

# Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: a PRISMA extension

Rose E. O'Dea<sup>1\*</sup>, Malgorzata Lagisz<sup>1</sup>, Michael D. Jennions<sup>2</sup>, Julia Koricheva<sup>3</sup>, Daniel W.A. Noble<sup>1,2</sup>, Timothy H. Parker<sup>4</sup>, Jessica Gurevitch<sup>5</sup>, Matthew J. Page<sup>6</sup>, Gavin Stewart<sup>7</sup>, David Moher<sup>8</sup> and Shinichi Nakagawa<sup>1\*</sup>

<sup>1</sup>*Evolution & Ecology Research Centre and School of Biological and Environmental Sciences, University of New South Wales, Sydney, NSW, 2052, Australia*

<sup>2</sup>*Research School of Biology, Australian National University, 46 Sullivans Creek Road, Canberra, 2600, Australia*

<sup>3</sup>*Department of Biological Sciences, Royal Holloway University of London, Egham, Surrey, TW20 0EX, U.K.*

<sup>4</sup>*Department of Biology, Whitman College, Walla Walla, WA, 99362, U.S.A.*

<sup>5</sup>*Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY, 11794-5245, U.S.A.*

<sup>6</sup>*School of Public Health and Preventative Medicine, Monash University, Melbourne, VIC, 3004, Australia*

<sup>7</sup>*School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, NE1 7RU, U.K.*

<sup>8</sup>*Centre for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, General Campus, 501 Smyth Road, Room L1288, Ottawa, ON, K1H 8L6, Canada*

## ABSTRACT

Since the early 1990s, ecologists and evolutionary biologists have aggregated primary research using meta-analytic methods to understand ecological and evolutionary phenomena. Meta-analyses can resolve long-standing disputes, dispel spurious claims, and generate new research questions. At their worst, however, meta-analysis publications are wolves in sheep's clothing: subjective with biased conclusions, hidden under coats of objective authority. Conclusions can be rendered unreliable by inappropriate statistical methods, problems with the methods used to select primary research, or problems within the primary research itself. Because of these risks, meta-analyses are increasingly conducted as part of systematic reviews, which use structured, transparent, and reproducible methods to collate and summarise evidence. For readers to determine whether the conclusions from a systematic review or meta-analysis should be trusted – and to be able to build upon the review – authors need to report what they did, why they did it, and what they found. Complete, transparent, and reproducible reporting is measured by 'reporting quality'. To assess perceptions and standards of reporting quality of systematic reviews and meta-analyses published in ecology and evolutionary biology, we surveyed 208 researchers with relevant experience (as authors, reviewers, or editors), and conducted detailed evaluations of 102 systematic review and meta-analysis papers published between 2010 and 2019. Reporting quality was far below optimal and approximately normally distributed. Measured reporting quality was lower than what the community perceived, particularly for the systematic review methods required to measure trustworthiness. The minority of assessed papers that referenced a guideline (~16%) showed substantially higher reporting quality than average, and surveyed researchers showed interest in using a reporting guideline to improve reporting quality. The leading guideline for improving reporting quality of systematic reviews is the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement. Here we unveil an extension of PRISMA to serve the meta-analysis community in ecology and evolutionary biology: PRISMA-EcoEvo (version 1.0). PRISMA-EcoEvo is a checklist of 27 main items that, when applicable, should be reported in systematic review and meta-analysis publications summarising primary research in ecology and evolutionary biology. In this explanation and elaboration document, we provide guidance for authors, reviewers, and editors, with explanations for each item on the checklist, including supplementary examples from published papers. Authors can

\* Authors for correspondence: (E-mail: [rose.eleanor.o.dea@gmail.com](mailto:rose.eleanor.o.dea@gmail.com)); (Tel.: +61 2 9385 9138; E-mail: [s.nakagawa@unsw.edu.au](mailto:s.nakagawa@unsw.edu.au))

consult this PRISMA-EcoEvo guideline both in the planning and writing stages of a systematic review and meta-analysis, to increase reporting quality of submitted manuscripts. Reviewers and editors can use the checklist to assess reporting quality in the manuscripts they review. Overall, PRISMA-EcoEvo is a resource for the ecology and evolutionary biology community to facilitate transparent and comprehensively reported systematic reviews and meta-analyses.

**Key words:** comparative analysis, critical appraisal, evidence synthesis, non-independence, open science, study quality, pre-registration, registration

## CONTENTS

|   |      |
|---|------|
| I. Introduction   | 1696 |
| II. Primer a: terminology   | 1697 |
| III. Abstract & introduction  | 1698 |
| IV. Primer b: types of questions  | 1702 |
| V. Registration   | 1703 |
| VI. Primer c: registration and registered reports                         | 1703 |
| VII. Finding and extracting information                                   | 1704 |
| VIII. Primer d: bias from primary studies                                 | 1706 |
| IX. Analysis methods  | 1707 |
| X. Primer e: phylogenetic non-independence                                | 1711 |
| XI. Reporting results   | 1711 |
| XII. Discussion and contributions   | 1714 |
| XIII. How and why to use prisma-ecoevo                                    | 1715 |
| (1) PRISMA-EcoEvo for authors   | 1715 |
| (2) PRISMA-EcoEvo for reviewers and editors                               | 1715 |
| (3) Valuing permanent online repositories                                 | 1716 |
| (4) Comparisons with PRISMA-2020  | 1716 |
| XIV. Conclusions  | 1716 |
| XV. Acknowledgements, author contributions and data and code availability | 1717 |
| XVI. References   | 1717 |
| XVII. Supporting information  | 1722 |

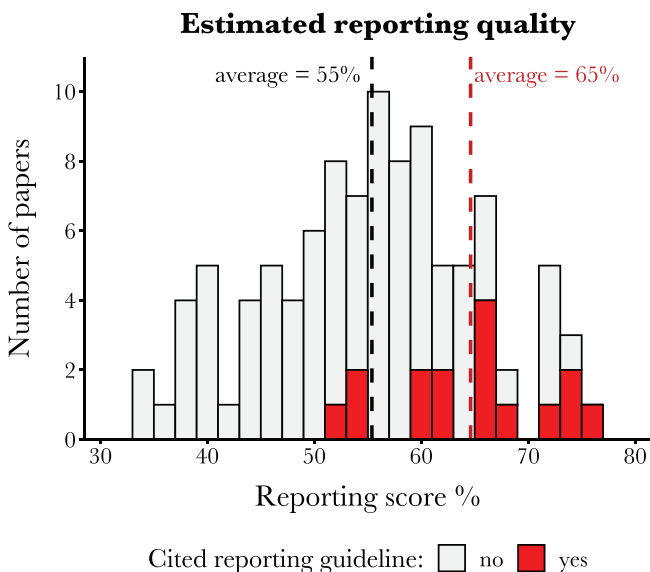
## I. INTRODUCTION

Ecological and evolutionary research topics are often distilled in systematic review and meta-analysis publications (Gurevitch *et al.*, 2018; Koricheva & Kulinskaya, 2019). Although terminology differs both across and within disciplines, here we use the term ‘meta-analysis’ to refer to the statistical synthesis of effect sizes from multiple independent studies, whereas a ‘systematic review’ is the outcome of a series of established, transparent, and reproducible methods to find and summarise studies (definitions are discussed further in Primer A below). As with any scientific project, systematic reviews and meta-analyses are susceptible to quality issues, limitations, and biases that can undermine the credibility of their conclusions. First, the strength of primary evidence included in the review might be weakened by selective reporting and research biases (Jennions & Möller, 2002; Forstmeier, Wagenmakers & Parker, 2017; Fraser *et al.*, 2018). Second, reviews might be conducted or communicated in ways that summarise existing evidence inaccurately (Whittaker, 2010; Ioannidis, 2016). Systematic review methods have been designed to identify and mitigate both these threats to credibility (Haddaway & Macura, 2018)

but, from the details that authors of meta-analyses report, it is often unclear whether systematic review methods have been used in ecology and evolution. For a review to provide a firm base of knowledge on which researchers can build, it is essential that review authors transparently report their aims, methods, and outcomes (Liberati *et al.*, 2009; Parker *et al.*, 2016a).

In evidence-based medicine, where biased conclusions from systematic reviews can endanger human lives, transparent reporting is promoted by reporting guidelines and checklists such as the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement. PRISMA, first published in 2009 (Moher *et al.*, 2009) and recently updated as PRISMA-2020 (Page *et al.*, 2021b), describes minimum reporting standards for authors of systematic reviews of healthcare interventions. PRISMA has been widely cited and endorsed by prominent journals, and there is evidence of improved reporting quality in clinical research reviews following its publication (Page & Moher, 2017). Several extensions of PRISMA have been published to suit different types of reviews (e.g. PRISMA for Protocols, PRISMA for Network Meta-Analyses, and PRISMA for individual patient data: Hutton *et al.*, 2015; Moher *et al.*, 2015; Stewart *et al.*, 2015).

Ecologists and evolutionary biologists seldom reference reporting guidelines in systematic reviews and meta-analyses. However, there is community support for wider use of reporting guidelines (based on our survey of 208 researchers; see online Supporting Information) and benefits to their adoption. In a representative sample of 102 systematic review and meta-analysis papers published between 2010 and 2019, the 16% of papers that mentioned a reporting guideline showed above-average reporting quality (Fig. 1). In all but one paper, the reporting guideline used by authors was PRISMA, despite it being focussed on reviews of clinical research. While more discipline-appropriate reporting checklists are available for our fields (e.g. 'ROSES RepOrting standards for Systematic Evidence Syntheses'; Haddaway *et al.*, 2018; and the Tools for Transparency in Ecology and Evolution'; Parker *et al.*, 2016a), these have so far focussed on applied topics in environmental evidence, and/or lack explanations and examples for meta-analysis reporting items. Ecologists and evolutionary biologists need a detailed reporting guideline for systematic review and meta-analysis papers.



**Fig 1.** Results from our assessment of reporting quality of systematic reviews and meta-analyses published between 2010 and 2019, in ecology and evolutionary biology ( $n = 102$ ). For each paper, the reporting score represents the mean 'average item % score' across all applicable items. Full details are provided in the Supporting Information and supplementary code. Red columns indicate the minority of papers that cited a reporting guideline ( $n = 15$  cited PRISMA, and  $n = 1$  cited Koricheva & Gurevitch, 2014). The subset of papers that referenced a reporting guideline tended to have higher reporting scores (note that these observational data cannot distinguish between checklists causing better reporting, or authors with better reporting practices being more likely to report using checklists). Welch's  $t$ -test:  $t$ -value = 5.21;  $df = 25.65$ ;  $P < 0.001$ .

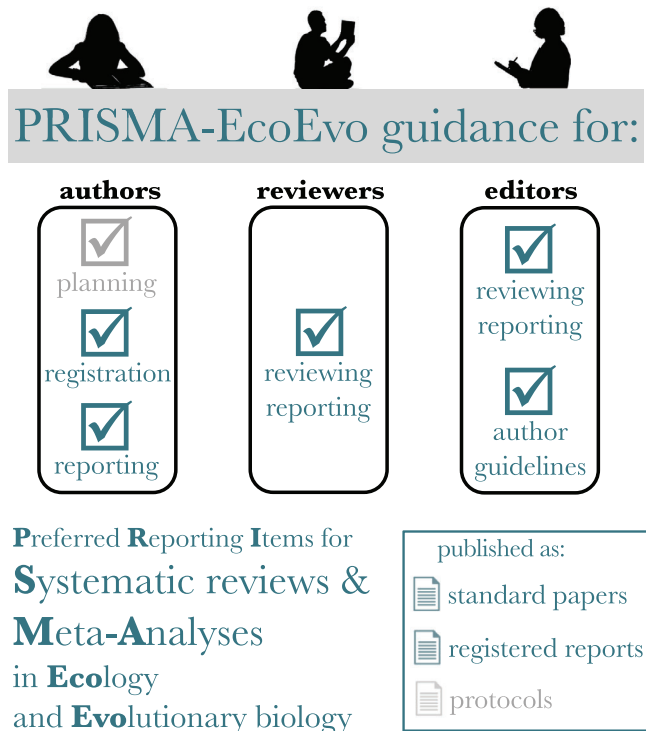
We have designed version 1.0 of a PRISMA extension for ecology and evolutionary biology: PRISMA-EcoEvo. This guideline caters for the types of reviews and methods common within our fields. For example, meta-analyses in ecology and evolutionary biology often combine large numbers of diverse studies to summarise patterns across multiple taxa and/or environmental conditions (Nakagawa & Santos, 2012; Senior *et al.*, 2016). Aggregating diverse studies often creates multiple types of statistical non-independence that require careful consideration (Noble *et al.*, 2017), and guidance on reporting these statistical issues is not comprehensively covered by PRISMA. Conversely, some of the items on PRISMA are yet to be normalised within ecology and evolution (e.g. risk of bias assessment, and duplicate data extraction). Without pragmatic consideration of these differences between fields, most ecologists and evolutionary biologists are unlikely to use a reporting guideline for systematic reviews and meta-analyses.

Here we explain every item of the PRISMA-EcoEvo checklist, for use by authors, peer-reviewers, and editors (Fig. 2). We also include extended discussion of the more difficult topics for authors in five 'Primer' sections (labelled A–E). Table 1 presents a checklist of sub-items, to aid the assessment of partial reporting. The full checklist applies to systematic reviews with a meta-analysis, but many of the items will be applicable to systematic reviews without a meta-analysis, and meta-analyses without a systematic review. Examples of each item from a published paper are presented in the Supporting Information, alongside text descriptions of current reporting practices.

## II. PRIMER A: TERMINOLOGY

Within the ecology and evolutionary biology community there are terminological differences regarding how 'meta-analysis' is defined (Vetter, Rucker & Storch, 2013). In the broadest sense, any aggregation of results from multiple studies is sometimes referred to as a 'meta-analysis' (including the common but inadvisable practice of tallying the number of significant *versus* non-significant results, i.e. 'vote-counting'; Vetter *et al.*, 2013; Koricheva & Gurevitch, 2014; Gurevitch *et al.*, 2018). Here, we reserve the term 'meta-analysis' for studies in which effect sizes from multiple independent studies are combined in a statistical model, to give an estimate of a pooled effect size and error. Each effect size represents a result, and the effect sizes from multiple studies are expressed on the same scale. Usually the effect sizes are weighted so that more precise estimates (lower sampling error) have a greater impact on the pooled effect size than imprecise estimates (although unweighted analyses can sometimes be justified; see Item 12).

In comparison with meta-analyses, which have been used in ecology and evolutionary biology for nearly 30 years (the first meta-analysis in ecology was published by Jarvinen, 1991), systematic reviews are only now becoming



**Fig 2.** PRISMA-EcoEvo for authors, peer-reviewers, and editors. Planning and protocols are shown in grey because, while PRISMA-EcoEvo can point authors in the right direction, authors should seek additional resources for detailed conduct guidance. Authors can use PRISMA-EcoEvo as a reporting guideline for both registered reports (Primer C) and completed manuscripts. Reviewers and editors can use PRISMA-EcoEvo to assess reporting quality of the systematic review and meta-analysis manuscripts they read. Editors can promote high reporting quality by asking submitting authors to complete the PRISMA-EcoEvo checklist, either by downloading a static file at <https://osf.io/t8qd2/>, or by using an interactive web application at <https://prisma-ecoevo.shinyapps.io/checklist/>.

an established method (Gurevitch *et al.*, 2018; Berger-Tal *et al.*, 2019; but see Pullin & Stewart, 2006). Systematic-review methods are concerned with how information was gathered and synthesised. In fields such as medicine and conservation biology, the required steps for a systematic review are as follows: defining specific review questions; identifying all likely relevant records; screening studies against pre-defined eligibility criteria; assessing the risk of bias both within and across studies (i.e. 'critical appraisal'; Primer D); extracting data; and synthesising results (which might include a meta-analysis) (Pullin & Stewart, 2006; Liberati *et al.*, 2009; Haddaway & Verhoeven, 2015; James, Randall & Haddaway, 2016; Cooke *et al.*, 2017; Higgins *et al.*, 2019). Under this formal definition, systematic reviews in ecology and evolutionary biology are exceedingly rare for two reasons. First, we tend not to conduct exhaustive searches to find all relevant records (e.g. we usually rely on a sample of

published sources from just one or two databases). Second, assessing the risk of bias in primary studies is very uncommon (based on the meta-analyses we assessed; see Supporting Information and Section VIII). Given current best practice and usage of the term 'systematic review' in ecology and evolutionary biology, the PRISMA-EcoEvo checklist is targeted towards meta-analyses that were conducted on data collected from multiple sources and whose methods were structured, transparent, and reproducible.

### III. ABSTRACT & INTRODUCTION

#### Item 1: Title and abstract

**In the title or abstract, identify the review as a systematic review, meta-analysis, or both. In the abstract provide a summary including the aims and scope of the review, description of the data set, results of the primary outcome, conclusions, and limitations.**

##### Explanation and elaboration

Identifying the report as a systematic review and/or meta-analysis in the title, abstract, or keywords makes these types of reviews identifiable through database searches. It is essential that the summary of the review in the abstract is accurate because this is the only part of the review that some people will read (either by choice or because of restricted access) (Beller *et al.*, 2013). While it is currently rare for abstracts to report limitations, this practice should change. Casual readers can be misled if the abstract does not disclose limitations of the review or fails to report the result of the primary outcome. Even very concise abstracts (e.g. for journals that allow a maximum of 150 words) should state obvious limitations. To signify greater accountability, authors can report when their review was registered in advance (Item 3).

#### Item 2: Aims and questions

**Explain the rationale of the study, including reference to any previous reviews or meta-analyses on the topic. State the aims and scope of the study (including its generality) and its primary questions (e.g. which moderators were tested), including whether effect sizes are derived from experimental and/or observational comparisons.**

##### Explanation and elaboration

An effective introduction sets up the reader so that they understand why a study was done and what it entailed, making it easier to process subsequent information (Liberati *et al.*, 2009). In this respect, the introduction of a systematic review is no different to that of primary studies (Heard, 2016). Previous review articles are likely to influence thinking around a research topic, so these reviews should be placed in context, as their absence signifies a research gap. If the introduction is well written and the study is well designed, then the reader can roughly infer what methods were used before reading the methods section. To achieve such harmony, authors should clearly lay out the scope and primary

**Table 1.** PRISMA-EcoEvo v1.0. Checklist of preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology, alongside an assessment of recent reporting practices (based on a representative sample of 102 meta-analyses published between 2010 and 2019; references to all assessed papers are provided in the reference list, while the Supporting Information presents details of the assessment). The proportion of papers meeting each sub-item is presented as a percentage. While all papers were assessed for each item, there was a set of reasons why some items might not be applicable (e.g. no previous reviews on the topic would make sub-item 2.2 not applicable). Only applicable sub-items contributed to reporting scores; sample sizes for each sub-item are shown in the column on the right. Asterisks (\*) indicate sub-items that are identical, or very close, to items from the 2009 PRISMA checklist. In the wording of each Sub-item, 'review' encompasses all forms of evidence syntheses (including systematic reviews), while 'meta-analysis' and 'meta-regression' refer to statistical methods for analysing data collected in the review (definitions are discussed further in Primer A)

| Checklist item       | Sub-item number | Sub-item   | Papers meeting component (%) | No. papers applicable |
|----------------------|-----------------|--|------------------------------|-----------------------|
| Title and abstract   | 1.1             | Identify the review as a systematic review, meta-analysis, or both*  | 100                          | 102                   |
|                      | 1.2             | Summarise the aims and scope of the review   | 97                           | 102                   |
|                      | 1.3             | Describe the data set  | 74                           | 102                   |
|                      | 1.4             | State the results of the primary outcome   | 96                           | 102                   |
|                      | 1.5             | State conclusions*   | 94                           | 102                   |
|                      | 1.6             | State limitations*   | 17                           | 96                    |
| Aims and questions   | 2.1             | Provide a rationale for the review*  | 100                          | 102                   |
|                      | 2.2             | Reference any previous reviews or meta-analyses on the topic   | 93                           | 75                    |
|                      | 2.3             | State the aims and scope of the review (including its generality)  | 91                           | 102                   |
|                      | 2.4             | State the primary questions the review addresses (e.g. which moderators were tested)   | 96                           | 102                   |
|                      | 2.5             | Describe whether effect sizes were derived from experimental and/or observational comparisons  | 57                           | 76                    |
| Review registration  | 3.1             | Register review aims, hypotheses (if applicable), and methods in a time-stamped and publicly accessible archive and provide a link to the registration in the methods section of the manuscript. Ideally registration occurs before the search, but it can be done at any stage before data analysis | 3                            | 102                   |
|                      | 3.2             | Describe deviations from the registered aims and methods   | 0                            | 3                     |
|                      | 3.3             | Justify deviations from the registered aims and methods  | 0                            | 3                     |
| Eligibility criteria | 4.1             | Report the specific criteria used for including or excluding studies when screening titles and/or abstracts, and full texts, according to the aims of the systematic review (e.g. study design, taxa, data availability)   | 84                           | 102                   |
|                      | 4.2             | Justify criteria, if necessary (i.e. not obvious from aims and scope)  | 54                           | 67                    |
| Finding studies      | 5.1             | Define the type of search (e.g. comprehensive search, representative sample)   | 25                           | 102                   |
|                      | 5.2             | State what sources of information were sought (e.g. published and unpublished studies, personal communications)*   | 89                           | 102                   |
|                      | 5.3             | Include, for each database searched, the exact search strings used, with keyword combinations and Boolean operators  | 49                           | 102                   |
| Study selection      | 5.4             | Provide enough information to repeat the equivalent search (if possible), including the timespan covered (start and end dates)   | 14                           | 102                   |
|                      | 6.1             | Describe how studies were selected for inclusion at each stage of the screening process (e.g. use of decision trees, screening software)   | 13                           | 102                   |
|                      | 6.2             | Report the number of people involved and how they contributed (e.g. independent parallel screening)  | 3                            | 102                   |



Table 1. (Cont.)

| Checklist item                         | Sub-item number | Sub-item   | Papers meeting component (%) | No. papers applicable |
|--|-----------------|--|------------------------------|-----------------------|
| Data collection process                | 7.1             | Describe where in the reports data were collected from (e.g. text or figures)  | 44                           | 102                   |
|  | 7.2             | Describe how data were collected (e.g. software used to digitize figures, external data sources)   | 42                           | 102                   |
|  | 7.3             | Describe moderator variables that were constructed from collected data (e.g. number of generations calculated from years and average generation time)  | 56                           | 41                    |
|  | 7.4             | Report how missing or ambiguous information was dealt with during data collection (e.g. authors of original studies were contacted for missing descriptive statistics, and/or effect sizes were calculated from test statistics)                                 | 47                           | 102                   |
| Data items                             | 7.5             | Report who collected data  | 10                           | 102                   |
|  | 7.6             | State the number of extractions that were checked for accuracy by co-authors   | 1                            | 102                   |
|  | 8.1             | Describe the key data sought from each study   | 96                           | 102                   |
|  | 8.2             | Describe items that do not appear in the main results, or which could not be extracted due to insufficient information   | 42                           | 53                    |
| Assessment of individual study quality | 8.3             | Describe main assumptions or simplifications that were made (e.g. categorising both 'length' and 'mass' as 'morphology')   | 62                           | 86                    |
|  | 8.4             | Describe the type of replication unit (e.g. individuals, broods, study sites)  | 73                           | 102                   |
|  | 9.1             | Describe whether the quality of studies included in the systematic review or meta-analysis was assessed (e.g. blinded data collection, reporting quality, experimental <i>versus</i> observational)  | 7                            | 102                   |
|  | 9.2             | Describe how information about study quality was incorporated into analyses (e.g. meta-regression and/or sensitivity analysis)   | 6                            | 102                   |
| Effect size measures                   | 10.1            | Describe effect size(s) used   | 97                           | 102                   |
|  | 10.2            | Provide a reference to the equation of each calculated effect size (e.g. standardised mean difference, log response ratio) and (if applicable) its sampling variance   | 63                           | 91                    |
|  | 10.3            | If no reference exists, derive the equations for each effect size and state the assumed sampling distribution(s)   | 7                            | 28                    |
| Missing data                           | 11.1            | Describe any steps taken to deal with missing data during analysis (e.g. imputation, complete case, subset analysis)   | 37                           | 57                    |
|  | 11.2            | Justify the decisions made to deal with missing data   | 21                           | 57                    |
| Meta-analytic model description        | 12.1            | Describe the models used for synthesis of effect sizes   | 97                           | 102                   |
|  | 12.2            | The most common approach in ecology and evolution will be a random-effects model, often with a hierarchical/multilevel structure. If other types of models are chosen (e.g. common/fixed effects model, unweighted model), provide justification for this choice | 50                           | 40                    |
| Software                               | 13.1            | Describe the statistical platform used for inference (e.g. <i>R</i> )  | 92                           | 102                   |
|  | 13.2            | Describe the packages used to run models   | 74                           | 80                    |
|  | 13.3            | Describe the functions used to run models  | 22                           | 69                    |
|  | 13.4            | Describe any arguments that differed from the default settings   | 29                           | 75                    |
|  | 13.5            | Describe the version numbers of all software used  | 33                           | 102                   |
| Non-independence                       | 14.1            | Describe the types of non-independence encountered (e.g. phylogenetic, spatial, multiple measurements over time)   | 32                           | 102                   |
|  | 14.2            | Describe how non-independence has been handled   | 74                           | 102                   |
|  | 14.3            | Justify decisions made   | 47                           | 102                   |

(Continues)

Table 1. (Cont.)

| Checklist item  | Sub-item number | Sub-item  | Papers meeting component (%) | No. papers applicable |
|---|-----------------|---|------------------------------|-----------------------|
| Meta-regression and model selection                   | 15.1            | Provide a rationale for the inclusion of moderators (covariates) that were evaluated in meta-regression models  | 81                           | 94                    |
|   | 15.2            | Justify the number of parameters estimated in models, in relation to the number of effect sizes and studies (e.g. interaction terms were not included due to insufficient sample sizes)   | 20                           | 94                    |
|   | 15.3            | Describe any process of model selection   | 80                           | 40                    |
| Publication bias and sensitivity analyses             | 16.1            | Describe assessments of the risk of bias due to missing results (e.g. publication, time-lag, and taxonomic biases)  | 65                           | 102                   |
|   | 16.2            | Describe any steps taken to investigate the effects of such biases (if present)   | 47                           | 30                    |
|   | 16.3            | Describe any other analyses of robustness of the results, e.g. due to effect size choice, weighting or analytical model assumptions, inclusion or exclusion of subsets of the data, or the inclusion of alternative moderator variables in meta-regressions | 35                           | 102                   |
| Clarification of <i>post hoc</i> analyses             | 17.1            | When hypotheses were formulated after data analysis, this should be acknowledged  | 14                           | 28                    |
| Metadata, data, and code                              | 18.1            | Share metadata (i.e. data descriptions)   | 44                           | 102                   |
|   | 18.2            | Share data required to reproduce the results presented in the manuscript  | 77                           | 102                   |
|   | 18.3            | Share additional data, including information that was not presented in the manuscript (e.g. raw data used to calculate effect sizes, descriptions of where data were located in papers)   | 39                           | 102                   |
| Results of study selection process                    | 18.4            | Share analysis scripts (or, if a software package with graphical user interface (GUI) was used, then describe full model specification and fully specify choices)   | 11                           | 102                   |
|   | 19.1            | Report the number of studies screened*  | 37                           | 102                   |
|   | 19.2            | Report the number of studies excluded at each stage of screening  | 22                           | 102                   |
|   | 19.3            | Report brief reasons for exclusion from the full-text stage   | 27                           | 102                   |
| Sample sizes and study characteristics                | 19.4            | Present a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)-like flowchart (www.prisma-statement.org)*  | 19                           | 102                   |
|   | 20.1            | Report the number of studies and effect sizes for data included in meta-analyses  | 96                           | 91                    |
|   | 20.2            | Report the number of studies and effect sizes for subsets of data included in meta-regressions  | 57                           | 93                    |
|   | 20.3            | Provide a summary of key characteristics for reported outcomes (either in text or figures; e.g. one quarter of effect sizes reported for vertebrates and the rest invertebrates)  | 62                           | 102                   |
| Meta-analysis   | 20.4            | Provide a summary of limitations of included moderators (e.g. collinearity and overlap between moderators)  | 22                           | 87                    |
|   | 20.5            | Provide a summary of characteristics related to individual study quality (risk of bias)   | 60                           | 5                     |
|   | 21.1            | Provide a quantitative synthesis of results across studies, including estimates for the mean effect size, with confidence/credible intervals  | 94                           | 87                    |
| Heterogeneity   | 22.1            | Report indicators of heterogeneity in the estimated effect (e.g. $I^2$ , $\tau^2$ and other variance components)  | 52                           | 84                    |
|   | 23.1            | Provide estimates of meta-regression slopes (i.e. regression coefficients) and confidence/credible intervals  | 78                           | 94                    |
| Outcomes of publication bias and sensitivity analyses | 23.2            | Include estimates and confidence/credible intervals for all moderator variables that were assessed (i.e. complete reporting)  | 59                           | 94                    |
|   | 23.3            | Report interactions, if they were included  | 59                           | 27                    |
|   | 23.4            | Describe outcomes from model selection, if done (e.g. $R^2$ and AIC)  | 81                           | 36                    |
|   | 24.1            | Provide results for the assessments of the risks of bias (e.g. Egger's regression, funnel plots)  | 60                           | 102                   |
|   | 24.2            | Provide results for the robustness of the review's results (e.g. subgroup analyses, meta-regression of study quality, results from alternative methods of analysis, and temporal trends)  | 44                           | 102                   |

Table 1. (Cont.)

| Checklist item            | Sub-item number | Sub-item  | Papers meeting component (%) | No. papers applicable |
|---------------------------|-----------------|---|------------------------------|-----------------------|
| Discussion                | 25.1            | Summarise the main findings in terms of the magnitude of effect   | 73                           | 102                   |
|                           | 25.2            | Summarise the main findings in terms of the precision of effects (e.g. size of confidence intervals, statistical significance)                                      | 57                           | 102                   |
|                           | 25.3            | Summarise the main findings in terms of their heterogeneity   | 47                           | 89                    |
|                           | 25.4            | Summarise the main findings in terms of their biological/practical relevance  | 98                           | 102                   |
|                           | 25.5            | Compare results with previous reviews on the topic, if available  | 88                           | 72                    |
|                           | 25.6            | Consider limitations and their influence on the generality of conclusions, such as gaps in the available evidence (e.g. taxonomic and geographical research biases) | 72                           | 100                   |
| Contributions and funding | 26.1            | Provide names, affiliations, and funding sources of all co-authors  | 92                           | 102                   |
|                           | 26.2            | List the contributions of each co-author  | 31                           | 102                   |
|                           | 26.3            | Provide contact details for the corresponding author  | 100                          | 102                   |
|                           | 26.4            | Disclose any conflicts of interest  | 0                            | 8                     |
| References                | 27.1            | Provide a reference list of all studies included in the systematic review or meta-analysis  | 92                           | 102                   |
|                           | 27.2            | List included studies as referenced sources (e.g. rather than listing them in a table or supplement)  | 18                           | 102                   |

aims of the study (e.g. which taxa and types of studies; hypothesis testing or generating/exploratory). The scope is crucial because this influences many aspects of the study design (e.g. eligibility criteria; Item 4) and interpretation of results. It is also important to distinguish between experimental and observational studies, as experiments provide an easier path to causal conclusions.

#### IV. PRIMER B: TYPES OF QUESTIONS

Broadly, systematic reviews with meta-analyses can answer questions of two kinds: the generality of a phenomenon, or its overall effect (Gurevitch *et al.*, 2018). When PRISMA was published in 2009 for reviews assessing the overall effect (i.e. the meta-analytic intercept/mean) it recommended that questions be stated with reference to ‘PICO’ or ‘PECO’: population (e.g. adult humans at risk of cardiovascular disease), intervention or exposure (e.g. statin medication), comparator (e.g. control group taking a placebo), and outcome (e.g. difference in the number of cardiovascular disease events between the intervention and control groups) (Liberati *et al.*, 2009). When a population is limited to one species, or one subset of a population, the number of studies available to quantify the overall effect of an intervention or exposure is typically small (e.g. <20 studies) (Gurevitch *et al.*, 2018). However, even when ecological and evolutionary questions can be framed in terms of ‘PICO’ or ‘PECO’, the ‘population’ is often broad (e.g. vertebrates, whole ecosystems) leading to larger and more diverse data sets (Gerstner *et al.*, 2017). Examples include the effect of the latitudinal gradient on global species richness (Kinlock *et al.*, 2018;  $n = 199$  studies), the effect of parasite infection on body condition in wild species (Sánchez *et al.*, 2018;  $n = 187$  studies), the effect of livestock grazing on ecosystem properties of salt marshes (Davidson *et al.*, 2017;  $n = 89$  studies), and the effect of cytoplasmic genetic variation on phenotypic variation in eukaryotes (Dobler *et al.*, 2014;  $n = 66$  studies).

In ecological and evolutionary meta-analyses, determining the average overall effect across studies is usually of less interest than exploring the extent, and sources, of variability in effect sizes among studies. Combining a large number of studies across species or contexts increases variability and makes estimation and interpretation of the average effect difficult and arguably meaningless. Instead, exploring variables which influence the magnitude or direction of an effect can be particularly fruitful; these variables could be biological or methodological (Gerstner *et al.*, 2017). To explore sources of variation in effects, and quantify statistical power, it is important to report the magnitude of heterogeneity among effect sizes (i.e. differences in effect sizes between studies beyond what is expected from sampling error; Item 22). Heterogeneity is typically high when diverse studies are combined in a single meta-analysis. High heterogeneity is considered problematic in medical meta-analyses (Liberati



*et al.*, 2009; Muka *et al.*, 2020), but it is the norm in ecology and evolutionary biology (Stewart, 2010; Senior *et al.*, 2016), and identifying sources of heterogeneity could produce both biological and methodological insights (Rosenberg, 2013). Authors of meta-analyses in ecology and evolutionary biology should be aware that high heterogeneity reduces statistical power for a given sample size, especially for estimates of moderator effects and their interactions (Valentine, Pigott & Rothstein, 2010). It is therefore important to report estimates of heterogeneity (Item 22) alongside full descriptions of sample sizes (Item 20), and communicate appropriate uncertainty in analysis results.

## V. REGISTRATION

### Item 3: Review registration

**Register study aims, hypotheses (if applicable), and methods in a time-stamped and publicly accessible archive. Ideally registration occurs before the search, but it can be done at any stage before data analysis. A link to the archived registration should be provided in the methods section of the manuscript. Describe and justify deviations from the registered aims and methods.**

#### Explanation and elaboration

Registering planned research and analyses, in a time-stamped and publicly accessible archive, is easily achieved with existing infrastructure (Nosek *et al.*, 2018) and is a promising protection against false-positive findings (Allen & Mehler, 2019) (discussed in Primer C). While ecologists and evolutionary biologists have been slower to adopt registrations compared to researchers in the social and medical sciences, our survey found that authors who had tried registrations viewed their experience favourably (see Supporting Information). Given that authors of systematic reviews and meta-analyses often plan their methods in advance (e.g. to increase the reliability of study screening and data extractions), only a small behavioural change would be required to register these plans in a time-stamped and publicly available archive. Inexperienced authors might feel ill-equipped to describe analysis plans in detail, but there are still benefits to registering conceptual plans (e.g. detailed aims, hypotheses, predictions, and variables that will be extracted for exploratory purposes only). Deviations from registered plans should be acknowledged and justified in the final report (e.g. when the data collected cannot be analysed using the proposed statistical model due to violation of assumptions). Authors who are comfortable with registration might consider publishing their planned systematic review or meta-analysis as a 'registered report', whereby the abstract, introduction and methods are submitted to a journal prior to the review being conducted. Some journals even publish review protocols before the review is undertaken (as is commonly done in environmental sciences, e.g. Greggor, Price & Shier, 2019).

## VI. PRIMER C: REGISTRATION AND REGISTERED REPORTS

Suboptimal reporting standards are often attributed to perverse incentives for career advancement (Ioannidis, 2005; Smaldino & McElreath, 2016; Moher *et al.*, 2020; Munafò *et al.*, 2020). 'Publish or perish' research cultures reward the frequent production of papers, especially papers that gain citations quickly. Researchers are therefore encouraged, both directly and indirectly, to extract more-compelling narratives from less-compelling data. For example, given multiple choices in statistical analyses, researchers might favour paths leading to statistical significance (i.e. 'P-hacking'; Simmons, Nelson & Simonsohn, 2011; Head *et al.*, 2015). Similarly, there are many ways to frame results in a manuscript. Results might be more impactful when framed as evidence for a hypothesis, even if data were not collected with the intention of testing that hypothesis (a problematic practice known as 'HARKing' — Hypothesising After the Results are Known; Kerr, 1998). Engaging in these behaviours does not require malicious intent or obvious dishonesty. Concerted effort is required to avoid the trap of self-deception (Forstmeier *et al.*, 2017; Aczel *et al.*, 2020). For researchers conducting a systematic review or meta-analysis, we need both to be aware that these practices could reduce the credibility of primary studies (Primer D), and guard against committing these practices when conducting and writing the review.

'Registration' or 'pre-registration' is an intervention intended to make it harder for researchers to oversell their results (Rice & Moher, 2019). Registration involves publicly archiving a written record of study aims, hypotheses, experimental or observational methods, and an analysis plan prior to conducting a study (Allen & Mehler, 2019). The widespread use of public archiving of study registrations only emerged in the 2000s when — in recognition of the harms caused by false-positive findings — the International Committee of Medical Journal Editors, World Medical Association, and the World Health Organisation, mandated that all medical trials should be registered (Goldacre, 2013). Since then, psychologists and other social scientists have adopted registrations too (which they term 'pre-registrations'; Rice & Moher, 2019), in response to high-profile cases of irreproducible research (Nelson, Simmons & Simonsohn, 2018; Nosek *et al.*, 2019). In addition to discouraging researchers from fishing for the most compelling stories in their data, registration may also help locate unpublished null results, which are typically published more slowly than 'positive' findings (Jennions & Møller, 2002) (i.e. registrations provide a window into researchers' 'file drawers', a goal of meta-analysts that seemed out of reach for decades; Rosenthal, 1979).

Beyond registration, a more powerful intervention is the 'registered report', because these not only make it harder for researchers to oversell their research and selectively report outcomes, but also prevent journals basing their publication decisions on study outcomes. In a registered report, the abstract, introduction, and method sections of a manuscript are submitted for peer review prior to conducting a study, and studies are provisionally accepted for publication before

their results are known (Parker, Fraser & Nakagawa, 2019). This publication style can therefore mitigate publication bias and helps to address flaws in researchers' questions and methods before it is too late to change them. Although the 'in principle' acceptance for publication does rely on authors closely following their registered plans, this fidelity comes with the considerable advantage of not requiring 'surprising' results for a smooth path to publication (and, if large changes reverse the initial decision of provisional acceptance, authors can still submit their manuscript as a new submission). Currently, a small number of journals that publish meta-analyses in ecology and evolutionary biology accept registered reports (see <https://cos.io/rr/> for an updated list) and, as with a regular manuscript, the PRISMA-EcoEvo checklist can be used to improve reporting quality in registered reports.

Systematic reviews and meta-analyses are well suited for registration and registered reports because these large and complicated projects have established and predictable methodology (Moher *et al.*, 2015; López-López *et al.*, 2018; Muka *et al.*, 2020). Despite these advantages, in ecology and evolutionary biology registration is rare (see Supporting Information). When we surveyed authors, reviewers, and editors, we found researchers had either not considered registration as an option for systematic reviews and meta-analyses or did not consider it worthwhile. Even in medical reviews, registration rates are lower than expected (Pussegoda *et al.*, 2017). Rather than a leap to perfect science, registration is a step towards greater transparency in the research process. Still, the practice has been criticised for not addressing underlying issues with research quality and external validity (Szollosi *et al.*, 2020). Illogical research questions and methods are not rescued by registration (Gelman, 2018), but registered reports provide the opportunity for them to be addressed before a study is conducted. Overall, wider adoption of registrations and registered reports is the clearest path towards transparent and reliable research.

## VII. FINDING AND EXTRACTING INFORMATION

### Item 4: Eligibility criteria

**Report the specific criteria used for including or excluding studies when screening titles and/or abstracts, and full texts, according to the aims of the meta-analysis (e.g. study design, taxa, data availability). Justify criteria, if necessary (i.e. not obvious from aims and scope).**

#### Explanation and elaboration

Fully disclosing which studies were included in the review allows readers to assess the generality, or specificity, of the review's conclusions (Vetter *et al.*, 2013). To decide upon the scope of the review, we typically use an iterative process of trial-and-error to refine the eligibility criteria, in conjunction with refining the research question. These planning stages should be conducted prior to registering study methods. Pragmatically, the scope of a systematic review

should be sufficiently broad to address the research question meaningfully, while being achievable within the authors' constrained resources (time and/or funding) (Forero *et al.*, 2019).

The eligibility criteria represent a key 'forking path' in any meta-analysis; slight modifications to the eligibility criteria could send the review down a path towards substantially different results (Palpacuer *et al.*, 2019). When planning a review, it is crucial to define explicit criteria for which studies will be included that are as objective as possible. These criteria need to be disclosed in the paper or supplementary information for the review to be replicable. It is especially important to describe criteria that do not logically follow from the aims and scope of the review (e.g. exclusion criteria chosen for convenience, such as excluding studies with missing data rather than contacting authors).

### Item 5: Finding studies

**Define the type of search (e.g. comprehensive search, representative sample), and state what sources of information were sought (e.g. published and unpublished studies, personal communications). For each database searched include the exact search strings used, with keyword combinations and Boolean operators. Provide enough information to repeat the equivalent search (if possible), including the timespan covered (start and end dates).**

#### Explanation and elaboration

Finding relevant studies to include in a systematic review is hard. Weeks can be spent sifting through massive piles of literature to find studies matching the eligibility criteria and yet, when reporting methods of the review, these details are typically skimmed over (average reporting quality <50%; see Supporting Information). While authors might deem it needlessly tedious to report the minutiae of their search methods, the supplementary information can service readers who wish to evaluate the appropriateness of the search methods (e.g. 'PRESS' – Peer Review of Electronic Search Strategies; McGowan *et al.*, 2016). Detailing search methods is also necessary for the study to be updatable using approximately the same methods (Garner *et al.*, 2016). Although journal subscriptions might vary over time and between different institutions (Mann, 2015), all authors can aim for approximately replicable searches. For instance, authors searching for studies through *Web of Science* should specify which databases were included in their search; institutions will typically only have access to a portion of the possible databases.

To recall how and why searches were conducted, authors should record the process and workflow of search strategy development. Often, multiple scoping searches are trialled before settling on a final search strategy (Siddaway, Wood & Hedges, 2019). For this process of trial-and-error, authors can check the ability of different searches to find a known set of suitable studies (studies that meet, or almost meet, the eligibility criteria; Item 4) (Bartels, 2013). The scoping

searches can be conducted using a single database, but it is preferable to use more than one database for the final search (Bramer *et al.*, 2018) (requiring duplicated studies to be removed prior to study selection, for which software is available; Rathbone *et al.*, 2015a; Westgate, 2019; Muka *et al.*, 2020). Sometimes potentially useful records will be initially inaccessible (e.g. when authors' home institutions do not subscribe to the journal), but efforts can be made to retrieve them from elsewhere (e.g. inter-library loans; directly contacting authors) (Stewart *et al.*, 2013). Authors should note whether the search strategy was designed to retrieve unpublished sources and grey literature. While most meta-analysts in ecology and evolutionary biology only search for published studies, the inclusion of unpublished data could substantially alter results (Sánchez-Tójar *et al.*, 2018).

Traditional systematic reviews aim to be comprehensive and find all relevant studies, published and unpublished, during a 'comprehensive search' (Primer A). In order to achieve comprehensive coverage in medical reviews, teams of systematic reviewers often employ an information specialist or research librarian. In ecology and evolutionary biology, it is more common to obtain a sample of available studies, sourced from a smaller number of sources and/or from a restricted time period. The validity of this approach depends on whether the sample is likely to be representative of all available studies; if the sampling strategy is not biased, aiming for a representative sample is justifiable (Cote & Jennions, 2013). We encourage authors to be transparent about the aim of their search and consider the consequences of sampling decisions. Further guidance on reporting literature searches is available from PRISMA-S (Rethlefsen *et al.*, 2021), and guidance on designing and developing searches is available from Bartels (2013), Bramer *et al.* (2018), Siddaway *et al.* (2019) and Stewart *et al.* (2013).

#### Item 6: Study selection

**Describe how studies were selected for inclusion at each stage of the screening process (e.g. use of decision trees, screening software). Report the number of people involved and how they contributed (e.g. independent parallel screening).**

##### Explanation and elaboration

As with finding studies, screening studies for inclusion is a time-consuming process that ecologists and evolutionary biologists rarely describe in their reports (average reporting quality <10%; see Supporting Information). Typically, screening is conducted in two stages. First, titles and abstracts are screened to exclude obviously ineligible studies (usually the majority of screened studies will be ineligible). Software can help speed up the process of title and abstract screening (e.g. Rathbone, Hoffmann & Glasziou, 2015b; Ouzzani *et al.*, 2016). Second, the full texts of potentially ineligible studies are downloaded (e.g. using a reference manager) and screened. At the full-text stage, the authors should record reasons why each full text did not meet the eligibility criteria (Item 19). Pre-determined, documented, and piloted eligibility criteria (Item 4) are essential for both stages of screening to

be reliable. Preferably, each study is independently screened by more than one person. Authors should report how often independent decisions were in agreement, and the process for resolving conflicting decisions (Littell, Corcoran & Pillai, 2008). To increase the reliability and objectivity of screening criteria, especially when complete independent screening is impractical, authors could restrict independent parallel screening to the piloting stage, informing protocol development. Regardless of how studies were judged for inclusion, authors should be transparent about how screening was conducted.

#### Item 7: Data collection process

**Describe where in the reports data were collected from (e.g. text or figures), how data were collected (e.g. software used to digitize figures, external data sources), and what data were calculated from other values. Report how missing or ambiguous information was dealt with during data collection (e.g. authors of original studies were contacted for missing descriptive statistics, and/or effect sizes were calculated from test statistics). Report who collected data and state the number of extractions that were checked for accuracy by co-authors.**

##### Explanation and elaboration

Describing how data were collected provides both information to the reader on the likelihood of errors and allows other people to update the review using consistent methods. Data extraction errors will be reduced if authors followed pre-specified data extraction protocols, especially when encountering missing or ambiguous data. For example, when sample sizes were only available as a range, were the minimum or mean sample sizes taken, or were corresponding authors contacted for precise numbers? Were papers excluded when contacted authors did not provide information, or was there a decision rule for the maximum allowable range (e.g. such that  $n = 10-12$  would be included, but  $n = 10-30$  would be excluded)? Another ambiguity occurs when effect sizes can be calculated in multiple ways, depending on which data are available (sensibly, the first priority should be given to raw data, followed by descriptive statistics – e.g. means and standard deviations, followed by test-statistics and then  $P$ -values). Data can also be duplicated across multiple publications and, to avoid pseudo-replication (Forstmeier *et al.*, 2017), the duplicates should be removed following objective criteria. Whatever precedent is set for missing, ambiguous, or duplicated information from one study should be applied to all studies. Without recording the decisions made for each scenario, interpretations can easily drift over time. Authors can record and report the percentages of collected data that were affected by missing, ambiguous, or duplicate information.

Data collection can be more efficient and accurate when authors invest time in developing and piloting a data collection form (or database), which can be made publicly available to facilitate updates (Item 18). The form should describe precisely where data were presented in the original studies, both to help re-extractions, and because some data sources are more reliable

than others. Using software to extract data from figures can improve reproducibility [e.g. metaDigitise (Pick, Nakagawa & Noble, 2019) and metagear (Lajeunesse, 2016)]. Ideally, all data should be collected by at least two people (which should correct for the majority of extraction errors). While fully duplicating extractions of large data sets might be impractical for small teams (Primer B), a portion of collected data could be independently checked. Authors can then report the percentage of collected data that were extracted or checked by more than one person, error rates, and how discrepancies were resolved.

#### Item 8: Data items

**Describe the key data sought from each study, including items that do not appear in the main results, or which could not be extracted due to insufficient information. Describe main assumptions or simplifications that were made (e.g. categorising both 'length' and 'mass' as 'morphology'). State the type of replication unit (e.g. individuals, broods, study sites).**

#### Explanation and elaboration

Data collection approaches fall on a spectrum between recording just the essential information to address the aim of the review, and recording all available information from each study. We recommend reporting both data that were collected and attempted to be collected. Complete reporting facilitates re-analyses, allows others to build upon previous reviews, and makes it easier to detect selective reporting of results (Primer C). For re-analyses, readers could be interested in the effects of additional data items (e.g. species information), and it is therefore useful to know whether those data are already available (Item 18). Similarly, stating which data were unavailable, despite attempts to collect them, identifies gaps in primary research or reporting standards. For selective reporting, authors could collect a multitude of variables but present only a selection of the most compelling results (inflating the risk of false positives; Primer C). Having a registered analysis plan is the easiest way to detect selective reporting (Item 3). Readers and peer-reviewers can also be alerted to this potential source of bias if it is clear that, for example, three different body condition metrics were collected, but the results of only one metric were reported in the paper.

### VIII. PRIMER D: BIAS FROM PRIMARY STUDIES

The conclusions drawn from systematic reviews and meta-analyses are only as strong as the studies that comprise them (Gurevitch *et al.*, 2018). Therefore, an integral step of a formal systematic review is to evaluate the quality of the information that is being aggregated (Pullin & Stewart, 2006; Haddaway & Verhoeven, 2015). If this evaluation reveals that the underlying studies are poorly conducted or biased, then a meta-analysis cannot answer the original research question. Instead, the synthesis serves a useful role in unearthing flaws in the existing primary studies and guiding newer studies (Ioannidis, 2016). While other fields emphasise

quality assessment, risk of bias assessment, and/or 'critical appraisal' (Cooke *et al.*, 2017), ecologists and evolutionary biologists seldom undertake these steps. When surveyed, authors, reviewers, and editors of systematic reviews and meta-analyses in ecology and evolutionary biology were largely oblivious to the existence of study quality or risk of bias assessments, sceptical of their importance, and somewhat concerned that such assessments could introduce more bias into the review (see Supporting Information). In this respect, little has changed since 2002, when Simon Gates wrote that randomization and blinding deserve more attention in meta-analyses in ecology (Gates, 2002).

It is difficult to decide upon metrics of 'quality' for the diverse types of studies that are typically combined in an ecological or evolutionary biology meta-analysis. We typically consider two types of quality – internal validity and external validity (James *et al.*, 2016). Internal validity describes methodological rigour: are the inferences of the study internally consistent, or are the inferences weakened by limitations such as biased sampling or confounds? External validity describes whether the study addresses the generalised research question. In ecology and evolutionary biology, the strongest causal evidence and best internal validity might come from large, controlled experiments that use 'best practice' methods such as blinding. If we want to generalise across taxa and understand the complexity of nature, however, then we need 'messier' evidence from wild systems. Note that in the medical literature, risk of bias (practically equivalent to internal validity) is considered a separate and preferable construct to 'study quality' (Büttner *et al.*, 2020), and there are well established constructs such as 'GRADE' for evaluating the body of evidence ('Grading of Recommendations, Assessment, Development and Evaluations'; Guyatt *et al.*, 2008). In PRISMA-EcoEvo we are broadly referring to study quality (Item 9) until such a time when more precise and accepted constructs are developed for our fields.

In PRISMA-EcoEvo we encourage ecologists and evolutionary biologists to consider the quality of studies included in systematic reviews and meta-analyses carefully, while recognising difficulties inherent in such assessments. A fundamental barrier is that we cannot see how individual studies were conducted. Usually, we only have the authors' reports to base our assessments on and, given problems with reporting quality, it is arguable whether the authors' reports can reliably represent the actual studies (Liberati *et al.*, 2009; Nakagawa & Lagisz, 2019). Quality assessments are most reliable when they measure what they claim to be measuring ('construct validity') with a reasonable degree of objectivity, so that assessments are consistent across reviewers (Cooke *et al.*, 2017). Despite the stated importance of quality assessment in evidence-based medicine, there are still concerns that poorly conducted assessments are worse than no assessments (Herbison, Hay-Smith & Gillespie, 2006), and these concerns were echoed in responses to our survey (see Supporting Information). Thoughtful research is needed on the best way to conduct study quality and/or risk-of-bias assessments. While internal validity (or risk of bias) will usually be

easier to assess, we urge review authors to be mindful of external validity too (i.e. generalisability).

## IX. ANALYSIS METHODS

### Item 9: Assessment of individual study quality

**Describe whether the quality of studies included in the meta-analysis was assessed (e.g. blinded data collection, reporting quality, experimental *versus* observational), and, if so, how this information was incorporated into analyses (e.g. meta-regression and/or sensitivity analysis).**

#### Explanation and elaboration

Meta-analysis authors in ecology and evolutionary biology almost never report study quality assessment, or the risk of bias within studies, despite these assessments being a defining feature of systematic reviews (average reporting quality <10%, see Supporting Information; Primer D). Potentially, authors are filtering out studies deemed unambiguously unreliable during the study selection process (Item 6), but this process is poorly reported, making reproducibility impractical. A more informative approach would be to code indicators of study quality and/or risk of bias within studies, and then use meta-regression or subgroup analyses to assess how these indicators impact the review's conclusions (Curtis *et al.*, 2013). While sensible in theory, quality assessment is difficult in practice (some might say impossible, given current reporting standards in the primary literature; Primer D). The principal difficulty is that we rely on authors' being reliable narrators of their conduct; omitting important information, such as the process of randomization, leaves us searching in the dark for a signal of study quality (O'Boyle, Banks & Gonzalez-Mulé, 2017). Uncertainty about the reliability of author reports is exacerbated by the absence of registration for most publications in ecology and evolutionary biology (Primer C). Until further research is conducted on reliable methods of quality assessment in our fields, we recommend review authors critically consider and report whether meaningful quality (or risk of bias) indicators could be collected from included studies. For example, indicators for experimental studies could include whether or not data collection and/or analysis was blinded for those collecting or analysing data [as blinding reduces the risk of bias (van Wilgenburg & Elgar, 2013; Holman *et al.*, 2015)] and whether the study showed full reporting of results [e.g. using a checklist such as Hillebrand & Gurevitch (2013); an example of the latter is shown in Parker *et al.* (2018a)]. Authors should then measure the impact that quality indicators have on the review's results (Items 16 and 24). Ultimately, as with collecting studies and data (Items 5 and 7), review authors are bound by the reporting quality of the primary literature.

### Item 10: Effect size measures

**Describe effect size(s) used. For calculated effect sizes (e.g. standardised mean difference, log response ratio) provide a reference to the equation**

**of each effect size and (if applicable) its sampling variance, or derive the equations and state the assumed sampling distribution(s).**

#### Explanation and elaboration

For results to be understandable, interpretable, and dependable, the choice of effect size should be carefully considered, and the justification reported (Harrison, 2010). For interpretable results it is essential to state the direction of the effect size clearly (e.g. for a mean difference, what was the control, was it subtracted from the treatment, or was the treatment subtracted from the control?). Sometimes, results will only be interpretable when the signs of some effect sizes are selectively reversed (i.e. positive to negative, or *vice versa*), and these instances need to be specified (and labelled as such in the available data; Item 18). For example, when measuring the effect of a treatment on mating success, both positive and negative differences could be 'good' outcomes (e.g. *more* offspring and *less* time to breed), so the signs of 'good' negative differences would be reversed.

Choosing an established effect size (such as Hedges' *g* for mean differences, or Fisher's *z* for correlations) carries the advantage of the effect size's statistical properties being sufficiently understood and described previously (Rosenberg, Rothstein & Gurevitch, 2013). When a non-conventional effect size is chosen, authors should provide equations for both the effect size and its sampling variance. Details should be provided on how the equations were derived, and how the sampling variance was determined (with analytic solutions or simulations) (Mengersen & Gurevitch, 2013).

### Item 11: Missing data

**Describe any steps taken to deal with missing data during analysis (e.g. imputation, complete case, subset analysis). Justify the decisions made.**

#### Explanation and elaboration

There are multiple methods to analyse data sets that are missing entries for one or more variables, therefore the chosen methods should be reported transparently. Statistical programs often default to 'complete case', deleting rows that contain missing data (empty cells) prior to analysis, but our assessment of reporting practices found it was uncommon for authors to state that complete case analysis was conducted (despite their data showing missing values for meta-regression moderator variables). Understandably, authors might not recognise the passive method of complete case analysis as a method of dealing with missing data, but it is important to be explicit about this step, both for the sample size implications (Item 20) and because of the potential to introduce bias when data are not 'missing completely at random' (Nakagawa & Freckleton, 2008; Little & Rubin, 2020). As an alternative to complete case analysis, authors can impute missing data based on the values of available correlated variables (e.g. multiple imputation methods, which retain uncertainty in the estimates of missing values; for discussion of these methods, see Ellington *et al.*, 2015). Data imputation can be used for missing moderator variables as well as information related to effect sizes (e.g. sampling



variances), thereby increasing the number of effect sizes included in analyses (Item 20) (Lajeunesse, 2013). Because imputation methods rely on the presence of correlated information, authors might extract additional data items to inform the imputation models, even if those data items are not of interest to the main analyses (Item 8). When justifying the chosen method, authors can conduct sensitivity analyses (Item 16) to assess the impact of missing data on estimated effects.

#### **Item 12: Meta-analytic model description**

**Describe the models used for synthesis of effect sizes. The most common approach in ecology and evolution will be a random-effects model, often with a hierarchical/multilevel structure. If other types of models are chosen (e.g. common/fixed effects model, unweighted model), this requires justification.**

##### **Explanation and elaboration**

Meta-analyses in ecology and evolutionary biology usually combine effect sizes from a broad range of studies, making it sensible to use a model that allows the 'true' effect size to vary between studies in a 'random-effects meta-analysis' (the alternative is a 'common' or 'fixed'-effect meta-analysis). Both frequentist and Bayesian statistical packages can implement random-effects meta-analyses. It is also common for multiple random effects to be included in a multilevel or hierarchical structure to account for non-independence (Item 14). Traditional meta-analytic models are weighted so that more precise effects have a greater influence on the pooled estimate than effects that are less certain (Primer A). In a random-effects meta-analysis, weights are usually taken from the sum of within-study sampling variance and the between-study variance. As a consequence of these variances being combined, large between-study variance will dilute the impact of within-study sampling variances. Alternatively, weights can be taken from the within-study sampling variances alone (as is done for common-effect models) (Henmi & Copas, 2010). When between-study variance is large (which can be assessed with heterogeneity statistics; Item 22), these two weighting structures could give different results. Authors could therefore assess the robustness of their results to alternative weighting methods as part of sensitivity analyses (Items 16 and 24).

Unweighted meta-analyses are regularly published in ecology and evolutionary biology journals, but we advise that these analyses be interpreted cautiously, and justified sufficiently. Theoretically, when publication bias is absent and effects have a normal sampling distribution, unweighted analyses can provide unbiased estimates (just with lower precision) (Morrissey, 2016). However, it is hard to detect effects that are inflated due to publication bias without sampling variances (Item 16), and from unweighted analyses we cannot estimate the contribution of sampling variance to the overall variation among effects (i.e. heterogeneity; Item 22). Unweighted analyses become more problematic for analyses of absolute values, because the 'folded' sampling distribution produces upwardly biased estimates (Nakagawa & Lagisz, 2016). Such analyses of

magnitudes, ignoring directions, are relatively common in ecology and evolutionary biology. There are two possible corrections for bias from analyses of absolute values (*sensu* Morrissey, 2016): (i) 'transform-analyse', where the folded distribution is converted to an unfolded distribution before analysis, or; (ii) 'analyse-transform', where the folded estimates are back-transformed to correct for bias.

#### **Item 13: Software**

**Describe the statistical platform used for inference (e.g. R), and the packages and functions used to run models, including version numbers. Specify any arguments that differed from the default settings.**

##### **Explanation and elaboration**

Given the many software options and methods available for conducting meta-analyses, transparent reporting is required for analyses to be reproducible (Sandve *et al.*, 2013). Authors should cite all software used and provide complete descriptions of version numbers. When describing software, it is easy to overestimate familiarity among the readership; changes from the default settings will not be obvious to some and should be described in full. That said, this item is less important than sharing data and code (Item 18) because shared code will convey much of the same information in a more reproducible form. Nonetheless it is helpful to describe software details in the text for the majority of readers who will not dig into the shared code.

#### **Item 14: Non-independence**

**Describe the types of non-independence encountered (e.g. phylogenetic, spatial, multiple measurements over time) and how non-independence has been handled, including justification for decisions made.**

##### **Explanation and elaboration**

Meta-analyses in ecology and evolutionary biology regularly violate assumptions of statistical non-independence, which can bias effect estimates and inflate precision. For example, studies containing pseudo-replication (Forstmeier *et al.*, 2017) have inflated sample sizes and downwardly biased sampling variances. When multiple effect sizes are derived from the same study, they are often not statistically independent, such as when multiple experimental groups are compared to the same control group (Gleser & Olkin, 2009). Alternatively, there may be non-independence among effect sizes across studies. There are numerous sources of non-independence at this level, including dependence among effect sizes due to phylogenetic relatedness (discussed in Primer E), and correlations between effect sizes originating from the same population or research group (Nakagawa *et al.*, 2019). Despite the ubiquity of non-independence in ecological and evolutionary meta-analyses, these issues are often not disclosed in the report (32% of 102 meta-analyses described the types of non-independence encountered; see Supporting Information). We recommend that authors report all potential sources of non-independence among effect sizes included in the meta-analysis, and the proportion of effect sizes that

are impacted (for further guidance, see Noble *et al.*, 2017; López-López *et al.*, 2018).

In addition to listing all sources of non-independence, authors should report and justify any steps that were taken to account for the stated non-independence. Steps range from the familiar (e.g. averaging multiple effect sizes from the same source, the inclusion of random effects, and robust variance estimation; Hedges, Tipton & Johnson, 2010) to the more involved (e.g. modelling correlations directly by including correlation or covariance matrices). Complicated methods of dealing with non-independence are best communicated through shared analysis scripts (Item 18). When primary studies are plagued by pseudo-replication (which could be considered in quality assessment; Item 9), an effect size could be chosen that is less sensitive to biased sample sizes (Item 10) (e.g. for mean differences, the log response ratio,  $\ln RR$ , rather than the standardised mean difference,  $d$ ), and (to be more conservative) sampling variances could be increased (Noble *et al.*, 2017).

It is not expected that non-independence can be completely controlled for. As with primary studies, problems of non-independence are complicated, and often the information necessary to solve the problem is unavailable (e.g. strength of correlation between non-independent samples or effect sizes, or an accurate phylogeny). Where there are multiple, imperfect, solutions, we encourage running sensitivity analyses (Item 16) and reporting how these decisions affect the magnitude and precision of results (Item 24).

#### **Item 15: Meta-regression and model selection**

**Provide a rationale for the inclusion of moderators (covariates) that were evaluated in meta-regression models. Justify the number of parameters estimated in models, in relation to the number of effect sizes and studies (e.g. interaction terms were not included due to insufficient sample sizes). Describe any process of model selection.**

##### **Explanation and elaboration**

When meta-regressions are used to assess the effects of moderator variables (i.e. meta-analyses with one or more fixed effects), the probability of false-positive findings increases with multiple comparisons. Therefore, rationales for each moderator variable should be provided in either the introduction or methods section of the meta-analysis manuscript. Analyses conducted solely for exploration and description should be distinguished from hypothesis-testing analyses (see Item 17). Authors should also report how closely the chosen moderator variables relate to the biological phenomena of interest (e.g. using mating call rate as a proxy for mating investment), and how the variables were categorised (Item 8).

For model selection and justification, principles from ordinary regression analyses apply to meta-regressions too (Gelman & Carlin, 2014; Harrison *et al.*, 2018; Meteyard & Davies, 2020). To avoid cryptic multiple hypothesis testing and associated high rates of false positive findings, authors should report full details of any model selection procedures (Forstmeier & Schielzeth, 2011). Under-powered meta-

regressions should be reported with obvious caveats (or else avoided completely), to discourage the results from being interpreted with unwarranted confidence (Tipton, Pustejovsky & Ahmadi, 2019). Meta-regressions have lower statistical power than meta-analyses, especially when including interaction terms between two or more moderator variables (Hedges & Pigott, 2004). Low statistical power can be due to any combination of too few available studies, small sample sizes within studies, or high amounts of variability between study effects, and therefore justification of meta-regression models should include consideration of sample sizes (Item 20) and heterogeneity (Item 22) (Valentine *et al.*, 2010).

#### **Item 16: Publication bias and sensitivity analyses**

**Describe assessments of the risk of bias due to missing results (e.g. publication, time-lag, and taxonomic biases), and any steps taken to investigate the effects of such biases (if present). Describe any other analyses of robustness of the results, e.g. due to effect size choice, weighting or analytical model assumptions, inclusion or exclusion of subsets of the data, or the inclusion of alternative moderator variables in meta-regressions.**

##### **Explanation and elaboration**

Reviews can produce biased conclusions if they summarise a biased subset of the available information, or if there is bias within the information itself. Authors should therefore assess risks of bias so that confidence in the conclusions (or lack thereof) can be accurately conveyed. Bias within the information itself was discussed in Item 9 and Primer D. Publication bias occurs when journals and authors prioritise the publication of studies with particular outcomes. For example, journals might prefer studies that support an exciting hypothesis, rather than publishing null or contradictory evidence (Rosenthal, 1979; Murtaugh, 2002; Leimu & Koricheva, 2004). Meta-analyses in ecology and evolutionary biology typically rely on published papers for data (Item 5) and hence are especially vulnerable to the effects of publication bias (e.g. Sánchez-Tójar *et al.*, 2018). Even if all research was published, the resulting papers would still provide information that was biased towards certain taxa (e.g. vertebrates), geographical locations (e.g. field sites close to Western universities), and study designs (e.g. short-term studies contained within the length of a PhD) (Pyšek *et al.*, 2008; Rosenthal *et al.*, 2017). Such ‘research biases’ (Gurevitch & Hedges, 1999) should be considered when categorising studies (Item 8).

While none are entirely satisfactory, multiple tools are available to detect publication bias in a meta-analytic data set (Møller & Jennions, 2001; Parekh-Bhurke *et al.*, 2011; Jennions *et al.*, 2013). Many readers will be familiar with funnel plots, whereby effect sizes (either raw, or residuals) are plotted against the inverse of their sampling variances, which should form a funnel shape. Asymmetries in the funnel indicate studies that are ‘missing’ due to publication bias, but could also be a benign outcome of heterogeneity (Item 22) (Egger *et al.*, 1997; Terrin *et al.*, 2003). Although funnel plots and Egger’s regression (a test of funnel plot asymmetry) were originally only useful for common-effect meta-analytic models (Egger *et al.*, 1997), modified methods have been

proposed to suit the random-effects meta-analytic models commonly used in ecology and evolutionary biology (Item 12; Moreno *et al.*, 2009; Nakagawa & Santos, 2012). Publication bias might also be indicated by a reduction in the magnitude of an effect through time (tested with a meta-regression using publication year as a moderator, or with a cumulative meta-analysis) (Jennions & Møller, 2002; Leimu & Koricheva, 2004; Koricheva & Kulinskaya, 2019). When biases are detected, we recommend authors report multiple sensitivity analyses to assess the robustness of the review's results (Rothstein, Sutton & Borenstein, 2005; Vevea, Coburn & Sutton, 2019). Subgroup analyses can be reported to test whether the original effect remains once the data set is restricted to recent studies, or studies that have been assessed to have a lower risk of bias (Item 9). To assess the sensitivity of the results to individual studies, authors can also report 'leave-one-out' analyses, and plot variability in the primary outcome (Willis & Riley, 2017).

#### **Item 17: Clarification of *post hoc* analyses**

**When hypotheses were formulated after data analysis, this should be acknowledged.**

##### **Explanation and elaboration**

Usually, a hypothesis should only be tested on data that were collected with the prior intention of testing that hypothesis. While a meta-analysis can test different hypotheses from those addressed in primary studies on which it is based, it is important that these hypotheses are formulated in advance. It is common, however, for researchers to be curious about patterns in their data after they have already been collected. Exploration and description are integral to research, but problems arise when such analyses are presented as hypothesis-testing, especially when they deviate from what is stated in a registration (also called 'Hypothesising After Results are Known' — see Primer C; Kerr, 1998). Ideally, authors will have a registered analysis plan (Item 3) to protect against the common self-deception that a *post-hoc* analysis was, in hindsight, the obvious *a-priori* choice (Parker *et al.*, 2016a). When a public registration is not provided, the reader is reliant on the memory and honesty of the authors. It is currently rare to see *post-hoc* acknowledgements in ecology and evolution meta-analyses (or indeed in primary studies) but, in the methods section, we encourage authors to state transparently which analyses were developed after data collection and, in the discussion, temper confidence in the results of such analyses accordingly.

#### **Item 18: Metadata, data, and code**

**Share metadata (i.e. data descriptions), data, and analysis scripts with peer-reviewers. Upon publication, upload this information to a permanently archived website in a user-friendly format. Provide all information that was collected, even if it was not included in the analyses presented in the manuscript (including raw data used to calculate effect sizes, and descriptions of where data were located in papers). If a software package with graphical user interface (GUI) was used, then describe full model specification and fully specify choices.**

##### **Explanation and elaboration**

Sharing data, metadata, and code scripts (or equivalent descriptions) is the only way for authors to achieve 'computational reproducibility' (the ability to reproduce all results presented in a paper) (Piccolo & Frampton, 2016). Data, metadata and code files also preserve important information in a format that some readers will find easier to understand than wordy summaries (e.g. descriptions of statistical methods). Thanks to a decadal shift in journal policies to mandate data sharing, this is one aspect of reporting where ecological and evolutionary biology meta-analyses are ahead of medical fields (Roche *et al.*, 2015; Sholler *et al.*, 2019). During this coming decade our community can raise the bar higher by mandating sharing of metadata and analysis scripts (or complete descriptions of workflow for point-and-click software). We strongly encourage authors to provide data in a user-friendly file format [such as a .csv file rather than a table in a .pdf or .doc file; see also the 'FAIR' principles by Wilkinson *et al.* (2016), urging for data to be findable, accessible, interoperable, and reusable]. Currently, more than half of meta-analyses in ecology and evolution share data without corresponding metadata (i.e. complete descriptions of variable names; see Supporting Information), and this can render the data itself unusable. Authors should also consider sharing other materials that, in a spirit of collegiality, could be helpful to other researchers (e.g. bibliographic files from a systematic search; Item 5).

Data and code should be provided from the peer-review stage (Goldacre, Morton & DeVito, 2019) (for double-blind reviews, files can be anonymised). Currently, the requirement by most journals in ecology and evolutionary biology for data to be shared upon publication reduces this important task to an afterthought. We recognise that many peer-reviewers, who are already overburdened, will not check computational reproducibility. But some will, and this background level of accountability should improve the standards of authors, who are ultimately responsible for the trustworthiness of their work. From our experiences, the analyses described in papers can differ markedly from what is contained in the code, and data collection and processing mistakes are common. Evidence from other fields suggests that data and/or code provided by authors often do not reproduce the results presented in a paper (Nosek *et al.*, 2015; Stodden, Seiler & Ma, 2018). For full transparency and computational reproducibility, authors should provide raw and pre-processed data and the accompanying scripts (Sandve *et al.*, 2013; Piccolo & Frampton, 2016). That way, the pipeline of data tidying, calculations (including effect size calculations), and any outlier exclusions can be reproduced. Rather than uploading these materials as static files along with the supplementary materials, which might not be stored permanently, we recommend authors create an active project repository (e.g. on the Open Science Framework) so that, if the authors wish, the materials and code can be improved (e.g. if readers spot small mistakes).

Occasionally it might be justifiable to withhold raw data (e.g. due to confidentiality or legal issues, or if future projects are planned with the data). In such a case, a dummy data set, approximately replicating the real data, could be made available to peer-reviewers along with analysis scripts. Because meta-analyses in ecology and evolutionary biology typically summarise published studies, which are

unlikely to contain ethically sensitive information (such as precise locations of endangered species), it is exceedingly rare that withholding data past an embargoed date is justifiable. Aiming for computational reproducibility can be burdensome but, increasingly, these efforts should be viewed favourably by academic reward systems (Moher *et al.*, 2018). Rewards aside, as authors of reviews we should practice what we preach. When collecting data for a meta-analysis, much time is wasted in frustration when data are not adequately reported in published papers, and requests for data from authors have varying success (Item 7). Archiving data from primary studies in an online repository would make the data collection process far easier, faster and more accurate. If we wish this practice from authors of the primary studies that systematic reviews and meta-analyses are reliant upon, then we should hold ourselves to the same standard.

## X. PRIMER E: PHYLOGENETIC NON-INDEPENDENCE

Meta-analyses in ecology and evolutionary biology often encounter phylogenetic non-independence (or relatedness). Phylogenetic non-independence – a special type of statistical non-independence – occurs because we usually combine data originating from different species, whose evolutionary history causes them to be related to each other to varying extents (i.e. each species is not an independent unit) (Noble *et al.*, 2017). Phylogenetic signal in a meta-analytic data set may impact the outcome of the analysis (Chamberlain *et al.*, 2012). In many cases we can model phylogenetic non-independence by converting a phylogenetic tree into a correlation matrix, which describes relationships among species in the data set. The matrix can then be incorporated into the meta-analytic model, which becomes a ‘phylogenetic meta-analysis’ (Adams, 2008; Lajeunesse, 2009; Hadfield & Nakagawa, 2010). A phylogenetic meta-analysis is mathematically identical to a phylogenetic comparative model that accounts for sampling error in the measured traits (Nakagawa & Santos, 2012).

Advances in phylogenetic comparative methods and software have made it superficially easy to incorporate phylogeny into meta-analytic models, but the particulars of the methods are contestable. Phylogenetic trees vary in quality and rely on simplifying assumptions (e.g. Brownian motion model of evolutionary divergence) (Harmon, 2018). When a meta-analysis combines data from a diverse collection of species, it becomes harder to resolve deep phylogenetic relationships, and some species might be excluded from existing trees (resulting in incomplete data). One solution to uncertainty in individual trees is to incorporate multiple trees into the analyses (Nakagawa & de Villemerreuil, 2019). Once you have a tree, there are different methods – corresponding to different models of evolution – to convert the relationships into a correlation matrix (Nakagawa & Santos, 2012). Given

that phylogenetic comparative methods remain an active area of research and different analysis assumptions could lead to different outcomes, authors could present results from multiple analyses as part of sensitivity analyses (Item 24) (and whether or not a meta-analytic data set contains a phylogenetic signal is itself a potentially revealing outcome). Regardless of how authors choose to handle phylogenetic non-independence, transparency is essential for a phylogenetic meta-analysis to be reproducible (Borries *et al.*, 2016).

## XI. REPORTING RESULTS

### Item 19: Results of study selection process

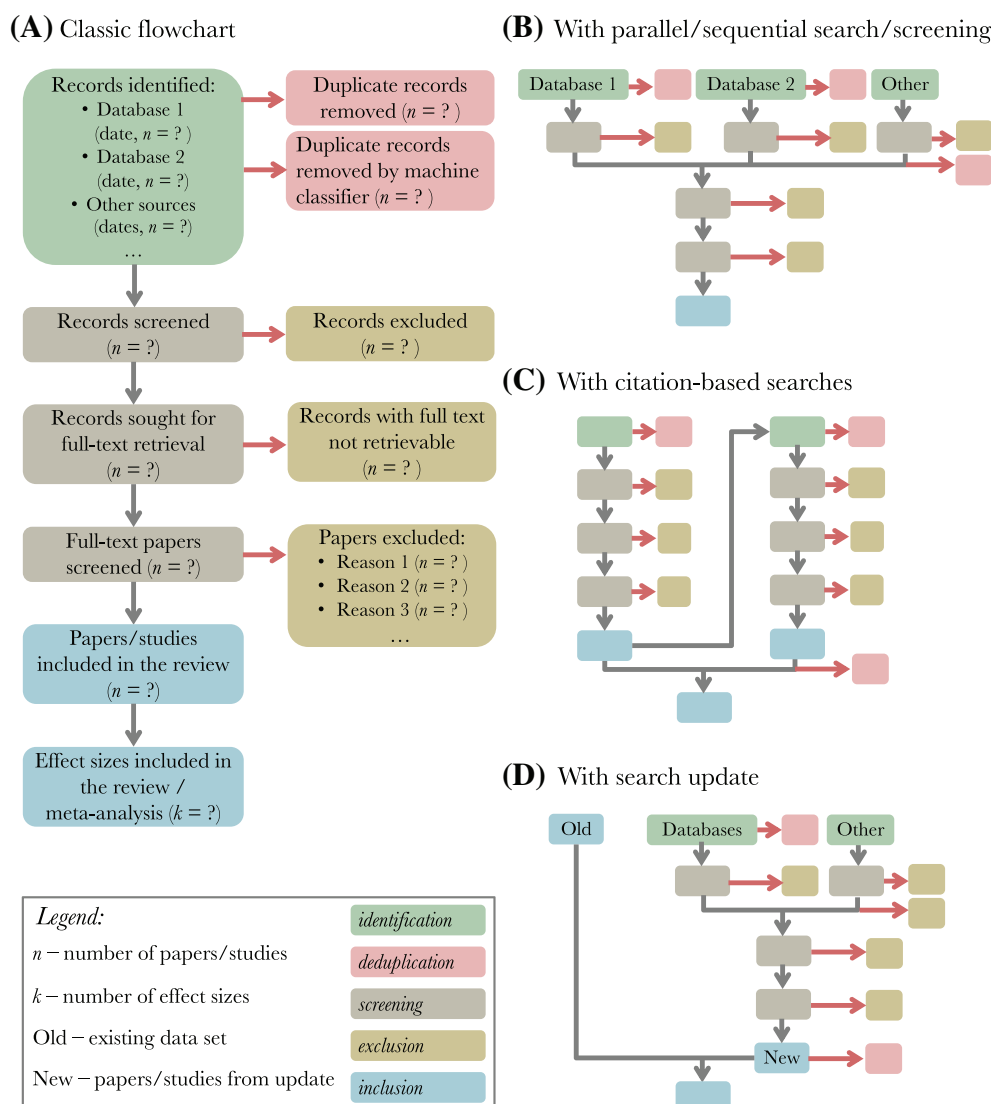
**Report the number of studies screened, and the number of studies excluded at each stage of screening (with brief reasons for exclusion from the full-text stage). Present a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)-like flowchart** ([www.prisma-statement.org](http://www.prisma-statement.org)).

#### Explanation and elaboration

Without knowing the number of studies found (Item 5) and screened (Item 6), and the reasons why full-text articles were excluded, the reliability of a systematic search is unclear. The flow-chart template, originally provided by PRISMA in 2009 and modified in 2020, presents this information in a concise and consistent format, and it is the aspect of PRISMA that is most commonly referenced by meta-analysis authors in ecology and evolutionary biology (see Supporting Information). Conceptual examples of PRISMA flowcharts are shown in Fig. 3; authors can customise the flowchart as they please, but the following should be presented: (i) the number of records that originated from each source of information; (ii) the number of records after duplicates were removed; (iii) the number of full texts screened; (iv) the tallied reasons why full texts were excluded; and (v) the total number of studies included in the systematic review and meta-analysis (which might differ), and the number of effect sizes. Tracking these five details during the search and screening process requires conscientious workflows. In addition, as recommended in the updated PRISMA flow diagram, authors could list the number of records that were removed by machine learning classifiers, and the number of full texts that could not be retrieved. While the flowchart summarises why full texts were excluded, for full accountability we recommend that authors provide the references of all articles excluded at the full text stage, alongside reasons for their exclusion.

### Item 20: Sample sizes and study characteristics

**Report the number of studies and effect sizes for data included in meta-analyses, and subsets of data included in meta-regressions. Provide a summary of key characteristics for reported outcomes (either in text or figures; e.g. one quarter of effect sizes reported for vertebrates and the rest invertebrates)**



**Fig 3.** PRISMA-style flowcharts and some variations. (A) The classic flow-chart: all searches are conducted around the same date, and screening occurs after de-duplication. (B) Records are obtained from different databases (or other sources, e.g. personal archives or requests) and screened separately. De-duplication occurs after at least one stage of screening. (C) The studies included after a classic search are then used as the ‘seed’ for a new search, based on citation information. Authors can retrieve all papers cited in included articles (backwards search), and all papers that cite the included articles (forwards search). A second round of de-duplication and screening then occurs. (D) When a systematic review is an update of an already existing one, the newly found papers are added to the existing (old) set of included papers. As a further extension, it would be beneficial to record, and report, how many of the included articles originate from each source. For example, if one database contributed none of the included articles, then updates of the review could save time by not screening articles from that database.

### and their limitations (e.g. collinearity and overlaps between moderators), including characteristics related to individual study quality (risk of bias).

#### Explanation and elaboration

Meta-analyses and meta-regressions cannot answer questions for which there are no, or almost no, data available. It is therefore essential to report complete sample sizes for every analysis that was conducted (these can be reported in supplementary tables, if brevity is required). Authors should provide sample sizes for the number of studies (or equivalent unit of analysis) because, for example, it would be misleading

to withhold that a sample size of  $k = 20$  effect sizes originated from only  $n = 2$  studies. Figures or tables can be used to report sample sizes concisely across multiple hierarchical levels of meta-analyses, as well as across and within different moderator variables included in meta-regressions [e.g. see fig. 3 in Lagisz *et al.* (2020) and table 1 in Chaplin-Kramer *et al.* (2011)].

When presenting results from meta-regressions, authors should report complete case sample sizes for moderators containing missing data (Item 11) and, in the case of categorical (i.e. discrete) moderators, sample sizes for all included



categories. In meta-regressions with more than one moderator it is important to consider the extent to which moderator variables overlap. In the case of multiple categorical variables, authors should report sample sizes across all possible combinations of categories. For example, in a meta-regression including the interaction between ‘vertebrate or invertebrate’ and ‘urban or wild’ moderator variables, authors could report  $n = 8$  studies on invertebrates were divided into  $n = 6$  studies in urban environments, and  $n = 2$  in wild environments, while  $n = 20$  studies on vertebrates were evenly split with  $n = 10$  studies in each category of urbanisation. When reporting meta-regressions with both continuous and categorical moderators, we recommend reporting the amount of coverage of the continuous moderator within each category using descriptive statistics or data visualisations (e.g. when including the continuous fixed effect of study year, authors should report whether studies conducted in ‘urban’ and ‘wild’ environments spanned a similar time period).

It is important to report sample sizes comprehensively so that the risk of inaccurate parameter estimates can be evaluated. Statistical power in a random-effects meta-analysis (Item 12), with a single random effect for study identity, is influenced by: the number of included studies and sample sizes within them (i.e. number of independent effect sizes and their precision); the amount of variation in effects between studies; the (pre-specified) size of the ‘true’ effect being investigated; and the accepted probability of falsely rejecting the null hypothesis (conventionally set at  $\alpha = 0.05$  for ecological and evolutionary analyses) (Valentine *et al.*, 2010). Statistical power is always lower for estimates of moderating effects compared to the meta-analytic mean, and much lower for estimates of the interaction between multiple moderating effects (Hedges & Pigott, 2004). Power calculations are further complicated by uneven sampling within studies and multiple types of statistical non-independence (as is common in ecology and evolutionary biology; Item 14). Data simulations are usually required to estimate the probability of false-positive results for these complex analyses [for further guidance on these issues, see Gelman & Carlin (2014), Tipton *et al.* (2019) and Valentine *et al.* (2010)].

A key feature of any review is to summarise what research has been conducted and highlight gaps in the literature. For broad topics this can be the sole purpose of the review. For example, ‘systematic maps’ take a snapshot of the current state of research, and ‘bibliometric maps’ chart a field’s development by analysing publication and citation histories (Nakagawa *et al.*, 2019). The topics summarised by meta-analyses might be comparatively narrow, but their results are still context-dependent; for example, if all available studies were on temperate bird species it would be misleading to generalise about fish, or even tropical birds (Pyšek *et al.*, 2008). Most ecology and evolution meta-analyses contain too many studies for the characteristics of each individual study to be conveyed to the reader. Authors could therefore adopt some ‘mapping’ tools (e.g. Haddaway *et al.*, 2019) to distil the key characteristics of their data set in a concise (and sometimes beautiful)

format and summarise the magnitude and direction of research biases (Item 16).

#### **Item 21: Meta-analysis**

**Provide a quantitative synthesis of results across studies, including estimates for the mean effect size, with confidence/credible intervals.**

##### **Explanation and elaboration**

The meta-analytic mean and associated confidence interval can be provided in the text or displayed graphically. For some questions the primary outcome will be slopes or contrasts from meta-regressions (Item 23), and the meta-analytic mean or its statistical significance might not be biologically interesting (e.g. for analyses of absolute values). In those cases, authors can justify not displaying the meta-analytic mean.

#### **Item 22: Heterogeneity**

**Report indicators of heterogeneity in the estimated effect (e.g.  $I^2$ ,  $\tau^2$  and other variance components).**

##### **Explanation**

Statistical heterogeneity in a meta-analysis describes variation in the outcome between studies that exceeds what would be expected by sampling variance alone. When heterogeneity is high, meta-regressions can be run to see whether moderator variables account for some of the unexplained variation in outcomes. It is trickier to quantify heterogeneity for multilevel models (which are commonly used to account for non-independence; Item 14), but methods are available (Nakagawa & Santos, 2012). For unweighted meta-analyses (Item 12) it is not possible to quantify heterogeneity as properly defined, but residual errors can be used as a surrogate (as sampling errors will be absorbed into the residual error component). High heterogeneity is the norm in ecology and evolution meta-analyses (Primer B; Senior *et al.*, 2016). Given that we usually summarise studies that differ in multiple ways, it is surprising and notable to find an effect with low heterogeneity (e.g. Rutkowska, Dubiec & Nakagawa, 2013). Opinions differ on the most informative heterogeneity statistics, and the best way to report them. In addition to presenting a single summary of heterogeneity for a meta-analysis, it might be beneficial to estimate heterogeneity on subsets of studies that are considered more homogeneous (as characterised in Item 20). Authors can also present prediction intervals alongside confidence/credible intervals, to capture uncertainty in the predicted effect from a future study (IntHout *et al.*, 2016; Nakagawa *et al.*, 2020).

#### **Item 23: Meta-regression**

**Provide estimates of meta-regression slopes (i.e. regression coefficients) for all variables that were assessed for their contribution to heterogeneity. Include confidence/credible intervals, and report interactions if they were included. Describe outcomes from model selection, if done (e.g.  $R^2$  and AIC).**

##### **Explanation and elaboration**

Meta-regressions test whether a given variable moderates the magnitude or direction of an effect, and can therefore provide ‘review’ or ‘synthesis-generated’ evidence (Cooper, 2009;

Nakagawa *et al.*, 2017). Moderator variables can be biological (e.g. sex), methodological (e.g. experimental design; Dougherty & Shuker, 2015), and sociological (e.g. publication status; Item 24; Murtaugh, 2002). Authors should distinguish between results from exploratory and hypothesis-testing models (Item 15 and Item 17), with the latter requiring complete reporting of the number of tests that were run (so that the false-positive discovery rate can be adjusted for multiple comparisons) (Forstmeier *et al.*, 2017). When justifying the choice of model, authors can present statistical parameters of model fit (e.g.  $R^2$ ) and information criterion [e.g. Akaike Information Criterion (AIC) and Deviance Information Criterion (DIC)]. When meta-regressions are run, authors can present  $Q$ -statistics to quantify whether moderator variables account for significant heterogeneity (Item 22). Model selection is a difficult and debated topic (Dennis *et al.*, 2019), but all methods share the principle of transparently reporting outcomes.

Authors should provide all estimates from all models (i.e. complete reporting), to avoid the 'Texas sharpshooter fallacy' (whereby someone randomly fires many gunshots at a barn, paints a target around the tightest cluster, and then claims they are a good shot; Evers, 2017) (Primer C; Item 3; Item 17). Complete reporting will often require supplementary tables (especially for extensive sensitivity analyses; Item 24). Figures are an effective way to communicate the main results from meta-regressions. For categorical moderator variables, it is common to plot the estimate of each category's intercept, with whiskers representing confidence (or credible) intervals (showing uncertainty in the mean effect for a given category). Authors can also include prediction intervals — showing uncertainty in the value of effect sizes from future studies — which provide intuitive displays of heterogeneity (Item 22) (for further ideas on displaying prediction intervals, see Nakagawa *et al.*, 2020). For categorical variables with more than two levels, authors who wish to make inferences about the differences between mean estimates ('intercepts') should report the precision of all 'slopes' or 'contrasts', not just the contrast from one baseline category. Continuous moderator variable slopes can be displayed on top of a scatterplot or, better yet, a 'bubble plot' of raw effect sizes (in a bubble plot, the size of points can represent weights from the meta-regression model; Lane *et al.*, 2012).

#### **Item 24: Outcomes of publication bias and sensitivity analyses**

**Provide results for the assessments of the risks of bias (e.g. Egger's regression, funnel plots) and robustness of the review's results (e.g. subgroup analyses, meta-regression of study quality, results from alternative methods of analysis, and temporal trends).**

##### **Explanation and elaboration**

Results from meta-analyses need to be considered alongside the risk that those results are biased, due to biases either across or within studies (i.e. publication bias, or problems with the included studies). When authors find evidence of bias, they should estimate the impact of suspected bias on the reviews' results (i.e. robustness of results; Item 16). More

generally, authors should consider if their results are robust to subjective decisions made during the review, such as: eligibility criteria (Item 4); data processing (including outlier treatment and choice of effect size; Item 10); and analysis methods (including how non-independence was handled; Items 12 and 14). There are usually multiple justifiable ways to conduct a meta-analysis: each subjective decision creates an alternative path, but results are robust when many paths lead to the same outcome (Palpacuer *et al.*, 2019). Multiple sensitivity analyses will generate an abundance of information that can be presented within the supplementary information.

## **XII. DISCUSSION AND CONTRIBUTIONS**

### **Item 25: Discussion**

**Summarise the main findings in terms of the magnitude of effects, their precision (e.g. size of confidence intervals, statistical significance), their heterogeneity, and biological/practical relevance. Compare results with previous reviews on the topic, if available. Discuss limitations and their influence on the generality of conclusions, such as gaps in the available evidence (e.g. taxonomic and geographical research biases).**

#### **Explanation and elaboration**

There are six notes that we think authors should hit in their discussions, while allowing for variety in personal styles and journal requirements. First and second, both the magnitude and precision of the main results should be discussed. Some readers will skip straight from the introduction to the discussion, so it should be clear whether the effects being discussed are small or large, and whether they have been estimated with confidence. Third, do not ignore variation among studies when summarising aggregated results (i.e. discuss heterogeneity; Item 22). Fourth, put the results into a wider biological context (or state if the available evidence does not provide one; do not overreach). Such discussions can include the generation of testable hypotheses from exploratory analyses (Item 17). Fifth, discuss how the results of previous reviews or influential studies are strengthened or undermined by the current evidence. Sixth, discuss limitations of the current research, caused by either the methods of the review itself, or the information that was available from the primary literature. Limitations in the primary literature can refer to both the quality of individual studies (Primer D), and knowledge gaps (research biases; Item 20), both of which can be addressed by future research. Conversely, authors could identify types of studies that have been sufficiently common such that that future resources would be better spent elsewhere.

### **Item 26: Contributions and funding**

**Provide names, affiliations, contributions, and all funding sources of all co-authors. Provide**

**contact details for the corresponding author. Disclose any conflicts of interest.**

**Explanation and elaboration**

The corresponding author should provide an email address with no intended expiry date (institutional email addresses might expire when authors leave the institution). The recent uptake in the author identifier ORCID has made it easier to contact authors when their corresponding email address changes (or to contact co-authors when the corresponding author is no longer available). In medical fields, systematic reviews are regularly financed by private companies that might have financial conflicts of interest in the review's conclusions. When such external and financial conflicts of interest are present (which is conceivable for applied topics in ecology and evolutionary biology) it is expected they are disclosed. Currently, it is not common to disclose internal conflicts of interests, such as the professional benefits authors expect from publishing a high-profile review paper. For contributions statements, authors usually only list their contributions when there is a dedicated section in the journal, but we encourage authors to take it upon themselves to precisely state their contributions in the methods or acknowledgements sections (Ewers *et al.*, 2019) (see also 'CRediT' – Contributor Roles Taxonomy; Holcombe, 2019).

**Item 27: References**

**Provide a reference list of all studies included in the meta-analysis. Whenever possible, list included studies as referenced sources (e.g. rather than listing them in a table or supplement).**

**Explanation and elaboration**

Studies included in a review article should be cited, so that they appear in the citation counts of scientific databases, to ensure that authors of primary studies are credited for their contribution (Kueffer *et al.*, 2011). For authors of primary studies, it can be frustrating to have their hard work included in reviews but not cited, as most commonly occurs when sources for a meta-analysis are listed in a table or the supplementary information. To give primary studies their due credit, studies included in the review can either (i) be included in the main reference list of the paper, and indicated by an asterisk (or other symbol) to distinguish them from papers cited elsewhere in the review; or (ii) be listed in a secondary reference list that appears at the bottom of the journal article. The latter option has the advantage of delineating between studies included in the meta-analysis and studies that are cited in the text (e.g. Li *et al.*, 2010; Kinlock *et al.*, 2018), while allowing all studies to be correctly indexed within citation databases. Recognising this choice is not always available to authors, we encourage journals, and editors of review articles, to require all studies included in a meta-analysis be included in the main reference list, or else to ensure that citations in supplementary information are appropriately indexed.

### XIII. HOW AND WHY TO USE PRISMA-ECOEOVO

We aim for PRISMA-EcoEvo to help the ecology and evolution community raise the quality of reporting in systematic

reviews and meta-analyses. Improved reporting in review articles is in the best interests of everyone. For authors of reviews, it will become easier to build upon earlier work; published methods and materials will make it easier to conduct both original reviews and update old ones. Clear reporting will help editors and reviewers provide reliable assessments of meta-analysis manuscripts. For the fields of ecology and evolutionary biology as a whole, well-reported systematic reviews provide clarity on what research has been done, what we can and cannot say with confidence, and which topics deserve attention from empiricists.

The PRISMA-EcoEvo checklist is available to download at <https://osf.io/t8qd2/>, and as an interactive web application at <https://prisma-ecoevo.shinyapps.io/checklist/>. The web application allows for automatic reporting quality assessments that are consistent with those detailed in the Supporting Information. Over time, PRISMA-EcoEvo can be updated to reflect improved reporting standards; users can provide feedback for an update by going to <https://doi.org/10.17605/OSF.IO/GB5VX> and following the instructions therein.

#### (1) PRISMA-EcoEvo for authors

Authors of systematic reviews and meta-analyses in ecology and evolutionary biology could make use of PRISMA-EcoEvo at any stage. Most of the items can be profitably planned ahead of time and registered in a time-stamped public repository or submitted to a journal as a registered report (Primer C; see also PRISMA for Protocols; Moher *et al.*, 2015). When conducting a systematic review and meta-analysis, heeding the checklist items will help with organisation and project management, as PRISMA-EcoEvo provides a guide for which pieces of information should be tracked and recorded. For in-depth conduct guidance, review authors should consult more specialised resources, many of which are cited in the items above. Once a systematic review and meta-analysis is completed and is being prepared for submission to a journal, PRISMA-EcoEvo provides a checklist for authors on what to report in their manuscript. Checklists are useful tools as they reduce cognitive burdens (Parker *et al.*, 2018b). Conscientious authors could assist reviewers by submitting the checklist alongside the key location for each item. When items are not applicable for their study, authors can provide a brief explanation. Upon first reading the checklist, the amount of detail and information might overwhelm some authors, but items that appear out of reach can be considered stretch goals (i.e. things to aim for in the future). We do not expect any authors, including ourselves, to achieve every item all at once, but we aim to make small improvements over time. Even if every manuscript reported only one extra item, we would raise reporting standards for the whole field.

#### (2) PRISMA-EcoEvo for reviewers and editors

Editors and reviewers of meta-analyses in ecology and evolutionary biology can use PRISMA-EcoEvo as a checklist for

complete reporting and transparency in the manuscripts they assess. Peer-reviewers should feel empowered to request additional information from authors, including data and analysis scripts. When we assessed reporting quality in the current meta-analytic literature (detailed in the Supporting Information), it was impossible to evaluate multiple aspects of reporting quality haphazardly; typically some items were reported well and others poorly, so different aspects of reporting (i.e. items on the checklist) needed to be assessed systematically. Journals could reduce the burden on reviewers by requesting authors to submit a completed PRISMA-EcoEvo checklist alongside their manuscript (as is done in many medical journals). Importantly, a poorly reported manuscript does not mean that the study itself was poorly conducted, but to be able to work out the quality of a study, reporting quality needs to be high.

### (3) Valuing permanent online repositories

For PRISMA-EcoEvo to be effective, authors will need to report information supplementary to the main paper. The legacy of print journals engendered brevity in papers, which remains, in part, for the benefit of the typical reader who does not want, or need, the tedium of complete reporting. Permanent, findable, and accessible online repositories are therefore essential for those readers who want to know the basis for the conclusions of a review article. Our survey of reporting standards suggested supplementary resources are currently under-used (data shown in the Supporting Information), especially given that free and permanent online repositories remove the barriers previously imposed by journals (such as finite archiving and fees). The community may need to undergo a cultural shift to better appreciate and value materials supplementary to the main paper. Currently, authors might understandably feel the time spent preparing such materials does not return adequate benefits, but this could change if editors and reviewers request the additional information, and if supplementary materials are cited independently of the main text (Moher *et al.*, 2018) (the materials can have their own DOI through platforms such as the Open Science Framework).

### (4) Comparisons with PRISMA-2020

PRISMA-EcoEvo continues a long and valuable tradition of meta-analysts in ecology and evolution learning from practices in evidence-based medicine. To improve review methods, we often look towards medical fields for inspiration, as they are continually improving (due to both more researchers, funding, and stakeholders). PRISMA-2020 requests greater reporting detail than the original PRISMA checklist did, and we encourage systematic review and meta-analysis authors in ecology and evolution to read the updated statement paper (Page *et al.*, 2021b) as well as the explanations and elaborations (Page *et al.*, 2021a). In comparing our respective fields, ecology and evolutionary biology

currently lags furthest behind in review registration (Primer C) and assessing individual study quality (Primer D). Both these areas are currently contentious (as voiced by respondents to our survey; see Supporting Information), but we hope to see improvements spurred by PRISMA-EcoEvo. Two areas where ecology and evolution might be ahead of some medical fields are in the consideration of statistical non-independence (because our data are often more 'complex'), and data sharing. We can further strive to improve the useability of our shared data (e.g. with better metadata), and lift code sharing to match the level of data sharing.

## XIV. CONCLUSIONS

- (1) Systematic reviews and meta-analyses are vital contributions to research fields when conducted well. However, when conducted poorly, erroneous conclusions can mislead readers. Transparent and complete reporting of these studies is therefore required, both so that confidence in the review's results can be accurately assessed, and to make the review updatable when new information is published.
- (2) Evidence suggests that reporting guidelines and checklists improve reporting quality. PRISMA, the most-cited guideline, was developed for reviews of medical trials. Multiple extensions of PRISMA have already been published to suit different types of reviews, but until now there has not been a checklist specifically designed for ecology and evolution meta-analyses. Having explanations and examples targeted at the ecology and evolutionary biology community should increase uptake of the guideline in this field.
- (3) We created an extension of PRISMA to serve the ecology and evolution systematic review and meta-analysis community: version 1.0 of PRISMA-EcoEvo. PRISMA-EcoEvo is a 27-item checklist that outlines best reporting practices as they currently stand.
- (4) Authors, editors, and reviewers can use PRISMA-EcoEvo for systematic review and meta-analysis publications (both traditional papers, and registered reports). Authors can use it before, during, and after conducting a review to assist with recording and reporting aims, methods, and outcomes. Editors and reviewers can use PRISMA-EcoEvo to increase reporting standards in the systematic review and meta-analysis manuscripts they review.
- (5) Collectively, the meta-analysis community can improve reporting standards by including (and requesting) more PRISMA-EcoEvo items in the manuscripts they prepare, and review, over time. Research will be more efficient and effective when published reviews are transparent and reproducible.

## XV. ACKNOWLEDGEMENTS, AUTHOR CONTRIBUTIONS AND DATA AND CODE AVAILABILITY

Data collection for this project was funded through Australia Research Council Discovery Grants: DP180100818 to S.N., and DP190100297 to M.D.J. We are grateful to 208 anonymous members of the ecology and evolution meta-analysis community for providing feedback on PRISMA-EcoEvo during its development. We thank Alison Bell and Bob Wong for providing feedback on earlier drafts of both the main text and supporting information, and two anonymous reviewers for their constructive comments on the submitted manuscript. Finally, we are grateful to the Cambridge Philosophical Society for making this article open access.

**Author contributions:** conceptualisation: R.E.O., M.L., M.D.J., J.K., D.W.A.N., T.H.P., J.G., M.J.P., G.S., D.M., S.N.; data curation: R.E.O.; formal analysis: R.E.O.; funding acquisition: S.N.; investigation: R.E.O., M.L., M.D.J., J.K., D.W.A.N., T.H.P., S.N.; methodology: R.E.O.; project administration: R.E.O., S.N.; software: R.E.O.; supervision: D.M., S.N.; visualisation: R.E.O., M.L.; writing – original draft: R.E.O.; writing – review & editing: R.E.O., M.L., M.D.J., J.K., D.W.A.N., T.H.P., J.G., M.J.P., G.S., D.M., S.N.

**Data and code availability:** data and code for the assessment of reporting standards and survey of community attitudes are available from <http://doi.org/10.17605/OSF.IO/2XPFG>

## XVI. REFERENCES

- References marked with an asterisk (\*) were included in the assessment of reporting quality. The symbol † identifies additional references that are cited in the online supporting information. \*ABE, J. & KAMIMURA, Y. (2012). Do female parasitoid wasps recognize and adjust sex ratios to build cooperative relationships? *Journal of Evolutionary Biology* **25**, 1427–1437.
- ACZEL, B., SZASZI, B., SARAFIOLOU, A., KEKECS, Z., KUCHARSKÝ, Š., BENJAMIN, D., CHAMBERS, C. D., FISHER, A., GELMAN, A., GERNSBACHER, M. A., IOANNIDIS, J. P., JOHNSON, E., JONAS, K., KOUSTA, S., LILIENTHAL, S. O., et al. (2020). A consensus-based transparency checklist. *Nature Human Behaviour* **4**, 4–6.
- ADAMS, D. C. (2008). Phylogenetic meta-analysis. *Evolution* **62**, 567–572.
- ALLEN, C. & MEHLER, D. M. A. (2019). Open science challenges, benefits and tips in early career and beyond. *PLoS Biology* **17**, e3000246.
- \*AMAT, F. & MEIRI, S. (2017). Geographical, climatic and biological constraints on age at sexual maturity in amphibians. *Biological Journal of the Linnean Society* **123**, 34–42.
- \*BARRIENTOS, R. (2015). Adult sex-ratio distortion in the native European polecat is related to the expansion of the invasive American mink. *Biological Conservation* **186**, 28–34.
- \*BARRIO, I. C., HIK, D. S., BUENO, C. G. & CAHILL, J. F. (2012). Extending the stress-gradient hypothesis - is competition among animals less common in harsh environments? *Oikos* **122**, 516–523.
- BARTELS, E. M. (2013). How to perform a systematic search. *Best Practice & Research in Clinical Rheumatology* **27**, 295–306.
- \*BARTO, E. K. & RILLIG, M. C. (2010). Does herbivory really suppress mycorrhiza? A meta-analysis. *Journal of Ecology* **98**, 745–753.
- \*BARTON, K. E. & KORICHEVA, J. (2010). The ontogeny of plant defense and herbivory: characterizing general patterns using meta-analysis. *The American Naturalist* **175**, 481–493.
- \*BATARY, P., BALDI, A., KLEIJN, D. & TSCHARNTKE, T. (2011). Landscape-moderated biodiversity effects of Agri-environmental management: a meta-analysis. *Proceedings of the Royal Society B: Biological Sciences* **278**, 1894–1902.
- \*BATEMAN, D. C. & BISHOP, M. J. (2017). The environmental context and traits of habitat-forming bivalves influence the magnitude of their ecosystem engineering. *Marine Ecology Progress Series* **563**, 95–110.
- †BEGG, C. B. & MAZUMDAR, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics* **50**, 1088–1101.
- BELLER, E. M., GLASZIOU, P. P., ALTMAN, D. G., HOPEWELL, S., BASTIAN, H., CHALMERS, I., GÖTZSCHE, P. C., LASSERSON, T., TOVEY, D. & FOR THE PRISMA FOR ABSTRACTS GROUP (2013). PRISMA for abstracts: reporting systematic reviews in journal and conference abstracts. *PLoS Medicine* **10**, e1001419.
- BERGER-TAL, O., GREGGOR, A. L., MACURA, B., ADAMS, C. A., BLUMENTHAL, A., BOUSKILA, A., CANDOLIN, U., DORAN, C., FERNÁNDEZ-JURICIC, E., GOTANDA, K. M., PRICE, C., PUTMAN, B. J., SEGOLI, M., SNIJDERS, L., WONG, B. B. M., et al. (2019). Systematic reviews and maps as tools for applying behavioral ecology to management and policy. *Behavioral Ecology* **30**, 1–8.
- BORRIES, C., SANDEL, A. A., KOENIG, A., FERNÁNDEZ-DUQUE, E., KAMILAR, J. M., AMOROSO, C. R., BARTON, R. A., BRAY, J., DI FIORE, A., GILBY, I. C., GORDON, A. D., MUNDY, R., PORT, M., POWELL, L. E., PUSEY, A. E., et al. (2016). Transparency, usability, and reproducibility: guiding principles for improving comparative databases using primates as examples. *Evolutionary Anthropology: Issues, News, and Reviews* **25**, 232–238.
- \*BOURDEAU, P. E., BUTLIN, R. K., BRÖNNMARK, C., EDGELL, T. C., HOVERMAN, J. T. & HOLLANDER, J. (2015). What can aquatic gastropods tell us about phenotypic plasticity? A review and meta-analysis. *Heredity* **115**, 312–321.
- BRAMER, W. M., DE JONGE, G. B., RETHLEFSEN, M. L., MAST, F. & KLEIJNEN, J. (2018). A systematic approach to searching: an efficient and complete method to develop literature searches. *Journal of the Medical Library Association* **106**, 531–541.
- \*BREED, M. F., OTTEWELL, K. M., GARDNER, M. G., MARKLUND, M. H. K., DORMONTT, E. E. & LOWE, A. J. (2015). Mating patterns and pollinator mobility are critical traits in forest fragmentation genetics. *Heredity* **115**, 108–114.
- BÜTTNER, F., WINTERS, M., DELAHUNT, E., ELBERS, R., LURA, C. B., KHAN, K. M., WEIR, A. & ARDERN, C. L. (2020). Identifying the 'incredible'! Part 1: assessing the risk of bias in outcomes included in systematic reviews. *British Journal of Sports Medicine* **54**, 798–800.
- \*ČAPEK, P., MANZONI, S., KAŠTOVSKÁ, E., WILD, B., DIÁKOVÁ, K., BARTA, J., SCHNECKER, J., BIASI, C., MARTIKAINEN, P. J., ALVES, R. J. E., GUGGENBERGER, G., GENTSCH, N., HUGELIUS, G., PALMTAG, J., MIKUTTA, R., et al. (2018). A plant-microbe interaction framework explaining nutrient effects on primary production. *Nature Ecology & Evolution* **2**, 1588–1596.
- \*CARUSO, C. M., EISEN, K. E., MARTIN, R. A. & SLETVOLD, N. (2018). A meta-analysis of the agents of selection on floral traits. *Evolution* **73**, 4–14.
- \*CARUSO, C. M., MARTIN, R. A., SLETVOLD, N., MORRISSEY, M. B., WADE, M. J., AUGUSTINE, K. E., CARLSON, S. M., MACCOLL, A. D. C., SIEPIELSKI, A. M. & KINGSOLVER, J. G. (2017). What are the environmental determinants of phenotypic selection? A meta-analysis of experimental studies. *The American Naturalist* **190**, 363–376.
- CHAMBERLAIN, S. A., HOVICK, S. M., DIBBLE, C. J., RASMUSSEN, N. L., VAN ALLEN, B. G., MAITNER, B. S., AHERN, J. R., BELL-DERESKE, L. P., ROY, C. L., MEZA-LOPEZ, M., CARRILLO, J., SIEMANN, E., LAJEUNESSE, M. J. & WHITNEY, K. D. (2012). Does phylogeny matter? Assessing the impact of phylogenetic information in ecological meta-analysis. *Ecology Letters* **15**, 627–636.
- \*CHAPLIN-KRAMER, R., O'ROURKE, M. E., BLITZER, E. J. & KREMEN, C. (2011). A meta-analysis of crop pest and natural enemy response to landscape complexity. *Ecology Letters* **14**, 922–932.
- \*CONORD, C., GUREVITCH, J. & FADY, B. (2012). Large-scale longitudinal gradients of genetic diversity: a meta-analysis across six phyla in the Mediterranean basin. *Ecology and Evolution* **2**, 2600–2614.
- COOKE, S. J., BIRNIE-GAUVIN, K., LENNOX, R. J., TAYLOR, J. J., RYTWINSKI, T., RUMMER, J. L., FRANKLIN, C. E., BENNETT, J. R. & HADDADAWAY, N. R. (2017). How experimental biology and ecology can support evidence-based decision-making in conservation: avoiding pitfalls and enabling application. *Conservation Physiology* **5**, 623.
- COOPER, H. (2009). In *Hypotheses and Problems in Research Synthesis* (ed. H. COOPER), pp. 19–35. Russell Sage Foundation, New York.
- †COPAS, J. B. & LI, H. G. (1997). Inference for non-random samples. *Journal of the Royal Statistical Society Series B-Methodological* **59**, 55–95.
- COTE, I. M. & JENNIONS, M. (2013). The procedure of meta-analysis in a nutshell. In *Handbook of Meta-Analysis in Ecology and Evolution* (eds J. KORICHEVA, J. GUREVITCH and K. MENGENSEN), pp. 14–24. Princeton University Press, Princeton.
- \*CRANE, A. L. & FERRARI, M. C. O. (2017). Patterns of predator neophobia: a meta-analytic review. *Proceedings of the Royal Society B: Biological Sciences* **284**, 20170583.
- CURTIS, P. S., MENGENSEN, K., LAJEUNESSE, M. J., ROTHSTEIN, H. R. & STEWART, G. B. (2013). Extraction and critical appraisal of data. In *Handbook of Meta-Analysis in Ecology and Evolution* (eds J. KORICHEVA, J. GUREVITCH and K. MENGENSEN), pp. 52–60. Princeton University Press, Princeton.



- \*DAVIDSON, K. E., FOWLER, M. S., SKOV, M. W., DOERR, S. H., BEAUMONT, N. & GRIFFIN, J. N. (2017). Livestock grazing alters multiple ecosystem properties and services in salt marshes: a meta-analysis. *Journal of Applied Ecology* **54**, 1395–1405.
- \*DE BOECK, H. J., BLOOR, J. M. G., KREYLING, J., RANSIJN, J. C. G., NIJS, I., JENTSCH, A. & ZEITER, M. (2018). Patterns and drivers of biodiversity-stability relationships under climate extremes. *Journal of Ecology* **106**, 890–902.
- \*DE KORT, H., VANDEPITTE, K. & HONNAY, O. (2012). A meta-analysis of the effects of plant traits and geographical scale on the magnitude of adaptive differentiation as measured by the difference between QST and FST. *Evolutionary Ecology* **27**, 1081–1097.
- DENNIS, B., PONCIANO, J. M., TAPER, M. L. & LELE, S. R. (2019). Errors in statistical inference under model misspecification: evidence, hypothesis testing, and AIC. *Frontiers in Ecology and Evolution* **7**, 631.
- \*DES ROCHES, S., POST, D. M., TURLEY, N. E., BAILEY, J. K., HENDRY, A. P., KINNISON, M. T., SCHWEITZER, J. A. & PALKOVACS, E. P. (2018). The ecological importance of intraspecific variation. *Nature Ecology & Evolution* **2**, 57–64.
- \*DOBLER, R., ROGELL, B., BUDAR, F. & DOWLING, D. K. (2014). A meta-analysis of the strength and nature of cytoplasmic genetic effects. *Journal of Evolutionary Biology* **27**, 2021–2034.
- DOUGHERTY, L. R. & SHUKER, D. M. (2015). The effect of experimental design on the measurement of mate choice: a meta-analysis. *Behavioral Ecology* **26**, 311–319.
- \*DUNN, R. E., WHITE, C. R. & GREEN, J. A. (2018). A model to estimate seabird field metabolic rates. *Biology Letters* **14**, 20180190.
- \*DURRANT, H. M. S., BURRIDGE, C. P., KELAHER, B. P., BARRETT, N. S., EDGAR, G. J. & COLEMAN, M. A. (2013). Implications of macroalgal isolation by distance for networks of marine protected areas. *Conservation Biology* **28**, 438–445.
- †DUVAL, S. & TWEEDIE, R. (2000a). A nonparametric 'trim and fill' method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association* **95**, 89–98.
- †DUVAL, S. & TWEEDIE, R. (2000b). Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* **56**, 455–463.
- \*EDGE, C. B., HOULAHAN, J. E., JACKSON, D. A. & FORTIN, M.-J. (2016). The response of amphibian larvae to environmental change is both consistent and variable. *Oikos* **125**, 1700–1711.
- \*EDWARDS, K. F., AQUILINO, K. M., BEST, R. J., SELLHEIM, K. L. & STACHOWICZ, J. J. (2010). Prey diversity is associated with weaker consumer effects in a meta-analysis of benthic marine experiments. *Ecology Letters* **13**, 194–201.
- EGGER, M., DAVEY SMITH, G., SCHNEIDER, M. & MINDER, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* **315**, 629–634.
- ELLINGTON, E. H., BASTILLE-ROUSSEAU, G., AUSTIN, C., LANDOLT, K. N., POND, B. A., REES, E. E., ROBAR, N. & MURRAY, D. L. (2015). Using multiple imputation to estimate missing data in meta-regression. *Methods in Ecology and Evolution* **6**, 153–163.
- \*ENGLISH, S. & ULLER, T. (2016). Does early-life diet affect longevity? A meta-analysis across experimental studies. *Biology Letters* **12**, 20160291.
- EVERS, J. L. H. (2017). The Texas sharpshooter fallacy. *Human Reproduction* **32**, 1363–1363.
- EWERS, R. M., BARLOW, J., BANKS-LEITE, C. & RAHBK, C. (2019). Separate authorship categories to recognize data collectors and code developers. *Nature Ecology & Evolution* **3**, 1610–1610.
- \*EZIZ, A., YAN, Z., TIAN, D., HAN, W., TANG, Z. & FANG, J. (2017). Drought effect on plant biomass allocation: a meta-analysis. *Ecology and Evolution* **7**, 11002–11010.
- \*FELKER-QUINN, E., SCHWEITZER, J. A. & BAILEY, J. K. (2013). Meta-analysis reveals evolution in invasive plant species but little support for evolution of increased competitive ability (EICA). *Ecology and Evolution* **3**, 739–751.
- FORERO, D. A., LOPEZ-LEON, S., GONZÁLEZ-GIRALDO, Y. & BAGOS, P. G. (2019). Ten simple rules for carrying out and writing meta-analyses. *PLoS Computational Biology* **15**, e1006922-7.
- FORSTMEIER, W. & SCHIELZETH, H. (2011). Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology* **65**, 47–55.
- FORSTMEIER, W., WAGENMAKERS, E.-J. & PARKER, T. H. (2017). Detecting and avoiding likely false-positive findings - a practical guide. *Biological Reviews* **92**, 1941–1968.
- \*FRASER, D. J., WEIR, L. K., BERNATCHEZ, L., HANSEN, M. M. & TAYLOR, E. B. (2011). Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. *Heredity* **106**, 404–420.
- FRASER, H., PARKER, T., NAKAGAWA, S., BARNETT, A. & FIDLER, F. (2018). Questionable research practices in ecology and evolution. *PLoS One* **13**, e0200303.
- \*FUZESSY, L. F., CORNELISSEN, T. G., JANSON, C. & SILVEIRA, F. A. O. (2016). How do primates affect seed germination? A meta-analysis of gut passage effects on neotropical plants. *Oikos* **125**, 1069–1080.
- \*GARAMSZEGI, L. Z., MARKÓ, G. & HERCZEG, G. (2012). A meta-analysis of correlated behaviours with implications for behavioural syndromes: mean effect size, publication bias, phylogenetic effects and the role of mediator variables. *Evolutionary Ecology* **26**, 1213–1235.
- GARNER, P., HOPEWELL, S., CHANDLER, J., MACLEHOSE, H., AKI, E. A., BEYENE, J., CHANG, S., CHURCHILL, R., DEARNESS, K., GUYATT, G., LEFEBVRE, C., LILES, B., MARSHALL, R., GARCIA, L. M., MAVERGAMES, C., et al. (2016). When and how to update systematic reviews: consensus and checklist. *British Medical Journal* **354**, i3507.
- GATES, S. (2002). Review of methodology of quantitative reviews using meta-analysis in ecology. *Journal of Animal Ecology* **71**, 547–557.
- GELMAN, A. (2018). Ethics in statistical practice and communication: five recommendations. *Significance* **15**, 40–43.
- GELMAN, A. & CARLIN, J. (2014). Beyond power calculations: assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science* **9**, 641–651.
- GERSTNER, K., MORENO-MATEOS, D., GUREVITCH, J., BECKMANN, M., KAMBACH, S., JONES, H. P. & SEPPELT, R. (2017). Will your paper be used in a meta-analysis? Make the reach of your research broader and longer lasting. *Methods in Ecology and Evolution* **8**, 777–784.
- GLESER, L. J. & OLKIN, I. (2009). Stochastically dependent effect sizes. In *The Handbook of Research Synthesis and Meta-Analysis* (eds H. COOPER, L. V. HEDGES and J. C. VALENTINE), pp. 357–376. Russell Sage Foundation, New York.
- GOLDACRE, B. (2013). *Bad Pharma: How Medicine Is Broken, and How We Can Fix It*. HarperCollins Publishers, London.
- GOLDACRE, B., MORTON, C. E. & DEVITO, N. J. (2019). Why researchers should share their analytic code. *British Medical Journal* **367**, i6365.
- \*GORNÉ, L. D. & DÍAZ, S. (2017). A novel meta-analytical approach to improve systematic review of rates and patterns of microevolution. *Ecology and Evolution* **7**, 5821–5832.
- \*GREEN, D. M. (2019). Rarity of size-assortative mating in animals: assessing the evidence with anuran amphibians. *The American Naturalist* **193**, 279–295.
- \*GREENFIELD, M. D., ALEM, S., LIMOUSIN, D. & BAILEY, N. W. (2014). The dilemma of Fisherian sexual selection: mate choice for indirect benefits despite rarity and overall weakness of trait-preference genetic correlation. *Evolution* **68**, 3524–3536.
- \*GREENWOOD, S., RUIZ-BENITO, P., MARTÍNEZ-VILALTA, J., LLORET, F., KITZBERGER, T., ALLEN, C. D., FENSHAM, R., LAUGHLIN, D. C., KATTGE, J., BÖNISCH, G., KRAFT, N. J. B. & JUMP, A. S. (2017). Tree mortality across biomes is promoted by drought intensity, lower wood density and higher specific leaf area. *Ecology Letters* **20**, 539–553.
- GREGGOR, A. L., PRICE, C. J. & SHIER, D. M. (2019). Examining the efficacy of anti-predator training for increasing survival in conservation translocations: a systematic review protocol. *Environmental Evidence* **8**, 1–9.
- GUREVITCH, J. & HEDGES, L. V. (1999). Statistical issues in ecological meta-analyses. *Ecology* **80**, 1142–1149.
- GUREVITCH, J., KORICHEVA, J., NAKAGAWA, S. & STEWART, G. (2018). Meta-analysis and the science of research synthesis. *Nature* **555**, 175–182.
- GUYATT, G. H., OXMAN, A. D., VIST, G. E., KUNZ, R., FALCK-YTTER, Y., ALONSO-COELLO, P. & SCHÜNEMANN, H. J. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *British Medical Journal* **336**, 924–926.
- HADDADAWAY, N. R., FEIERMAN, A., GRAINGER, M. J., GRAY, C. T., TANRIVERAYDER, E., DHAUBANJAR, S. & WESTGATE, M. J. (2019). EviAtlas: a tool for visualising evidence synthesis databases. *Environmental Evidence* **8**, 258–210.
- HADDADAWAY, N. R. & MACURA, B. (2018). The role of reporting standards in producing robust literature reviews. *Nature Climate Change* **8**, 444–447.
- HADDADAWAY, N. R., MACURA, B., WHALEY, P. & PULLIN, A. S. (2018). ROSES RepOrting standards for systematic evidence syntheses: pro forma, flow-diagram and descriptive summary of the plan and conduct of environmental systematic reviews and systematic maps. *Environmental Evidence* **7**, 7. <https://doi.org/10.1186/s13750-018-0121-7>.
- HADDADAWAY, N. R. & VERHOEVEN, J. T. A. (2015). Poor methodological detail precludes experimental repeatability and hampers synthesis in ecology. *Ecology and Evolution* **5**, 4451–4454.
- HADFIELD, J. D. & NAKAGAWA, S. (2010). General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *Journal of Evolutionary Biology* **23**, 494–508.
- \*HALUPKA, L. & HALUPKA, K. (2017). The effect of climate change on the duration of avian breeding seasons: a meta-analysis. *Proceedings of the Royal Society B: Biological Sciences* **284**, 20171710.
- HARMON, L. (2018). *Phylogenetic Comparative Methods: Learning from Trees*. CreateSpace Independent Publishing Platform, Scotts Valley.
- HARRISON, F. (2010). Getting started with meta-analysis. *Methods in Ecology and Evolution* **2**, 1–10.
- HARRISON, X. A., DONALDSON, L., CORREA-CANO, M. E., EVANS, J., FISHER, D. N., GOODWIN, C. E. D., ROBINSON, B. S., HODGSON, D. J. & INGER, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ* **6**, e4794.
- \*HARVEY, B. P., GWYNN-JONES, D. & MOORE, P. J. (2013). Meta-analysis reveals complex marine biological responses to the interactive effects of ocean acidification and warming. *Ecology and Evolution* **3**, 1016–1030.

- HEAD, M. L., HOLMAN, L., LANFEAR, R., KAHN, A. T. & JENNIONS, M. D. (2015). The extent and consequences of P-hacking in science. *PLoS Biology* **13**, e1002106.
- HEARD, S. B. (2016). *The Scientist's Guide to Writing*. Princeton University Press, Princeton.
- \*HECTOR, K. L., LAGISZ, M. & NAKAGAWA, S. (2012). The effect of resveratrol on longevity across species: a meta-analysis. *Biology Letters* **8**, 790–793.
- †HEDGES, L. V. (1984). Estimation of effect size under nonrandom sampling: the effects of censoring studies yielding statistically insignificant mean differences. *Journal of Educational Statistics* **9**, 61–85.
- HEDGES, L. V. & PIGOTT, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods* **9**, 426–445.
- HEDGES, L. V., TIPTON, E. & JOHNSON, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods* **1**, 39–65.
- HENMI, M. & COPAS, J. B. (2010). Confidence intervals for random effects meta-analysis and robustness to publication bias. *Statistics in Medicine* **29**, 2969–2983.
- HERBISON, P., HAY-SMITH, J. & GILLESPIE, W. J. (2006). Adjustment of meta-analyses on the basis of quality scores should be abandoned. *Journal of Clinical Epidemiology* **59**, 1249.e1–1249.e11.
- HIGGINS, J. P. T., THOMAS, J., CHANDLER, J., CUMPSTON, M., LI, T., PAGE, M. J. & WELCH, V. A. (2019). *Cochrane Handbook for Systematic Reviews of Interventions*, 2nd Edition. John Wiley & Sons, Chichester.
- HILLEBRAND, H. & GUREVITCH, J. (2013). Reporting standards in experimental studies. *Ecology Letters* **16**, 1419–1420.
- HOLCOMBE, A. O. (2019). Contributorship, not authorship: use CRediT to indicate who did what. *Publications* **7**(3), 48.
- HOLMAN, L., HEAD, M. L., LANFEAR, R. & JENNIONS, M. D. (2015). Evidence of experimental bias in the life sciences: why we need blind data recording. *PLoS Biology* **13**, e1002190.
- HUTTON, B., SALANTI, G., CALDWELL, D. M., CHAIMANI, A., SCHMID, C. H., CAMERON, C., IOANNIDIS, J. P. A., STRAUS, S., THORLUND, K., JANSEN, J. P., MULROW, C., CATALÁ-LÓPEZ, F., GÖTZSCHE, P. C., DICKERSIN, K., BOUTRON, I., et al. (2015). The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Annals of Internal Medicine* **162**, 777–772.
- INT'HOUT, J., IOANNIDIS, J. P. A., ROVERS, M. M. & GOEMAN, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open* **6**, e010247.
- IOANNIDIS, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine* **2**, e124.
- IOANNIDIS, J. P. A. (2016). The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *The Milbank Quarterly* **94**, 485–514.
- †IYENGAR, S. & GREENHOUSE, J. B. (1988). Selection models and the file drawer problem. *Statistical Science* **3**, 109–117.
- \*JACTEL, H., GRITTI, E. S., DRÖSSLER, L., FORRESTER, D. I., MASON, W. L., MORIN, X., PRETZSCH, H. & CASTAGNEYROL, B. (2018). Positive biodiversity–productivity relationships in forests: climate matters. *Biology Letters* **14**, 20170747. <https://doi.org/10.1098/rsbl.2017.0747>.
- JAMES, K. L., RANDALL, N. P. & HADDADAWAY, N. R. (2016). A methodology for systematic mapping in environmental sciences. *Environmental Evidence* **1**–13.
- JARVINEN, A. (1991). A meta-analytic study of the effects of female age on laying-date and clutch-size in the great tit *Parus major* and the pied flycatcher *Ficedula hypoleuca*. *Ibis* **133**, 62–67.
- JENNIONS, M. D., LORTIE, C. J., ROSENBERG, M. S. & ROTHSTEIN, H. R. (2013). Publication and related biases. In *Handbook of Meta-Analysis in Ecology and Evolution* (eds J. KORICHEVA, J. GUREVITCH and K. MENGENSEN), pp. 207–236. Princeton University Press, Princeton.
- JENNIONS, M. D. & MØLLER, A. P. (2002). Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution. *Proceedings of the Royal Society B: Biological Sciences* **269**, 43–48.
- \*JESSOP, T. S., LANE, M. L., TEASDALE, L., STUART-FOX, D., WILSON, R. S., CAREAU, V. & MOORE, I. T. (2016). Multiscale evaluation of thermal dependence in the glucocorticoid response of vertebrates. *The American Naturalist* **188**, 342–356.
- \*JIANG, Y., BOLNICK, D. I. & KIRKPATRICK, M. (2013). Assortative mating in animals. *The American Naturalist* **181**, E125–E138.
- \*KAMIYA, T., O'DWYER, K., WESTERDAHL, H., SENIOR, A. & NAKAGAWA, S. (2014). A quantitative review of MHC-based mating preference: the role of diversity and dissimilarity. *Molecular Ecology* **23**, 5151–5163.
- \*KARBAN, R., YANG, L. H. & EDWARDS, K. F. (2013). Volatile communication between plants that affects herbivory: a meta-analysis. *Ecology Letters* **17**, 44–52.
- KERR, N. L. (1998). HARKing: hypothesizing after the results are known. *Personality and Social Psychology Review* **2**, 196–217.
- †KETTERING, K. M. & ADAMS, C. R. (2011). Lessons learned from invasive plant control experiments: a systematic review and meta-analysis: invasive plant control experiments. *Journal of Applied Ecology* **48**, 970–979.
- \*KINLOCK, N. L., PROWANT, L., HERSTOFF, E. M., FOLEY, C. M., AKIN-FAJIYE, M., BENDER, N., UMARANI, M., RYU, H. Y., ŞEN, B. & GUREVITCH, J. (2018). Explaining global variation in the latitudinal diversity gradient: meta-analysis confirms known patterns and uncovers new ones. *Global Ecology and Biogeography* **27**, 125–141.
- †KORICHEVA, J., GUREVITCH, J. & MENGENSEN, K. (2013). *Handbook of Meta-Analysis in Ecology and Evolution*. Princeton University Press, Princeton.
- KORICHEVA, J. & GUREVITCH, J. (2014). Uses and misuses of meta-analysis in plant ecology. *Journal of Ecology* **102**, 828–844.
- KORICHEVA, J. & KULINSKAYA, E. (2019). Temporal instability of evidence base: a threat to policy making? *Trends in Ecology & Evolution* **34**, 895–902.
- \*KORICHEVA, J. & ROMERO, G. Q. (2012). You get what you pay for: reward-specific trade-offs among direct and ant-mediated defences in plants. *Biology Letters* **8**, 628–630.
- \*KORNER, N. A., RIEGL, B. M. & FIGUEIREDO, J. (2018). Thresholds and drivers of coral calcification responses to climate change. *Global Change Biology* **24**, 5084–5095.
- KUEFFER, C., NIINEMETS, Ü., DRENOVSKY, R. E., KATTGE, J., MILBERG, P., POORTER, H., REICH, P. B., WERNER, C., WESTOBY, M. & WRIGHT, I. J. (2011). Fame, glory and neglect in meta-analyses. *Trends in Ecology & Evolution* **26**, 493–494.
- LAGISZ, M., ZIDAR, J., NAKAGAWA, S., NEVILLE, V., SORATO, E., PAUL, E. S., BATESON, M., MENDL, M. & LØVLIE, H. (2020). Optimism, pessimism and judgement bias in animals: a systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews* **118**, 3–17.
- LAJEUNESSE, M. J. (2009). Meta-analysis and the comparative phylogenetic method. *The American Naturalist* **174**, 369–381.
- LAJEUNESSE, M. J. (2013). Recovering missing or partial data from studies: a survey of conversions and imputations for meta-analysis. In *Handbook of Meta-Analysis in Ecology and Evolution* (eds J. KORICHEVA, J. GUREVITCH and K. MENGENSEN), pp. 195–206. Princeton University Press, Princeton.
- LAJEUNESSE, M. J. (2016). Facilitating systematic reviews, data extraction and meta-analysis with the metagear package for R. *Methods in Ecology and Evolution* **7**, 323–330.
- LANE, P. W., ANZURES-CABRERA, J., LEWIS, S. & TOMLINSON, J. (2012). Graphics for meta-analysis. In *A Picture Is Worth a Thousand Tables* (eds A. KRAUSE and M. O'CONNELL), pp. 295–308. Springer US, Boston.
- \*LEHMANN, A., ZHENG, W. & RILLIG, M. C. (2017). Soil biota contributions to soil aggregation. *Nature Ecology & Evolution* **1**, 1828–1835.
- LEIMU, R. & KORICHEVA, J. (2004). Cumulative meta-analysis: a new tool for detection of temporal trends and publication bias in ecology. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **271**, 1961–1966.
- \*LI, F.-R., PENG, S.-L., CHEN, B.-M. & HOU, Y.-P. (2010). A meta-analysis of the responses of woody and herbaceous plants to elevated ultraviolet-B radiation. *Acta Oecologica* **36**, 1–9.
- LIBERATI, A., ALTMAN, D. G., TETZLAFF, J., MULROW, C., GÖTZSCHE, P. C., IOANNIDIS, J. P. A., CLARKE, M., DEVEREAUX, P. J., KLEIJNEN, J. & MOHER, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *British Medical Journal* **339**, b2700.
- LITTILL, J. A., CORCORAN, J. & PILLAI, V. (2008). *Systematic Reviews and Meta-Analysis*. Oxford University Press, New York.
- LITTLE, R. J. A. & RUBIN, D. B. (2020). *Statistical Analysis with Missing Data*, 3rd Edition. Wiley, Hoboken.
- LÓPEZ-LÓPEZ, J. A., PAGE, M. J., LIPSEY, M. W. & HIGGINS, J. P. T. (2018). Dealing with effect size multiplicity in systematic reviews and meta-analyses. *Research Synthesis Methods* **68**, 1282.
- \*LOYDI, A., ECKSTEIN, R. L., OTTE, A. & DONATH, T. W. (2012). Effects of litter on seedling establishment in natural and semi-natural grasslands: a meta-analysis. *Journal of Ecology* **101**, 454–464.
- \*MA, C., LI, S., PU, Z., TAN, J., LIU, M., ZHOU, J., LI, H. & JIANG, L. (2016). Different effects of invader–native phylogenetic relatedness on invasion success and impact: a meta-analysis of Darwin's naturalization hypothesis. *Proceedings of the Royal Society B: Biological Sciences* **283**, 20160663.
- \*MA, Z. & CHEN, H. Y. H. (2016). Effects of species diversity on fine root productivity in diverse ecosystems: a global meta-analysis. *Global Ecology and Biogeography* **25**, 1387–1396.
- MANN, T. (2015). *The Oxford Guide to Library Research*, 4th Edition. Oxford University Press, Oxford.
- \*MASSAD, T. J. (2013). Ontogenetic differences of herbivory on woody and herbaceous plants: a meta-analysis demonstrating unique effects of herbivory on the young and the old, the slow and the fast. *Oecologia* **172**, 1–10.
- \*MAZIA, N., MOYANO, J., PEREZ, L., AGUIAR, S., GARIBALDI, L. A. & SCHLICHTER, T. (2016). The sign and magnitude of tree–grass interaction along a global environmental gradient. *Global Ecology and Biogeography* **25**, 1510–1519.
- McGOWAN, J., SAMPSON, M., SALZWEDEL, D. M., COGO, E., FOERSTER, V. & LEFEBVRE, C. (2016). PRESS peer review of electronic search strategies: 2015 guideline statement. *Journal of Clinical Epidemiology* **75**, 40–46.

- \*MCLEAN, M. J., BISHOP, P. J. & NAKAGAWA, S. (2012). Male quality, signal reliability and female choice: assessing the expectations of inter-sexual selection. *Journal of Evolutionary Biology* **25**, 1513–1520.
- †MC SHANE, B. B., BÖCKENHOLT, U. & HANSEN, K. T. (2016). Adjusting for publication bias in meta-analysis: an evaluation of selection methods and some cautionary notes. *Perspectives on Psychological Science* **11**, 730–749.
- MENGERSEN, K. & GUREVITCH, J. (2013). Using other metrics of effect size in meta-analysis. In *Handbook of Meta-Analysis in Ecology and Evolution* (eds J. KORICHEVA, J. GUREVITCH and K. MENGERSEN), pp. 72–85. Princeton University Press, Princeton.
- \*MERKLING, T., NAKAGAWA, S., LAGISZ, M. & SCHWANZ, L. E. (2018). Maternal testosterone and offspring sex-ratio in birds and mammals: a meta-analysis. *Evolutionary Biology* **45**, 96–104.
- METEYARD, L. & DAVIES, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language* **112**, 104092.
- \*MILLER, S. E., BARRUETO, M. & SCHLUTER, D. (2017). A comparative analysis of experimental selection on the stickleback pelvis. *Journal of Evolutionary Biology* **30**, 1165–1176.
- \*MOATT, J. P., NAKAGAWA, S., LAGISZ, M. & WALLING, C. A. (2016). The effect of dietary restriction on reproduction: a meta-analytic perspective. *BMC Evolutionary Biology* **16**, 401.
- MOHER, D., BOUTER, L., KLEINERT, S., GLASZIOU, P., SHAM, M. H., BARBOUR, V., CORIAT, A.-M., FOEGER, N. & DIRNAGL, U. (2020). The Hong Kong principles for assessing researchers: fostering research integrity. *PLoS Biology* **18**, e3000737.
- †MOHER, D., COOK, D. J., EASTWOOD, S., OLKIN, I., RENNIE, D. & STROUP, D. F. (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet* **354**, 1896–1900.
- MOHER, D., LIBERATI, A., TETZLAFF, J., ALTMAN, D. G. & THE PRISMA GROUP (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine* **6**, e1000097.
- MOHER, D., NAUDET, F., CRISTEA, I. A., MIEDEMA, F., IOANNIDIS, J. P. A. & GOODMAN, S. N. (2018). Assessing scientists for hiring, promotion, and tenure. *PLoS Biology* **16**, e2004089.
- MOHER, D., SHAMSEER, L., CLARKE, M., GHERSI, D., LIBERATI, A., PETTICREW, M., SHEKELLE, P., STEWART, L. A., ALTMAN, D. G., BOOTH, A., CHAN, A. W., CHANG, S., CLIFFORD, T., DICKERSON, K., EGGER, M., et al. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews* **4**, e1000326.
- MÖLLER, A. P. & JENNIONS, M. D. (2001). Testing and adjusting for publication bias. *Trends in Ecology & Evolution* **16**, 580–586.
- \*MOORE, M. P., RIESCH, R. & MARTIN, R. A. (2016). The predictability and magnitude of life-history divergence to ecological agents of selection: a meta-analysis in livebearing fishes. *Ecology Letters* **19**, 435–442.
- MORENO, S. G., SUTTON, A. J., ADES, A., STANLEY, T. D., ABRAMS, K. R., PETERS, J. L. & COOPER, N. J. (2009). Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC Medical Research Methodology* **9**, 2.
- MORRISSEY, M. B. (2016). Meta-analysis of magnitudes, differences and variation in evolutionary parameters. *Journal of Evolutionary Biology* **29**, 1882–1904.
- MUKA, T., GLISIC, M., MILIC, J., VERHOOG, S., BOHLIUS, J., BRAMER, W., CHOWDHURY, R. & FRANCO, O. H. (2020). A 24-step guide on how to design, conduct, and successfully publish a systematic review and meta-analysis in medical research. *European Journal of Epidemiology* **35**, 49–60.
- MUNAFÒ, M. R., CHAMBERS, C. D., COLLINS, A. M., FORTUNATO, L. & MACLEOD, M. R. (2020). Research culture and reproducibility. *Trends in Cognitive Sciences* **24**, 91–93.
- \*MUNDIM, F. M. & BRUNA, E. M. (2016). Is there a temperate bias in our understanding of how climate change will alter plant-herbivore interactions? A meta-analysis of experimental studies. *The American Naturalist* **188**, S74–S89.
- \*MURPHY, G. E. P. & ROMANUK, T. N. (2012). A meta-analysis of community response predictability to anthropogenic disturbances. *The American Naturalist* **180**, 316–327.
- \*MURPHY, G. E. P. & ROMANUK, T. N. (2013). A meta-analysis of declines in local species richness from human disturbances. *Ecology and Evolution* **4**, 91–103.
- \*MURREN, C. J., MACLEAN, H. J., DIAMOND, S. E., STEINER, U. K., HESKEL, M. A., HANDELSMAN, C. A., GHALAMBOR, C. K., AULD, J. R., CALLAHAN, H. S., PFENNIG, D. W., RELYEA, R. A., SCHLICHTING, C. D. & KINGSOLVER, J. (2014). Evolutionary change in continuous reaction norms. *The American Naturalist* **183**, 453–467.
- MURTAUGH, P. A. (2002). Journal quality, effect size, and publication bias in meta-analysis. *Ecology* **83**, 1162–1166.
- NAKAGAWA, S. & DE VILLEMEREUIL, P. (2019). A general method for simultaneously accounting for phylogenetic and species sampling uncertainty via Rubin's rules in comparative analysis. *Systematic Biology* **68**, 632–641.
- NAKAGAWA, S. & FRECKLETON, R. P. (2008). Missing inaction: the dangers of ignoring missing data. *Trends in Ecology & Evolution* **23**, 592–596.
- NAKAGAWA, S. & LAGISZ, M. (2016). Visualizing unbiased and biased unweighted meta-analyses. *Journal of Evolutionary Biology* **29**, 1914–1916.
- NAKAGAWA, S. & LAGISZ, M. (2019). How good does our map of knowledge have to be?: a comment on Berger-Tal et al. *Behavioral Ecology* **30**, 13–14.
- NAKAGAWA, S., LAGISZ, M., O'DEA, R. E., RUTKOWSKA, J., YANG, Y., NOBLE, D. W. A. & SENIOR, A. M. (2020). The orchard plot: cultivating a forest plot for use in ecology, evolution, and beyond. *Research Synthesis Methods* **12**, 4–12.
- NAKAGAWA, S., NOBLE, D. W. A., SENIOR, A. M. & LAGISZ, M. (2017). Meta-evaluation of meta-analysis: ten appraisal questions for biologists. *BMC Biology* **15**, 3.
- NAKAGAWA, S., SAMARASINGHE, G., HADDADAWAY, N. R., WESTGATE, M. J., O'DEA, R. E., NOBLE, D. W. A. & LAGISZ, M. (2019). Research weaving: visualizing the future of research synthesis. *Trends in Ecology & Evolution* **34**, 224–238.
- NAKAGAWA, S. & SANTOS, E. S. A. (2012). Methodological issues and advances in biological meta-analysis. *Evolutionary Ecology* **26**, 1253–1274.
- NELSON, L. D., SIMMONS, J. & SIMONSOHN, U. (2018). Psychology's renaissance. *Annual Review of Psychology* **69**, 511–534.
- \*NIEBEL, P. T. & DINGEMANSE, N. J. (2018). Meta-analysis reveals weak associations between intrinsic state and personality. *Proceedings of the Royal Society B: Biological Sciences* **285**, 20172823.
- NOBLE, D. W. A., LAGISZ, M., O'DEA, R. E. & NAKAGAWA, S. (2017). Nonindependence and sensitivity analyses in ecological and evolutionary meta-analyses. *Molecular Ecology* **26**, 2410–2425.
- NOSEK, B. A., ALTER, G., BANKS, G. C., BORSBOOM, D., BOWMAN, S. D., BRECKLER, S. J., BUCK, S., CHAMBERS, C. D., CHIN, G., CHRISTENSEN, G., CONTESTABLE, M., DAFOE, A., EICH, E., FREESE, J., GLENNESTER, R., et al. (2015). Promoting an open research culture. *Science* **348**, 1422–1425.
- NOSEK, B. A., BECK, E. D., CAMPBELL, L., FLAKE, J. K., HARDWICKE, T. E., MELLOR, D. T., VAN T VEER & VAZIRE, S. (2019). Preregistration is hard, and worthwhile. *Trends in Cognitive Sciences* **23**, 815–818.
- NOSEK, B. A., EBERSOLE, C. R., DEHAVEN, A. C. & MELLOR, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 2600–2606.
- O'BOYLE, E. H. JR., BANKS, G. C. & GONZALEZ-MULÉ, E. (2017). The Chrysalis effect. *Journal of Management* **43**, 376–399.
- \*O'BRIEN, J. M. & SCHEIBLING, R. E. (2018). Turf wars: competition between foundation and turf-forming species on temperate and tropical reefs and its role in regime shifts. *Marine Ecology Progress Series* **590**, 1–17.
- †O'DEA, R. E., LAGISZ, M., HENDRY, A. P. & NAKAGAWA, S. (2019). Developmental temperature affects phenotypic means and variability: a meta-analysis of fish data. *Fish and Fisheries* **12**, 523–518.
- \*ORD, T. J., KING, L. & YOUNG, A. R. (2011). Contrasting theory with the empirical data of species recognition. *Evolution* **65**, 2572–2591.
- †ORWIN, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics* **8**, 157.
- OUZZANI, M., HAMMADY, H., FEDOROWICZ, Z. & ELMAGARMID, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews* **5**, 31.
- PAGE, M. J., MOHER, D., BOSSUYT, P. M., BOUTRON, I., HOFFMANN, T. C., MULROW, C. D., SHAMSEER, L., TETZLAFF, J. M., AKL, E. A., BRENNAN, S. E., CHOU, R., GLANVILLE, J., GRIMSHAW, J. M., HRÓBJARTSSON, A., LALU, M. M., et al. (2021a). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *British Medical Journal* **372**, n160.
- PAGE, M. J., MCKENZIE, J. E., BOSSUYT, P. M., BOUTRON, I., HOFFMANN, T. C., MULROW, C. D., SHAMSEER, L., TETZLAFF, J. M., AKL, E. A., BRENNAN, S. E., CHOU, R., GLANVILLE, J., GRIMSHAW, J. M., HRÓBJARTSSON, A., LALU, M. M., et al. (2021b). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *British Medical Journal* **372**, n71.
- PAGE, M. J. & MOHER, D. (2017). Evaluations of the uptake and impact of the preferred reporting items for systematic reviews and meta-analyses (Prisma). Statement and extensions: a scoping review. *Systematic Reviews* **6**, 263.
- PALPACUER, C., HAMMAS, K., DUPREZ, R., LAVIOLLE, B., IOANNIDIS, J. P. A. & NAUDET, F. (2019). Vibration of effects from diverse inclusion/exclusion criteria and analytical choices: 9216 different ways to perform an indirect comparison meta-analysis. *BMC Medicine* **17**, 1–13.
- PAKEKH-BHURKE, S., KWOK, C. S., PANG, C., HOOPER, L., LOKE, Y. K., RYDER, J. J., SUTTON, A. J., HING, C. B., HARVEY, I. & SONG, F. (2011). Uptake of methods to deal with publication bias in systematic reviews has increased over time, but there is still much scope for improvement. *Journal of Clinical Epidemiology* **64**, 349–357.
- PARKER, T., FRASER, H. & NAKAGAWA, S. (2019). Making conservation science more reliable with preregistration and registered reports. *Conservation Biology* **33**, 747–750.
- PARKER, T. H., FORSTMEIER, W., KORICHEVA, J., FIDLER, F., HADFIELD, J. D., CHEE, Y. E., KELLY, C. D. & NAKAGAWA, S. (2016a). Transparency in ecology and evolution: real problems, real solutions. *Trends in Ecology & Evolution* **31**, 711–719.

- PARKER, T. H., GREIG, E. I., NAKAGAWA, S., PARRA, M. & DALISIO, A. C. (2018a). Subspecies status and methods explain strength of response to local versus foreign song by oscine birds in meta-analysis. *Animal Behaviour* **142**, 1–17.
- PARKER, T. H., GRIFFITH, S. C., BRONSTEIN, J. L., FIDLER, F., FOSTER, S., FRASER, H., FORSTMEIER, W., GUREVITCH, J., KORICHEVA, J., SEPPELT, R., TINGLEY, M. W. & NAKAGAWA, S. (2018b). Empowering peer reviewers with a checklist to improve transparency. *Nature Ecology & Evolution* **2**, 929–935.
- †PARKER, T. H., NAKAGAWA, S., GUREVITCH, J. & IIEE (IMPROVING INFERENCE IN EVOLUTIONARY BIOLOGY AND ECOLOGY) WORKSHOP PARTICIPANTS (2016b). Promoting transparency in evolutionary biology and ecology. *Ecology Letters* **19**(7), 726–728.
- \*PEARSON, D. E., ORTEGA, Y. K., RUNYON, J. B. & BUTLER, J. L. (2016). Secondary invasion: the bane of weed management. *Biological Conservation* **197**, 8–17.
- \*PERE, C., JACTEL, H. & KENIS, M. (2013). Response of insect parasitism to elevation depends on host and parasitoid life-history strategies. *Biology Letters* **9**, 20130028.
- PICCOLO, S. R. & FRAMPTON, M. B. (2016). Tools and techniques for computational reproducibility. *GigaScience* **5**, 1226.
- PICK, J. L., NAKAGAWA, S. & NOBLE, D. W. A. (2019). Reproducible, flexible and high-throughput data extraction from primary literature: the metaDigitise R package. *Methods in Ecology and Evolution* **10**, 426–431.
- \*PIETSCH, K. A., OGLE, K., CORNELISSEN, J. H. C., CORNWELL, W. K., BÖNISCH, G., CRAINE, J. M., JACKSON, B. G., KATTGE, J., PELTZER, D. A., PENUELAS, J., REICH, P. B., WARDLE, D. A., WEEDON, J. T., WRIGHT, I. J., ZANNE, A. E., et al. (2014). Global relationship of wood and leaf litter decomposability: the role of functional traits within and across plant organs. *Global Ecology and Biogeography* **23**, 1046–1057.
- \*POMBI, M., KENGNE, P., GIMONNEAU, G., TENE-FOSSOG, B., AYALA, D., KAMDEM, C., SANTOLAMAZZA, F., GUELBEIGO, W. M., SAGON, N., PETRARCA, V., FONTENILLE, D., BESANSKY, N. J., ANTONIO-NKONDJIO, C., DABIRÉ, R. K., DELLA TORRE, A., et al. (2017). Dissecting functional components of reproductive isolation among closely related sympatric species of the *Anopheles gambiae* complex. *Evolutionary Applications* **10**, 1102–1120.
- \*PROKOP, Z. M., MICHALCZYK, L., DROBNAK, S. M., HERDEGEN, M. & RADWAN, J. (2012). Meta-analysis suggests choosy females get sexy sons more than “good genes”. *Evolution* **66**, 2665–2673.
- \*PRZESLAWSKI, R., BYRNE, M. & MELLIN, C. (2015). A review and meta-analysis of the effects of multiple abiotic stressors on marine embryos and larvae. *Global Change Biology* **21**, 2122–2140.
- PULLIN, A. S. & STEWART, G. B. (2006). Guidelines for systematic review in conservation and environmental management. *Conservation Biology* **20**, 1647–1656.
- PUSSEGODA, K., TURNER, L., GARRITTY, C., MAYHEW, A., SKIDMORE, B., STEVENS, A., BOUTRON, I., SARKIS-ONOFRE, R., BJERRE, L. M., HRÓBJARTSSON, A., ALTMAN, D. G. & MOHER, D. (2017). Systematic review adherence to methodological or reporting quality. *Systematic Reviews* **6**, 131.
- PYŠEK, P., RICHARDSON, D. M., PERGL, J., JAROŠÍK, V., SIXTOVÁ, Z. & WEBER, E. (2008). Geographical and taxonomic biases in invasion ecology. *Trends in Ecology & Evolution* **23**, 237–244.
- RATHBONE, J., CARTER, M., HOFFMANN, T. & GLASZIOU, P. (2015a). Better duplicate detection for systematic reviewers: evaluation of systematic review assistant-Deduplication module. *Systematic Reviews* **4**, 1–6.
- RATHBONE, J., HOFFMANN, T. & GLASZIOU, P. (2015b). Faster title and abstract screening? Evaluating Abstrackr, a semi-automated online screening program for systematic reviewers. *Systematic Reviews* **4**, 1.
- †R CORE TEAM. (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- RETHLESEN, M. L., KIRTLEY, S., WAFFENSCHMIDT, S., AYALA, A. P., MOHER, D., PAGE, M. J., KOFFEL, J. B. & PRISMA-S GROUP (2021). PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews* **10**, 39.
- RICE, D. B. & MOHER, D. (2019). Curtailing the use of preregistration: a misused term. *Perspectives on Psychological Science* **14**, 1105–1108.
- \*RIFKIN, J. L., NUNN, C. L. & GARAMSZEGI, L. Z. (2012). Do animals living in larger groups experience greater parasitism? A meta-analysis. *The American Naturalist* **180**, 70–82.
- ROCHE, D. G., KRUK, L. E. B., LANFEAR, R. & BINNING, S. A. (2015). Public data archiving in ecology and evolution: how well are we doing? *PLoS Biology* **13**, e1002295.
- \*ROFF, D. A. & FAIRBAIRN, D. J. (2012). A test of the hypothesis that correlational selection generates genetic correlations. *Evolution* **66**, 2953–2960.
- \*ROSA, R., RUMMER, J. L. & MUNDAY, P. L. (2017). Biological responses of sharks to ocean acidification. *Biology Letters* **13**, 20160796.
- †ROSENBERG, M. S. (2005). The file-drawer problem revisited: a general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution* **59**, 464–468.
- ROSENBERG, M. S. (2013). Moment and least-squares based approaches to meta-analytic inference. In *Handbook of Meta-Analysis in Ecology and Evolution* (eds J. KORICHEVA, J. GUREVITCH and K. MENGENSEN), pp. 108–124. Princeton University Press, Princeton.
- ROSENBERG, M. S., ROTHSTEIN, H. R. & GUREVITCH, J. (2013). Effect sizes: conventional choices and calculations. In *Handbook of Meta-Analysis in Ecology and Evolution* (eds J. KORICHEVA, J. GUREVITCH and K. MENGENSEN), pp. 61–71. Princeton University Press, Princeton.
- ROSENTHAL, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin* **86**, 638–641.
- ROSENTHAL, M. F., GERTLER, M., HAMILTON, A. D., PRASAD, S. & ANDRADE, M. C. B. (2017). Taxonomic bias in animal behaviour publications. *Animal Behaviour* **127**, 83–89.
- ROTHSTEIN, H. R., SUTTON, A. J. & BORENSTEIN, M. (2005). *Publication Bias in Meta-Analysis*. John Wiley & Sons, Ltd, Chichester.
- \*ROWIŃSKI, P. K. & ROGELL, B. (2017). Environmental stress correlates with increases in both genetic and residual variances: a meta-analysis of animal studies. *Evolution* **71**, 1339–1351.
- RUTKOWSKA, J., DUBIEC, A. & NAKAGAWA, S. (2013). All eggs are made equal: meta-analysis of egg sexual size dimorphism in birds. *Journal of Evolutionary Biology* **27**, 153–160.
- \*SAMIA, D. S. M., NOMURA, F. & BLUMSTEIN, D. T. (2013). Do animals generally flush early and avoid the rush? A meta-analysis. *Biology Letters* **9**, 20130016.
- \*SÁNCHEZ, C. A., BECKER, D. J., TEITELBAUM, C. S., BARRIGA, P., BROWN, L. M., MAJEWSKA, A. A., HALL, R. J. & ALTIZER, S. (2018). On the relationship between body condition and parasite infection in wildlife: a review and meta-analysis. *Ecology Letters* **21**, 1869–1884.
- SÁNCHEZ-TÓJAR, A., NAKAGAWA, S., SÁNCHEZ-FORTÚN, M., MARTIN, D. A., RAMANI, S., GIRNIT, A., BÓKONY, V., KEMPEAERS, B., LIKER, A., WESTNEAT, D. F., BURKE, T. & SCHROEDER, J. (2018). Meta-analysis challenges a textbook example of status signalling and demonstrates publication bias. *eLife* **7**, aac4716.
- SANDVE, G. K., NEKRUTENKO, A., TAYLOR, J. & HOVIG, E. (2013). Ten simple rules for reproducible computational research. *PLoS Computational Biology* **9**, e1003285.
- \*SEIBOLD, S., BÄSSLER, C., BRANDL, R., GOSSNER, M. M., THORN, S., ULYSHEN, M. D. & MÜLLER, J. (2015). Experimental studies of dead-wood biodiversity — a review identifying global gaps in knowledge. *Biological Conservation* **191**, 139–149.
- SENIOR, A. M., GRUEBER, C. E., KAMIYA, T., LAGISZ, M., O'DWYER, K., SANTOS, E. S. A. & NAKAGAWA, S. (2016). Heterogeneity in ecological and evolutionary meta-analyses: its magnitude and implications. *Ecology* **97**, 3293–3299.
- †SHAMSEER, L., MOHER, D., CLARKE, M., GHESI, D., LIBERATI, A., PETTICREW, M., SHEKELLE, P., STEWART, L. A. & THE PRISMA-P GROUP (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *British Medical Journal* **349**, g7647.
- \*SHANTZ, A. A., LEMOINE, N. P. & BURKEPILE, D. E. (2015). Nutrient loading alters the performance of key nutrient exchange mutualisms. *Ecology Letters* **19**, 20–28.
- SHOLLER, D., RAM, K., BOETTIGER, C. & KATZ, D. S. (2019). Enforcing public data archiving policies in academic publishing: a study of ecology journals. *Big Data & Society* **6**, 205395171983625.
- SIDDAWAY, A. P., WOOD, A. M. & HEDGES, L. V. (2019). How to do a systematic review: a best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology* **70**, 747–770.
- SIMMONS, J. P., NELSON, L. D. & SIMONSOHN, U. (2011). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science* **22**, 1359–1366.
- \*SINGH, A. & PUNZALAN, D. (2018). The strength of sex-specific selection in the wild. *Evolution* **72**, 2818–2824.
- \*SLATYER, R. A., HIRST, M. & SEXTON, J. P. (2013). Niche breadth predicts geographical range size: a general ecological pattern. *Ecology Letters* **16**, 1104–1114.
- \*SLOT, M. & KITAJIMA, K. (2014). General patterns of acclimation of leaf respiration to elevated temperatures across biomes and plant types. *Oecologia* **177**, 885–900.
- SMALDINO, P. E. & McELREATH, R. (2016). The natural selection of bad science. *Royal Society Open Science* **3**, 160384.
- \*SMITH, R. K., PULLIN, A. S., STEWART, G. B. & SUTHERLAND, W. J. (2010). Effectiveness of predator removal for enhancing bird populations. *Conservation Biology* **24**, 820–829.
- \*SMITH, R. K., PULLIN, A. S., STEWART, G. B. & SUTHERLAND, W. J. (2011). Is nest predator exclusion an effective strategy for enhancing bird populations? *Biological Conservation* **144**, 1–10.
- \*SORENSEN, G. H., DEY, C. J., MADLIGER, C. L. & LOVE, O. P. (2016). Effectiveness of baseline corticosterone as a monitoring tool for fitness: a meta-analysis in seabirds. *Oecologia* **183**, 353–365.
- STEWART, G. (2010). Meta-analysis in applied ecology. *Biology Letters* **6**, 78–81.
- STEWART, G. B., COTE, I. M., ROTHSTEIN, H. R. & CURTIS, P. S. (2013). First steps in beginning a meta-analysis. In *Handbook of Meta-Analysis in Ecology and Evolution* (eds J. KORICHEVA, J. GUREVITCH and K. MENGENSEN), pp. 27–36. Princeton University Press, Princeton.

- STEWART, L. A., CLARKE, M., ROVERS, M., RILEY, R. D., SIMMONDS, M., STEWART, G. & TIERNEY, J. F. (2015). Preferred reporting items for a systematic review and meta-analysis of individual participant data. *Journal of the American Medical Association* **313**, 1657–1659.
- STODDEN, V., SEILER, J. & MA, Z. (2018). An empirical analysis of journal policy effectiveness for computational reproducibility. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 2584–2589.
- SZOLLOSI, A., KELLEN, D., NAVARRO, D. J., SHIFFRIN, R., VAN ROOIJ, I., VAN ZANDT, T. & DONKIN, C. (2020). Is preregistration worthwhile? *Trends in Cognitive Sciences* **24**, 94–95.
- \*TEMME, A. A., CORNWELL, W. K., CORNELISSEN, J. H. C. & AERTS, R. (2013). Meta-analysis reveals profound responses of plant traits to glacial CO<sub>2</sub> levels. *Ecology and Evolution* **3**, 4525–4535.
- TERRIN, N., SCHMID, C. H., LAU, J. & OLKIN, I. (2003). Adjusting for publication bias in the presence of heterogeneity. *Statistics in Medicine* **22**, 2113–2126.
- \*THOMSEN, M. S., ALTIERI, A. H., ANGELINI, C., BISHOP, M. J., GRIBBEN, P. E., LEAR, G., HE, Q., SCHIEL, D. R., SILLIMAN, B. R., SOUTH, P. M., WATSON, D. M., WERNBERG, T. & ZOTZ, G. (2018). Secondary foundation species enhance biodiversity. *Nature Ecology & Evolution* **2**, 634–639.
- \*THURMAN, T. J. & BARRETT, R. D. H. (2016). The genetic consequences of selection in natural populations. *Molecular Ecology* **25**, 1429–1448.
- TIPTON, E., PUSTEJOVSKY, J. E. & AHMADI, H. (2019). A history of meta-regression: technical, conceptual, and practical developments between 1974 and 2018. *Research Synthesis Methods* **10**, 161–179.
- \*ULLER, T., NAKAGAWA, S. & ENGLISH, S. (2013). Weak evidence for anticipatory parental effects in plants and animals. *Journal of Evolutionary Biology* **26**, 2161–2170.
- VALENTINE, J. C., PIGOTT, T. D. & ROTHSTEIN, H. R. (2010). How many studies do you need?: a primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics* **35**, 215–247.
- †VAN AERT, R. C. M., WICHERTS, J. M. & VAN ASSEN, M. A. L. M. (2016). Conducting meta-analyses based on p values: reservations and recommendations for applying p-uniform and p-curve. *Perspectives on Psychological Science* **11**, 713–729.
- \*VAN KATWIJK, M. M., THORHAUG, A., MARBÀ, N., ORTH, R. J., DUARTE, C. M., KENDRICK, G. A., ALTHUIZEN, I. H. J., BALESTRI, E., BERNARD, G., CAMBRIDGE, M. L., CUNHA, A., DURANCE, C., GIESEN, W., HAN, Q., HOSOKAWA, S., et al. (2015). Global analysis of seagrass restoration: the importance of large-scale planting. *Journal of Applied Ecology* **53**, 567–578.
- VAN WILGENBURG, E. & ELGAR, M. A. (2013). Confirmation bias in studies of nestmate recognition: a cautionary note for research into the behaviour of animals. *PLoS One* **8**, e53548.
- \*VERESOGLOU, S. D. & RILLIG, M. C. (2012). Suppression of fungal and nematode plant pathogens through arbuscular mycorrhizal fungi. *Biology Letters* **8**, 214–217.
- VETTER, D., RÜCKER, G. & STORCH, I. (2013). Meta-analysis: a need for well-defined usage in ecology and conservation biology. *Ecosphere* **4**, 1–24.
- VEVEA, J. L., COBURN, K. & SUTTON, A. (2019). Publication bias. In *The Handbook of Research Synthesis and Meta-Analysis* (eds H. COOPER, L. V. HEDGES and J. C. VALENTINE), pp. 383–430. Russell Sage Foundation, New York.
- \*VOGWILL, T. & MACLEAN, R. C. (2014). The genetic basis of the fitness costs of antimicrobial resistance: a meta-analysis approach. *Evolutionary Applications* **8**, 284–295.
- \*VOJE, K. L. (2016). Scaling of morphological characters across trait type, sex, and environment: a meta-analysis of static allometries. *The American Naturalist* **187**, 89–98.
- \*VRANCKX, G., JACQUEMYN, H., MUYS, B. & HONNAY, O. (2011). Meta-analysis of susceptibility of woody plants to loss of genetic diversity through habitat fragmentation. *Conservation Biology* **26**, 228–237.
- \*WAN, J. S. H., PANG, C. K. & BONSER, S. P. (2017). Does the cost of adaptation to extremely stressful environments diminish over time? A literature synthesis on how plants adapt to heavy metals and pesticides. *Evolutionary Biology* **44**, 411–426.
- †WANG, M. C. & BUSHMAN, B. J. (1998). Using the normal quantile plot to explore meta-analytic data sets. *Psychological Methods* **3**, 46–54.
- \*WEHI, P. M., NAKAGAWA, S., TREWICK, S. A. & MORGAN-RICHARDS, M. (2011). Does predation result in adult sex ratio skew in a sexually dimorphic insect genus? *Journal of Evolutionary Biology* **24**, 2321–2328.
- †WELCH, V., PETTICREW, M., PETKOVIC, J., MOHER, D., WATERS, E., WHITE, H., TUGWELL, P., ATUN, R., AWASTHI, S., BARBOUR, V., BHUTTA, Z. A., CUERVO, L. G., GROVES, T., KOEHLMOOS-PEREZ, T., KRISTJANSSON, E., et al. (2016). Extending the PRISMA statement to equity-focused systematic reviews (PRISMA-E 2012): explanation and elaboration. *Journal of Clinical Epidemiology* **70**, 68–89.
- WESTGATE, M. J. (2019). Revtools: an R package to support article screening for evidence synthesis. *Research Synthesis Methods* **10**, 606–614.
- WHITTAKER, R. J. (2010). Meta-analyses and mega-mistakes: calling time on meta-analysis of the species richness–productivity relationship. *Ecology* **91**, 2522–2533.
- WILKINSON, M. D., DUMONTIER, M., AALBERSBERG, I. J., APPLETON, G., AXTON, M., BAAK, A., BLOMBERG, N., BOITEN, J.-W., DA SILVA SANTOS, L. B., BOURNE, P. E., BOUWMAN, J., BROOKES, A. J., CLARK, T., CROSAS, M., DILLO, I., et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data* **3**, 160018.
- WILLIS, B. H. & RILEY, R. D. (2017). Measuring the statistical validity of summary meta-analysis and meta-regression results for use in clinical practice. *Statistics in Medicine* **36**, 3283–3301.
- \*WOOD, J. L. A., YATES, M. C. & FRASER, D. J. (2016). Are heritability and selection related to population size in nature? Meta-analysis and conservation implications. *Evolutionary Applications* **9**, 640–657.
- \*XI, N., ZHANG, C. & BLOOR, J. M. G. (2017). Species richness alters spatial nutrient heterogeneity effects on above-ground plant biomass. *Biology Letters* **13**, 20170510.
- \*YATES, M. C. & FRASER, D. J. (2014). Does source population size affect performance in new environments? *Evolutionary Applications* **7**, 871–882.
- \*ZHANG, L. & LU, X. (2012). Amphibians live longer at higher altitudes but not at higher latitudes. *Biological Journal of the Linnean Society* **106**, 623–632.
- †ZORZELA, L., LOKE, Y. K., IOANNIDIS, J. P., GOLDER, S., SANTAGUIDA, P., ALTMAN, D. G., MOHER, D., VOHRA, S. & PRISMA HARMS GROUP (2016). PRISMA harms checklist: improving harms reporting in systematic reviews. *British Medical Journal* **352**, i157.
- \*ZVEREVA, E. L. & KOZLOV, M. V. (2011). Sources of variation in plant responses to belowground insect herbivory: a meta-analysis. *Oecologia* **169**, 441–452.
- \*ZVEREVA, E. L. & KOZLOV, M. V. (2014). Effects of herbivory on leaf life span in woody plants: a meta-analysis. *Journal of Ecology* **102**, 873–881.

## XVII. Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.  
**Appendix S1.** Supporting Information

(Received 2 November 2020; revised 5 April 2021; accepted 8 April 2021; published online 7 May 2021)