marine reserves. *Ecological Applications*, **13**, 25-31.	*Focus is on reviewing results from models, not field data collection.*
Caveen, A.J., C.J. Sweeting, T.J. Willis & N.V.C. Polunin (2012) Are the scientific foundations of temperate marine reserves too warm and hard? *Environmental Conservation*, **39**, 199-203.	*Summarises scope of MPA research*
Christie, P. (2004) Marine protected areas as biological successes and social failures in Southeast Asia. *American Fisheries Society Symposium*, **42**, 155-164	*Primary research*
Dee, L.E., S.S. Horii & D.J. Thornhill (2014) Conservation and management of ornamental coral reef wildlife: successes, shortcomings, and future directions. *Biological Conservation*, **169**, 225-237	*Synthesis of regulations and management strategies relating to wildlife trade*
Diaz-Guisado, D., R.G. Cole, R.J. Davidson, D.J. Freeman, S. Kelly, A. Macdiarmid *et al.* (2012) Comparison of methodologies to quantify the effects of age and area of marine reserves on the density and size of targeted species. *Aquatic Biology*, **14**, 185-200	**Borderline relevant.** *Primary focus is in comparing the results from different meta-analytical approaches*

Review	Reason for Exclusion
Edgar, G.J. & N.S. Barrett (2002) Long-term monitoring of Tasmanian coastal MPA's. In: Examining Underwater Visual Census Techniques for the Assessment of Population Structure and Biodiversity in Temperate Coastal marine protected areas. Eds: N Barrett & C Buxton. *Tasmanian Aquaculture & Fisheries Institute Technical Report Series*, **11**, 46-55	*Workshop report describing approaches and findings from specific MPA monitoring programmes*
Fletcher, S., J. Saunders & R.J.H. Herbert (2011) A review of the ecosystem services provided by broad-scale marine habitats in England's MPA network. *Journal of Coastal Research*, **64**, 378-383	**Borderline relevant**. *Reviews types of ecosystem services potentially provided by MPAs, and types of evidence available (peer-reviewed, unpublished etc.), but does not synthesise the findings from primary research*
Fonteneau, A. (2007) Tuna management and closed areas. *International Commission for the Conservation of Atlantic Tunas: Collective Volume of Scientific Papers*, **60**, 190-223	**Borderline relevant.** *Judged to be an overview/report on MPA, with emphasis on underlying concepts and knowledge gaps, rather than synthesising primary research*
Francis, J., A. Nilsson & D. Waruinge (2002) Marine protected areas in the Eastern African region: how successful are they? *Ambio*, **31**, 503-511	*Some information on MPA effectiveness, but main focus judged to be on the history, management, and characteristics of MPAs*
Fraschetti, S., A. Terlizzi, F. Micheli, L. Benedetti-Cecchi & F. Boero (2002) Marine protected areas in the Mediterranean Sea: objectives, effectiveness and monitoring. *Marine Ecology*, 23, 190-200	*Focus on evaluating and informing the design of research into MPA effectiveness*
Gaines, S.D., C. White, M.H. Carr & S.R. Palumbi. (2010) Designing marine reserve networks for both conservation and fisheries management. *Proceedings of the National Academy of Sciences*, **107**, 18286-18293	**Borderline relevant.** *Judged to be focused more on synthesising ecological concepts than primary research into MPA effectiveness.*
Game, E.T., H.S. Grantham, A.J. Hobday, R.L. Pressey, A.T. Lombard, L.E. Beckley, K. Gjerde, R. Bustamente, H.P. Possinghame & A.J. Richardson. (2009) Pelagic protected areas: the missing dimension in ocean conservation. *Trends in Ecology & Evolution*, **24**, 360-369	*Article is marked as an Opinion piece (i.e. not presented as an objective assessment)*
Garcia-Charton, J.A., I.D. Williams, A. Perez Ruzafa, M. Milazzo, R. Chemello, C. Marcos, M.-S. Kitsos, A. Koukouras & S. Riggio. (2000) Evaluating the ecological effects of Mediterranean marine protected areas: habitat, scale and the natural variability of ecosystems. *Environmental Conservation*, **27**, 159-178	**Borderline relevant.** *Stated aim to: '*address the methodological implications of marine ecosystem heterogeneity for the design of sampling strategies to evaluate the ecological effects of MPAs, and identify knowledge gaps...' *Not principally intended to synthesise findings from primary research into MPA effectiveness.*

Review	Reason for Exclusion
Gaston, K.J., S. F. Jackson, L. Cantu-Salazar & G. Cruz-Pinon (2008) The ecological performance of protected areas. *Annual Review of Ecology, Evolution and Systematics*, **39**, 93-113	*Broad overview of protected area concepts, and considers all types of reserve (not just marine)*
Gell, F.R. & Roberts, C.M. (2003) Benefits beyond boundaries: the fishery effects of marine reserves. *TRENDS in Ecology and Evolution*, **18**, 448-455	*Article is marked as an Opinion piece (i.e. not presented as an objective assessment)*
Gianni, F., F. Bartolini, F., L. Airoldi, E. Ballesteros, P. Francour, P. Guidetti, A. Meinesz, T. Thibaut & L. Mangialajo (2013) Conservation and restoration of marine forests in the Mediterranean Sea and the potential role of marine protected areas. *Advances in Oceanography and Limnology*, **4**, 83-101	*Focus on the conservation of marine forests, with MPAs as a potential tool*
Goodsell, P.J. & A.J. Underwood (2009) Protection of biota and the value of marine protected areas. *Ecological Studies*, **206**, 345-355	*Some consideration of evidence for effectiveness of MPAs, but a broad overview with a stated focus on* 'the threats and processes that necessitate protection'
Gotz, A., S.E. Kerwath & C.G. Attwood (2013) A step-by-step framework to assess benefits of established temperate marine protected areas. *South African Journal of Science*, **102**, 36-44	*Focus on developing a framework for assessing MPA effectiveness, and testing this using a case study*
Grafton, R.Q., T. Kompas & V. Schneider (2005) The bioeconomics of marine reserves: a selected review with policy implications. *Journal of Bioeconomics*, **7**, 161-178	*Includes a synthesis of the effects of marine reserves, but focus is more on bioeconomic modelling of reserve effects*
Green, A.L., L. Fernandes, G. Almany, R. Abesamis, E. McLeod, P. M. Alino, A.T. White, R. Salm, J. Tanzer & R.L. Pressey (2014) Designing marine reserves for fisheries, management, biodiversity conservation, and climate change adaptation. *Coastal Management*, **42**, 143-159	**Borderline relevant.** *Judged to focus more on providing advice for MPA design rather than evaluating whether or not reserves are effective*
Gruss, A., D.M. Kaplan, S. Guenetter, C.M. Roberts & L.W. Botsford (2011) Consequences of adult and juvenile movement for marine protected areas. *Biological Conservation*, **144**, 692-702	*Using information on fish movement patterns to make recommendations on reserve design – i.e. not synthesising primary research on MPA effectiveness*
Guidetti, P., M. Milazzo, S. Bussotti, A. Molinari, M. Murenu, A. Pais, N. Spano, R. Balzano, T. Agardy, F. Boero, G. Carrada, R. Cattaneo-Vietti, A. Cau, R. Chemello, S. Greco, A. Manganaro, G. N. di Sciara, G. F. Russo & L. Tunesi (2008) Italian marine reserve effectiveness: Does enforcement matter? *Biological Conservation*, **141**, 699-709.	*Uses meta-analysis but appears to be analysing primary data*

Review	Reason for Exclusion
Halpern, B.S., S.E. Lester & J.B. Kellner (2010) Spillover from marine reserves and the replenishment of fished stocks. *Environmental Conservation*, **36**, 267-276	*Uses synthesis to develop model parameters*
Hatton-Ellis, T. (2005) Fish communities and fisheries in Wales's National nature Reserves: A review. *Freshwater Forum*, **24**, 82-104	*Focus on freshwater protected areas*
Higgins, R.M., F. Vandeperre, A. Perez-Ruzafa & R.S. Santos (2008) Priorities for fisheries in marine protected area design and management: implications for artisanal-type fisheries as found in southern Europe. *Journal for Nature Conservation*, **16**, 222-233.	*Judged to be a broad overview of MPA management and design*
Hilborn, R., K. Stokes, J.-J. Maguire, T. Smith, L.W. Botsford, M. Mangel, J. Orensanz, A. Parma, J. Rice, J. Bell, K.L. Cochrane, S. Garcia, S.J. Hall, G.P. Kirkwood, K. Sainsbury, G. Stefansson & C. Walters (2004) When can marine reserves improve fisheries management? *Ocean & Coastal Management*, **47**, 197-205	**Borderline relevant.** *Excluded because provides a broad overview of advantages and disadvantages of MPAs (including value in understanding long-term trends, potential increased fishing effort elsewhere, social impacts) – i.e. not primarily focused on direct effects of MPAs on biodiversity.*
Hooker, S.K. & L.R. Gerber (2004) Marine reserves as a tool for ecosystem-based management: the potential importance of megafauna. *BioScience*, **54**, 27-39	*MPA design, effectiveness and importance primarily assessed by synthesis and development of concepts*
Horigue, V., P.M. Alino & R.L. Pressey (2014) Evaluating management performance of marine protected area networks in the Philippines. *Ocean & Coastal Management*, **95**, 11-25	*Evaluating MPA management practices*
Huntingdon, B.E. (2011) Confronting publication bias in marine reserve meta-analysis. *Frontiers in Ecology & the Environment*, **9**, 375-376	*Assessing rigour of meta-analyses on MPA effectiveness with respect to treatment of potential publication bias*
Kemp, J., G.P. Jenkins, D.C. Smith & E. Fulton (2012) Measuring the performance of spatial management in marine protected areas. *Oceanography & Marine Biology*, **50**, 287-314	*Excluded because (i) expressly states a focus on selected case study reserves and (ii) emphasis is more on trends within selected MPAs over time*
Koldeway, H.J., D. Curnick, S. Harding, L.R. Harrison & M. Gollock (2010) Potential benefits to fisheries and biodiversity of the Chagos Archipelago/British Indian Ocean Territory as a no-take marine reserve. *Marine Pollution Bulletin*, **60**, 1906-1915	*Evaluates evidence to determine potential impacts and benefits of a proposed MPA*

Review	Reason for Exclusion
Magdaong, E.T., Fujii, M., Yamano, H., Licuanan, W.Y., Maypa, A., Campos, W.L., Alcala, A.C., White, A.T., Apistar, D. & Martinex, R. (2014) Long-term change in coral cover and the effectiveness of MPAs in Philippines: A meta-analysis. *Hydrobiologia*, **733**, 5-17	**Borderline relevant** *Excluded because it is a synthesis of studies that have been monitoring coral cover, not a synthesis of BA/CI/BACI studies.*
Maliao, R.K., R.S. Pomeroy & R.G. Turingan (2009) Performance of community-based coastal resource management (CBCRM) *Marine Policy*, **33**, 818-825	*Focus on perceptions of MPA management/governance, assessed through questionnaire*
Martin, K., M.A. Samoilys, A.K. Hurd, I. Meliane & C.G. Landin (2007) Experiences in the use of marine protected areas with fisheries management objectives – a review of case studies. Background Paper 1. In: Report and Documentation of the Expert Workshop on Marine Protected Areas and Fisheries Management: Review of Issues and Considerations. *Food & Agriculture Organisation Fisheries Report*, **825**, p21-108	*Broad overview, expressly based on selected case studies*
McCook, L.J., T. Ayling, M. Cappo, J.H. Choat, R.D. Evans, D.M. De Freitas, M. Heupel, T.P. Hughes, G.P. Jones, B. Mapstone, H. Marsh, M. Mills, F.J. Molloy, C.R. Pitcher, R.L. Pressey, G.R. Russ, S. Sutton, H. Sweatman, R. Tobin, D.R. Wachenfeld & D.H. Williamson (2010) *Proceedings of the National Academy of Sciences USA*, **107**, 18278-18285	*Focus is on the effects of zoning, rather than the effects of the MPA. Also marked as a Perspective*
Mora, C. & P.F. Sale (2011) Ongoing global biodiversity loss and the need to move beyond protected areas: a review of the technical and practical shortcomings of protected areas on land and sea. *Marine Ecology Progress Series*, **434**, 251-266	*Broad overview of strengths and weaknesses of the protected areas concept in relation to aquatic and terrestrial conservation*
Nagelkerken, I., M.G.G. Grol & P.J. Mumby (2012) Effects of marine reserves versus nursery habitat availability on structure of reef fish communities. *PLOS ONE*, **7**, e36906. doi:10.1371/journal.pone.0036906	*Primary research*
Nikolaeva, N.G., V.A. Spiridonov & Y.V. Krasnov (2006) Existing and proposed marine protected areas and their relevance for seabird conservation: a case study in the Barents Sea region. In: *Waterbirds around the world.* Eds. G.C. Boere, C.A. Galbraith & D.A. Stroud. The Stationery Office, Edinburgh, UK. pp743-749	*Outline of bird distributions with reference to proposed and existing MPAs*

Review	Reason for Exclusion
Natural England Research Report (2010) Guidance on the size and spacing of marine protected areas in England. NECR037	*Recommendations on MPA design based on species ecology and distributions*
Ojeda-Martinez, C., J.T. Bayle-Sempere, P. Sanchez-Jerez, A. Forcada & C. Valle (2007) Detecting conservation benefits in spatially protected fish populations with meta-analysis of long term monitoring data. *Marine Biology*, **151**, 1153-1161	*Synthesis of long-term monitoring data – judged not to be intended as a review of primary research that examines MPA effectiveness using a BA or CI approach*
Ojeda-Martinez, C., J.T. Bayle-Sempere, P. Sanchez-Jerez, F. Salas, B. Stobart, R. Goni, J.M. Falcon, M. Graziano, I. Guala, R. Higgins, F. Vandeperre, L. Le Direach, P. Martin-Sosa & S. Vaselli (2011) Review of the effects of protection in marine protected areas: current knowledge and gaps. *Animal Biodiversity & Conservation*, **34**, 191-203	*Describing and summarising the characteristics of studies conducted on MPAs*
Pelc, R.A., R.R. Warner, S.D. Gaines & C.B. Paris (2010) Detecting larval export from marine reserves. *Proceedings of the National Academy of Sciences USA*, **107**, 18266-18271	*Primarily modelling study that uses a review to check model findings*
Perera, N. & A. de Vos (2007) Marine protected areas in Sri Lanka: a review. *Environmental Management*, **40**, 727-738	*Overview of status (history, design, management, governance etc.) of MPAs*
Pinnegar, J.K., N.V.C. Polunin, P. Francour, F. Badalmenti, R. Chemello, M.-L. Harmelin-Vivien *et al.* (2000) Trophic cascades in benthic marine ecosystems: lessons for fisheries and protected-area management. *Environmental Conservation*, **27**, 179-200	*Documenting examples of trophic cascades (predicted or actual) inside and outside MPAs, rather than examining the effects of MPAs*
Planes, S., R. Galzin, A. Garcia Rubies, R. Goni, J.-G. Harmelin, L. Le Direacg, P. Lenfant & A. Quetglas (2000) Effects of marine protected areas on recruitment processes with special reference to Mediterranean littoral ecosystems. *Environmental Conservation*, **27**, 126-143	**Borderline relevant.** *Judged to focus largely on the ecological concepts influencing MPA effectiveness, rather than reviewing primary research that directly examines MPA effects.*
Rife, A.N., B. Erisman, A. Sanchez & O. Aburto-Oropeza (2013) When good intentions are not enough...Insights on networks of "paper park" marine protected areas. *Conservation Letters*, **6**, 200-212	*Article marked as Policy Perspective, so not intended as an objective synthesis*
Sale, P.F., R.K. Cowen, B.S. Danilowicz, G.P. Jones, J.P. Kritzer, K.C. Lindeman, *et al.* (2005) Critical science gaps impede use of no-take fishery reserves. *Trends in Ecology & Evolution*, **20**, 74-80	*Article marked as Opinion, so not intended as an objective synthesis*

Review	Reason for Exclusion
Samy, M. J.L. Sanchez Lizaso & A. Forcada (2011) Status of marine protected areas in Egypt. *Animal Biodiversity and Conservation*, **34**, 165-177	*Some consideration of effects of MPA on biodiversity, but part of a general overview describing location, governance, threats etc.*
Sanchez Lizaso, J.L., R. Goni, O. Renones, J.A. Garcia Charton, R. Galzin, J.T. Bayle, P. Sanchez Jerez, A. Perez Ruzafa & A.A. Ramos (2000) Density dependence in marine protected populations: a review. *Environmental Conservation*, **27**, 144-158	*Some synthesis of primary research on the effects of MPAs, but focus is more on synthesising general ecological concepts and research then applying this to reserves*
Sumaila, U.R., S. Guenette, J. Alder & R. Chuenpagdee (2000) Addressing ecosystem effects of fishing using marine protected areas. *ICES Journal of Marine Science*, **57**, 752-760	*Broad overview of MPAs with extensive treatment of modelling approaches for MPA management and limited synthesis of primary research examining MPA effects*
Valentine, J.F. & K.L. Heck Jr (2005) Perspective review of the impacts of overfishing on coral reef food web linkages. *Coral Reefs*, **24**, 209-213	*Article marked as Perspective*
Vandeperre, F., R.M. Higgins, J. Sanchez-Meca, F. Maynou, R. Goni, P. Martin-Sosa, A. Perez-Ruzafa, P. Afonso, I. Bertocci, R. Crec'hiou, G. D'Anna, M. Dimech, C. Dorta, O. Esparza, J.M. Falcon, A. Forcada, I. Guala, L. Le Direach, C. Marcos, C. Ojeda-Martinez, C. Pipitone, P.J. Schembri, V. Stelzenmuller, B. Stobart & R.S. Santos (2011) Effects of no-take area size and age of marine protected areas on fisheries yields: a meta-analytical approach. *Fish & Fisheries*, **12**, 412-426	**Borderline relevant**. *Excluded because meta-analysis is used to explore selected datasets and case studies to analyse trends*
Wenzel, L., J. Gass, M. D'Iorio & J. Blackburn (2013) A national perspective on the role of marine protected areas in sustaining fisheries. *Fisheries Research*, **144**, 23-27	*Overview of MPA management and locations*
Willis, T.J., R.B. Millar, R.C. Babcock & N. Tolimieri (2003) Burdens of evidence and the benefits of marine reserves: putting Descartes before des horse? *Environmental Conservation*, **30**, 97-103	*Mostly focused on summarising the types of MPA studies conducted and evaluating the methods used. Article also marked as Comment*
Willis, T.J. (2013) Scientific and biodiversity values of marine reserves: a review. *DOC Research and Development Series.* New Zealand Department of Conservation, Wellington	**Borderline relevant.** *More of a broad overview of the field, with emphasis on summarising the types of research conducted, providing information on study design, describing research conducted within MPAs (whether or not this examines effectiveness)*

Review	Reason for Exclusion
Wilson, B. (2005) Marine protected areas for whales, dolphins and porpoises: a world handbook for cetacean habitat. *Environmental Conservation*, **32**, 283-284	*Book review. *NB: Based on the review, the book appears to focus primarily on describing the history, management, context and distribution of MPAs*
Witherell, D. & D. Woodby (2005) Application of marine protected areas for sustainable production and marine biodiversity off Alaska. *Marine Fisheries Review*, **67**, 1-27	*Overview of management and history of MPAs. Some data synthesis, but based on trends detected in monitoring programmes*

Table A4: Scope of meta-analyses that examined the effectiveness of marine protected areas. Each number corresponds to a review – see Table A2 for details. Dashes indicate that no meta-analyses addressing particular questions were located.

	SCOPE	REVIEW
REGION	Global	3, 6-9, 11-14, 16, 17
	Tropical [a]	8, 10
	Temperate [a]	1, 2, 4, 5, 8, 15, 18
	Polar [a]	-
TAXA	Fish [b, c, d]	1-3, 5, 6, 8, 10-18
	Invertebrate [b, c, d]	6, 8, 15, 16, 18
	Algae [c]	8
	Bird	-
	Mammal	-
	Reptile	-
MPA CHARACTERISTIC	Size [e, f]	1-3, 5, 6, 8, 10, 11, 15, 17, 18
	Age[e]	1-3, 5, 7, 10-13, 15, 17, 18
	Connectivity	1
	Buffer Zone Size	1, 2
	Protection Level	9, 10, 17
OUTCOME MEASURE	Abundance	1-18
	Biomass [g, h, i]	4-8, 17, 18
	Species Richness [g, i, j]	1, 3, 4, 6-8, 17
	Organism Size [g, i]	4, 6-8, 15, 17

Notes:
[a] Review [3] found no influence of latitude on MPA effectiveness, and Review [13] included latitude as a potential confounding factor. Neither review directly reports effect sizes for reserves in different regions.
[b] Review [7] found consistent results across fish and invertebrates but did not present taxon-specific data.
[c] Review [9] recognised the potential importance of differences amongst taxa but was unable to explore this due to low number relevant primary studies. Review [17] investigated the effects of MPAs on invertebrates, but with <10 primary studies. Review [18] investigated the effects of MPAs on algae, but with <10 primary studies.
[d] Review [15] investigated MPA effectiveness at the level of two individual species (one fish, one invertebrate) rather than across assemblages.
[e] Review [9] recognised the potential importance of MPA size and MPA age, but was unable to explore this due to low number of relevant primary studies.
[f] Review [13] included MPA size as a potential confounding factor, but did not directly report effect size.
[g] Review [9] examined effect of MPAs on biomass, species richness, and organism size, but with <10 primary studies.
[h] Review [11] unable to consider effects of MPAs on biomass due to low number of relevant primary studies.
[i] Review [16] collected data on other outcome measures, but only directly presents data on abundance.
[j] Review [18] investigated effects of MPAs on species richness, but with <10 primary studies.

Table A5: Scope of meta-analyses that examine the effectiveness of marine protected areas. Combinations of Region and Taxa are shown. Each number corresponds to a review – see Table A2 for details.

		REGION		
		Global	Temperate	Tropical
TAXA	Fish	3, 6, 8, 11-14, 16, 17	1-2, 5, 8, 15, 18, [b, c]	8, 10, [b, c]
	Invertebrates	6, 8, 16, [a]	8, 15, 18	8
	Algae	8	8	[e]

Notes: [a] Review [17] investigated the effects of MPAs on invertebrates at a global scale, but with <10 primary studies.
[b] Review [3] found no influence of latitude on MPA effectiveness for fish, but did not directly report effect size for reserves in different regions.
[c] Review [13] examined effectiveness of MPAs for fish and included latitude as a potential confounding factor, but did not directly report effect size for reserves in different regions.
[d] Review [18] investigates effect of MPAs on temperate algae, but based on 9 studies.
[e] Review [8] investigates effect of MPAs on tropical algae, but only 5 studies were available.

Table A6: Scope of meta-analyses that examine the effectiveness of marine protected areas. Combinations of Taxa and MPA Characteristics are shown. Each number corresponds to a review – see reference list for details. Dashes indicate that no meta-analyses addressing particular combinations were located.

		MPA CHARACTERISTIC				
		Size	Age	Protection	Buffer Zone Size	Connectivity
TAXA	Fish	1-3, 5, 6, 10, 11, 15[a], 17, 18, [b, c]	1-3, 5, 10-13, 15[a], 17, 18, [e]	10, 17, [f]	1, 2	1
	Invertebrates	6, 15[a], [b, d]	15[a, d]	[d, f]	-	-
	Algae	-	-	[f]	-	-

Notes: [a] Review [15] focuses on two species (one invertebrate, one fish).
[b] Review [8] examined effect of MPA size but did not present this relationship for specific taxa.
[c] Review [13] examined effectiveness of MPAs for fish and included MPA size as a potential confounding factor, but did not directly report effect size for relationship between reserve size and effectiveness.
[d] Review [17] unable to investigate effect of MPA characteristics (size, age, level of protection) on invertebrates due to low number of relevant primary studies.
[e] Review [7] states that effect of MPA age is consistent for fish and invertebrates but does not present effect sizes for each taxon separately.
[f] Review [9] unable to investigate effect of MPA protection level on different taxa separately due to low number of relevant primary studies.

Table A7: Scope of meta-analyses that examine the effectiveness of marine protected areas. Combinations of taxa and Outcome Measure are shown. Each number corresponds to a review – see Table A2 for details.

		OUTCOME MEASURE			
		Abundance	Biomass	Species Richness	Organism Size
TAXA	Fish	1-3, 5, 6, 8, 10-18 [a, b]	5, 6, 8, 17, [a, b, c, d]	1, 3, 6, 8, 17, [a, b, d]	8, 15, 17, [a, b, f]
	Invertebrates	6, 8, 15, 16, 18, [a, b, e]	8, [a, b, d, e, f]	a, b, d, e, f, g	8, 15 , [a, b, e, f]
	Algae	8, [b, d]	b, d, g	b, d, g	b, g

Notes: [a] Review [7] states that effectiveness of MPAs is consistent for fish and invertebrates but does not present effect sizes for each taxon separately.
[b] Review [9] unable to disaggregate results by taxa due to low number of relevant primary studies.
[c] Review [11] unable to investigate effects of MPAs on fish biomass due to low number of relevant primary studies.
[d] Review [18] investigated effects of MPAs but with <10 primary studies for the following combinations: fish biomass and species richness, invertebrate biomass and species richness, algae density, biomass, and species richness.
[e] Review [17[attempted to investigate effects of MPAs on invertebrates, but <10 relevant primary studies were available.
[f] Review [6] investigated the effect of MPAs on fish size, but with <10 primary studies for each of the trophic groups considered, and also investigated the effects of MPAs on invertebrate biomass, species richness, and size, but again <10 relevant primary studies were available in each case.
[g] Review [8] investigated the effect of MPAs on invertebrate species richness, and on biomass, species and organism size for algae, but with <10 primary studies.

Table A8: Scope of meta-analyses that examine the effectiveness of marine protected areas. Combinations of Region and MPA Characteristic are shown. Each number corresponds to a review – see Table A2 for details. Dashes indicate that no meta-analyses addressing particular combinations were located.

		MPA CHARACTERISTIC				
		Size	Age	Protection	Buffer Zone Size	Connectivity
REGION	Global	3, 6, 8, 10, 17, [a]	3, 7, 10, 12, 13, 17	9, 17	-	-
	Temperate	1, 2, 5, 8, 15, 18	1, 2, 5, 15, 18	-	1, 2	1
	Tropical	8, 10	10	10	-	-

Notes: [a] Review [13] included MPA size and latitude as potential confounding factors, but did not directly explore these properties.

Table A9: Scope of meta-analyses that examine the effectiveness of marine protected areas. Combinations of MPA Characteristic and Outcome Measure are shown. Each number corresponds to a review – see Table A2 for details. Dashes indicate that no meta-analyses addressing particular combinations were located.

		MPA CHARACTERISTIC				
		Size	Age	Protection	Buffer Zone Size	Connectivity
OUTCOME MEASURE	Abundance	1-3, 5, 6, 8, 10, 11, 15, 17, 18, [a, d]	1-3, 5, 7, 10-13, 15, 17, 18	9, 10, 17	1, 2	1
	Biomass	6, 8, 17, 18, [a, c]	7, 17, 18, [a, c]	17, [b]	-	-
	Richness	1, 3, 6, 8, [e, f]	1, 3, 7, [e, f]	17, [b]	1	1
	Organism Size	6, 8, 15, [e]	7, 15, [e]	[b, e]	-	-

Notes: [a] Review [5] used density data to investigate the effect of MPA size and MPA age on community composition, but was unable to investigate effects of MPA size and age on biomass due to low number of relevant primary studies.
[b] Review [9] compared no-take MPAs with partially protected MPAs, but <10 primary studies were available for changes in biomass, species richness, and organism size.
[c] Review [11] did not investigate effects of MPA size and age on biomass due to low number of primary studies.
[d] Review [13] included MPA size as a potential confounding factor, but did not directly investigate the effects of reserve size on abundance.
[e] Review [17] was limited in investigation of effects of MPA size and age on species richness and on organism size due to low number of relevant primary studies.
[f] Review [18] was unable to investigate the effects of MPA size and age on species richness due to the low number of relevant primary studies.

Table A10: Scope of meta-analyses that examine the effectiveness of marine protected areas. Combinations of Region and Outcome Measures are shown. Each number corresponds to a review – see Table A2 for details.

		OUTCOME MEASURE			
		Abundance	Biomass	Species Richness	Size
REGION	Global	3, 6-9, 11-14, 16, 17, [b]	6-8, 17 , [b, c, d]	3, 6-8, 17, [b, c]	6-8, 17, [b, c]
	Temperate	1, 2, 4, 5, 8, 15, 18, [a, e]	4, 5, 8, 18	1, 4, 8, [f]	4, 8, 15
	Tropical	8, 10 [a, e]	8	8	8

Notes: [a] Review [3] examined effect of latitude on density, but did not report separate effect sizes for temperate and tropical regions.
[b] Review [4] includes data from a previous global-scale synthesis (Lester et al. 2009, see main text for full reference).
[c] Review [9] examined effects of MPAs on biomass, species richness, and organism size, but <10 relevant primary studies were found.
[d] Review [11] did not investigate effects of MPAs on biomass due to low number of relevant primary studies.
[e] Review [13] included latitude as potential confounding factor but did not directly investigate the effect sizes for tropical and temperate regions.
[f] Review [18] investigated effects of MPAs on species richness, but with <10 primary studies.

Table A11: Scope of narrative syntheses that examine the effectiveness of marine protected areas based on Broad Focus, Region, Type of Protected Area; MPA Characteristic. Each letter corresponds to a review – see Table A2 for details.

	SCOPE	REVIEW
BROAD FOCUS	Biodiversity	a, b, c, d, f, h, i
	Fisheries	b, d, e, f, g, h, i
REGION	Global	b, c, e, f, g, h, i
	Tropical [a]	a
	Temperate [a]	d
TYPE OF PROTECTED AREA	No-take	a, c
	MPA	b, d, e, f, g, h, i
MPA CHARACTERISTIC	Size [e, f]	a, e, g, i
	Connectivity	a

Table A12: Scope of narrative syntheses that examine the effectiveness of marine protected areas: Focus x Region. Each letter corresponds to a review – see Table A2 for details.

		FOCUS	
		Biodiversity	Fisheries
REGION	Global	b, c, f, h, i	b, e, f, g, h, i
	Temperate	d	d
	Tropical	a	

Table A13: Scope of narrative syntheses that examine the effectiveness of marine protected areas: Type of Protected Area x Region

		TYPE OF PROTECTED AREA	
		No-take	MPA
REGION	Global	c	b, e, f, g, h, i
	Temperate		d
	Tropical	a	

Table A14: Scope of narrative syntheses that examine the effectiveness of marine protected areas: MPA Characteristic x Region

		MPA CHARACTERISTIC	
		Size	Connectivity
REGION	Global	e, g, i	
	Temperate		
	Tropical	a	a

Table A15: Scope of narrative syntheses that examine the effectiveness of marine protected areas: MPA Characteristic x Focus

		MPA CHARACTERISTIC		
		Size	Connectivity	Spacing
FOCUS	Biodiversity	a	a	a
	Fisheries	e, g, i		

Table A16: Scope of narrative syntheses that examine the effectiveness of marine protected areas: MPA Characteristic x Type of Protected Area

		MPA CHARACTERISTIC		
		Size	Connectivity	Spacing
TYPE OF PROTECTED AREA	No-take	a	a	a
	MPA	e, g, i		

Table A17: Scope of narrative syntheses that examine the effectiveness of marine protected areas: Type of Protected Area x Focus

		FOCUS	
		Biodiversity	Fisheries
TYPE OF PROTECTED AREA	No-take	a, c	
	MPA	b, d, f, h, i	b, d, e, f, g, h, i