Received: 9 April 2020 | Revised: 3 February 2021 | Accepted: 5 February 2021
DOI: 10.1111/fwb.13695

REVIEW

Freshwater Biology WILEY

Quality of meta-analyses in freshwater ecology: A systematic review

Sara Lodi[1] | **Bruno S. Godoy**[2] | **Jean C. G. Ortega**[3] | **Luis M. Bini**[4]

[1]Programa de Pós-Graduação em Ecologia e Evolução, Universidade Federal de Goiás, Goiânia, Brazil
[2]Núcleo de Ciências Agrárias e Desenvolvimento Rural, Universidade Federal do Pará, Belém, Brazil
[3]Programa de Pós-Graduação em Ecologia e Manejo de Recursos Naturais, Universidade Federal do Acre, Rio Branco, Brazil
[4]Departamento de Ecologia, Universidade Federal de Goiás, Goiânia, Brazil

Correspondence
Sara Lodi, Programa de Pós-Graduação em Ecologia e Evolução, Universidade Federal de Goiás, Goiânia, Brazil.
Email: lodi.sara@gmail.com

Funding information
Fundação de Amparo à Pesquisa do Estado de Goiás; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; Instituto de Ciência e Tecnologia de Ecologia Evolução Conservação da Biodiversidade.; Conselho Nacional de Desenvolvimento Científico e Tecnológico

Abstract

1. Given the increasing use of systematic reviews and meta-analyses in ecology, their protocols should be closely followed to ensure quality. Several checklists are available to guide researchers towards a high-quality meta-analytic study. Freshwater ecology studies have a tradition of using experimental studies, which provide the ideal data to test hypotheses using meta-analysis.
2. Here, we evaluated the quality of 114 meta-analyses in freshwater ecology and 86 meta-analyses in ecology and evolution for comparative purposes.
3. We found that many studies are still using the term meta-analysis incorrectly and that this error persisted over time. The quality of the studies that did conduct a formal meta-analysis has improved. Thus, we speculate that available guidelines are being effective in improving the quality of meta-analytic studies. Quality was not associated with the impact factor of the journal where the meta-analyses were published or with the average number of citations.
4. In addition to the incorrect use of the term, we found that many studies failed to: report heterogeneity statistics, evaluate temporal changes in effect size, conduct publication bias analyses, address the collinearity among moderators, and provide the data. In general, meta-analyses in ecology and evolution have only a slightly better average score than meta-analyses in freshwater ecology.
5. Although the quality of meta-analyses in freshwater ecology has improved over time, there is much room for improvement. Authors should not label their studies as meta-analyses if these methods were not used. Compliance with checklists should be widely fostered as meta-analyses are increasingly being used to summarise findings in different areas of ecology. Authors, reviewers, and editors should use checklists to improve the quality of meta-analyses in freshwater ecology.

KEYWORDS
checklist, limnology, Prisma statement, quantitative review, systematization

1 | INTRODUCTION

Most review articles in ecology can be classified as narrative reviews, where a given research theme is qualitatively summarised in a narrative structure (Gates, 2002; Koricheva & Gurevitch, 2013). Narrative reviews are useful to explore perspectives, conceptual frameworks, and historical advances, and to provide a broader interpretation of topics (Collins & Fauser, 2005; Gurevitch et al., 2018). However, narrative reviews tend to be plagued with subjectivity and biases (e.g. by not stating the methods used to compile the primary studies reviewed; Koricheva & Gurevitch, 2013; Lortie, 2014). In recent decades, a growing number of studies have shown that systematic reviews and meta-analyses, if properly conducted, can minimise these issues (Gates, 2002; Gurevitch et al., 2018; Koricheva & Gurevitch, 2013; Lortie, 2014). These methods are especially important considering the need for predictive power and generalisation in the field of ecology (e.g. Houlahan et al., 2017).

Systematic reviews value transparency, systematic reporting, and reproducibility (Gurevitch et al., 2018). A meta-analysis may be conducted when enough data are available, going a step further by quantitatively summarising empirical findings (Gurevitch et al., 2018). In a nutshell, a meta-analysis calculates a weighted mean effect size and its statistical significance, and, importantly, explores possible causes of variation in effect sizes using moderators (Borenstein et al., 2009). Weighting effect sizes by their precision is a pivotal step (Borenstein et al., 2009), as it ensures that estimates will be more influenced by studies with high precision.

However, as emphasised by Olkin (1992), "doing a meta-analysis is easy, doing one well is hard" (see also Berman & Parker, 2002; Felson, 1992). The devil is in the details, and poorly conducted meta-analysis may do more harm than good. For example, results from meta-analyses, in the medical sciences, are generally regarded as providing stronger evidence than a single randomised control trial (but see Murad et al., 2016). However, if the meta-analysis is conducted poorly it may deliver misleading conclusions that may be believed as truthful and determine medical treatment regimens or health policies. Similarly, poorly conducted meta-analyses in ecology, in addition to delaying scientific progress, may lead to inappropriate conservation policies. For example, it is common that several primary studies in a meta-analysis contribute more than one effect size (e.g. when the same control is compared with two experimental groups or when the control group is compared with an experimental group at different time-points). Accordingly, a meta-analysis ignoring the fact that multiple effect sizes from the same primary studies are not independent would produce results with inflated type I error rates (see also Song et al., 2020; Mengersen et al., 2013; Van den Noortgate et al., 2014). In general, issues associated with lack of transparency (e.g. in selecting studies and reporting), misinterpretation of results (Morrissey, 2016), incorrect methodology (López-López et al., 2018), premature use of meta-analyses where there are insufficient data (Borenstein et al., 2009; Ioannidis, 2010), and lack of systematisation (Gurevitch et al., 2018), for example, have prompted critiques and challenged the credibility of meta-analytical results (Ioannidis, 2016; Morrissey, 2016; Whittaker, 2010). This shows a need to better disseminate good practices among different research fields and to improve the quality of meta-analysis studies (Gurevitch et al., 2018; Hillebrand & Cardinale, 2010).

Clear guidelines for conducting a systematic review have been published and updated as a strategy to improve the quality of meta-analysis (Fleming et al., 2014; Moher et al., 2009, 2015; Shamseer et al., 2015). Well-established guidelines for systematic reviews include, but are not limited to, the Cochrane Handbook of Systematic Reviews (Higgins & Green, 2011), PRISMA statement (Moher et al., 2009, 2015), Tools for Transparency in Ecology and Evolution (Parker et al., 2018), and other relevant checklists (Higgins et al., 2013; Nakagawa et al., 2017; see applications in Delaney et al., 2005; Willis & Quigley, 2011). In ecology, Koricheva and Gurevitch (2014) have proposed a checklist of quality criteria for meta-analysis for research synthesists, peer reviewers, and editors (see also Nakagawa et al., 2017). However, despite these calls for systematisation, misuses and poor reporting are still recurrent in meta-analyses in many fields of research (Gates, 2002; Koricheva & Gurevitch, 2014; Philibert et al., 2012; Roberts et al., 2006; Senior et al., 2016; Vetter et al., 2013).

Here, we performed a systematic review of meta-analyses published within the field of freshwater ecology and evaluated the quality of these studies considering established criteria. We used the checklist proposed by Koricheva and Gurevitch (2014) as the main reference to evaluate the quality of meta-analyses in freshwater ecology. For comparative purposes, we also evaluated the quality of meta-analyses in the field of ecology and evolution (Senior et al., 2016). We expected that our measure of quality (see Methods section) would increase over time due to increased knowledge of best practices by authors, reviewers, and editors. We also asked whether quality was associated with citation metrics to test whether better developed articles were published in journals with higher visibility. We also expected that the quality score would be correlated with the average number of citations, assuming that readers would seek well-developed studies as references. We expected that the statistical issues related to phylogenetic relationship would be the least addressed criterion due to the lack of recognition that species cannot be regarded as independent points in statistical analyses (Felsenstein, 1985). Also, phylogenetic data are not readily available for many biological groups in the freshwater realm. We expected that the assessment of temporal changes in effect sizes using, for example, cumulative meta-analysis, would not often be conducted. Despite being a well-established tool (e.g. Leimu & Koricheva, 2004), ecologists have only recently considered its potential and importance (e.g. Koricheva & Kulinskaya, 2019; Ortega et al., 2018). Methods to address the use of multiple outcomes or dependent effect sizes have only recently been published (e.g. Hedges et al., 2010; Van den Noortgate et al., 2014) and used in ecological studies (e.g. Stein et al., 2014). Thus, we predict that this quality criterion would also often be neglected.

2 | METHODS

2.1 | Search strategy and criteria

We searched for meta-analyses in freshwater ecology using the Web of Science (WoS) and Scopus databases (from 1990 to 8 August 2017). Our search string (see details in Table 1) resulted in 443 and 382 hits for WoS and Scopus, respectively. We removed

TABLE 1 Search string used to obtain meta-analytical studies on freshwater ecology published between 1991 to 2017 in the Web of Science and Scopus databases. The Boolean codes were adjusted to each search database, but the same words and wildcard functions were used in both

Terms	Search string
Meta-analysis terms	(meta-anal* OR metanal* OR "quantitative review")
Freshwater ecology terms	((freshwater OR aquatic* OR limnol* OR "inland water*" OR river* OR stream* OR creek* OR reservoir* OR lake* OR lagoon* OR pond* OR mere* OR loch* OR lakelet*) NOT (ocean* OR marine))

duplicates using the function *mergeDbSources* from the R package Bibliometrix (Aria & Cuccurullo, 2017) and screened the articles following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher et al., 2009, 2015). The PRISMA statement is a checklist, including a flow diagram, designed to help authors improve the reporting of systematic reviews (Moher et al., 2009, 2015). The first step consisted of reading the abstracts to exclude obviously irrelevant literature. Then, we read the papers for eligibility (full-text assessment). We included studies that contained formal meta-analyses and we excluded all reviews that were not quantitative reviews using meta-analysis procedures. Also, we only included studies with results from freshwater ecosystems and excluded studies that did not focus on freshwater organisms (Figure 1, Table S1).

A second reviewer re-evaluated 30% of the articles collected from the databases (174 out of 577 papers, selected using the function *sample*() in the software R; R Core Team, 2020). We then calculated the agreement rate (Orwin & Vevea, 2009) and Cohen's κ (Cohen, 1960) to assess the reliability of the screening procedures conducted in this study.

2.2 | Meta-analysis quality

We assigned scores to each study following the checklist of criteria proposed by Koricheva and Gurevitch (2014), which consisted of 16 criteria with references to support them (see their Table 3). We altered the criterion “use of a meta-regression to appraise the issue of exploring existent heterogeneity” to “exploring the heterogeneity among studies” (Table 2). In this case, the criterion was met whenever the authors explored causes of heterogeneity among studies using moderators irrespective of the approach (e.g. meta-regression or analyses). For each criterion met, we assigned one point to the study. Half points were assigned when the information was incomplete (e.g. lack of the specific search terms). To fulfil the third criterion, the studies had to have calculated

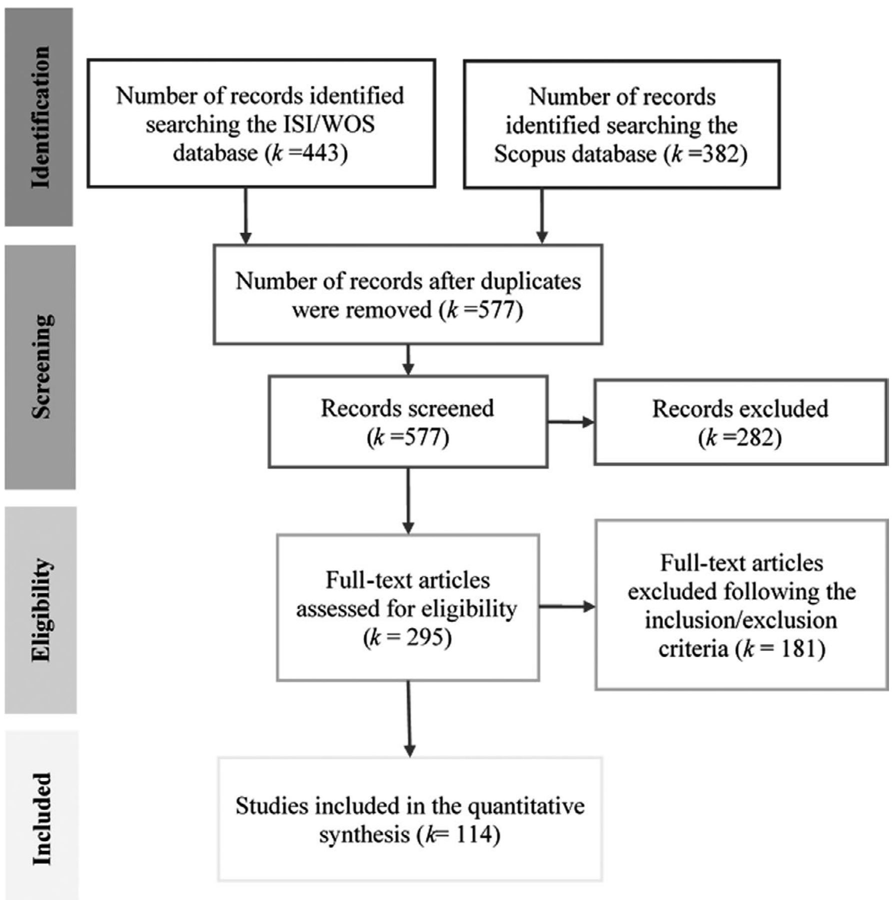


FIGURE 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram summarising study inclusion and exclusion phases (adapted from Moher et al., 2009)

TABLE 2 Quality criteria used to assign scores to the meta-analysis papers selected and observations explaining the criteria used and exceptions in punctuation. The quality criteria were obtained from Koricheva and Gurevitch (2014)

	Quality criteria	Observations
1	Are details of bibliographic search (electronic data bases used, keyword combinations, years) reported in sufficient detail to allow replication?	Half a point when the search terms are not provided. Not relevant for meta-analysis conducted with primary data.
2	Are criteria for study inclusion/exclusion explicitly listed?	Half a point when the criteria provided are not clear enough for reproducibility. Not relevant for meta-analysis conducted with primary data.
3	Have effect sizes been weighted by study precision or has the rationale for using unweighted approach been provided?	
4	Has the statistical model for meta-analysis been described?	
5	Has heterogeneity of effect sizes between studies been quantified?	
6	Have the causes of existent heterogeneity in effect sizes been explored?	
7	If effects of multiple moderators have been tested, have potential non-independence of and interactions between moderators been considered?	Not relevant if moderators were not considered
8	Have tests for publication bias been conducted?	
9	Have sensitivity analysis been performed to test the robustness of results?	
10	If meta-analysis combines studies published over considerable time span, have possible temporal changes in effect size been tested?	
11	If meta-analysis combined studies conducted on different species, has phylogenetic relatedness of species been considered?	Not relevant for studies considering single species or focused on community (e.g. diversity or species richness) or ecosystem responses
12	Has the software used been described?	
13	Have full bibliographic details of primary studies included in a meta-analysis been provided?	Not relevant for meta-analysis conducted with primary data.
14	Has the data set used for meta-analysis, including effect sizes and variances/sample sizes from individual primary studies and moderator variables, been provided as electronic appendix?	
15	Have standard metrics of effect size been used or, if nonstandard metrics have been employed, is the distribution of these parameters known and have the authors explained how they calculated variances for such metrics?	
16	If more than one estimate of effect size per study was included in the analysis, has potential non-independence of these estimates been considered?	Not relevant for studies that used one estimate of effect size per study

weighted effect sizes before modelling. We also assigned one point to the third criterion when effect sizes were weighted by the number of observations (see Gurevitch et al., 2018). Some studies only sought to explain the variation among studies (without estimating a weighted mean effect size). Consequently, we assumed that studies reporting variance partitioning results met the requirements for criterion 3. Some criteria were not relevant for some studies depending on their design (Table 2). For example, taking phylogenetic relatedness into account is not relevant for studies focused on a single species or ecosystem-level variables. Therefore, we calculated the ratio between the sum of points and the maximum possible number of points for the study. Accordingly, the maximum rating is no longer 16 as initially proposed by Koricheva and Gurevitch (2014) but ranged from 0 (lowest) to 1 (highest quality).

We tested for a temporal trend in the rating scores using the Spearman rank correlation between scores and publication year. Similarly, we tested for a correlation between scores and impact factor, and between scores and Normalised Citation Impact Index (NCII) using the Spearman rank correlation. The NCII is given by the ratio between the total number of citations and the time elapsed (in years) since publication (see Coursaris & Van Osch, 2014 and references therein).

We also rated the 86 meta-analyses in ecology and evolution compiled by Senior et al. (2016), to compare with our results. Senior et al. (2016) only included studies that quantified heterogeneity among effect sizes. Thus, for comparative purposes, we rated each of the 86 meta-analyses in ecology and evolution without considering the following criteria: “Has heterogeneity of effect sizes between studies been quantified?” (criterion 5), “Have the causes of existent heterogeneity in effect sizes been explored?" (criterion 6) and "Have standard metrics of effect size been used (…)" (criterion 15). We checked whether freshwater studies were included by Senior et al. (2016) and found that a total of 9 studies also evaluated the freshwater realm, none of which were included in our study, thus ensuring independence between datasets.

3 | RESULTS

Most of the 577 studies were excluded in abstract and full text screening because they did not conduct a formal meta-analysis (224 studies) or because they were not focused on freshwater systems (58 studies). The studies that were excluded during the analysis of the abstracts (282 papers) were clearly out of the objectives of our review. Specifically, 118 and 87 studies were not focused on freshwater systems and on ecological questions, respectively. Also, 76 studies did not conduct a formal meta-analysis and one was clearly methodological.

After reading the full texts, we excluded 147 and 34 studies for not being formal meta-analyses or focused on aquatic ecosystems, respectively. Most of these studies consisted of a re-analysis of primary data, without calculating an effect size per study. We found 114 meta-analytic studies that met our criteria (Figure 1) after following the different steps to select the studies. We found an agreement rate of 0.90 and a Cohen κ of 0.73 (95% CI = 0.62–0.84), indicating that most of the studies were consistently selected by both reviewers during the screening of the abstracts.

The number of meta-analytic studies focused on freshwater ecology increased over time, from the first study published by Wooster (1994), among the studies we selected, to 11 studies published in 2016 and eight in 2017 (until the search date; Figure 2a). These papers were published mostly in *Ecology*, *Oikos*, *Freshwater Biology*, and other specific journals within the field of freshwater ecology (Figure 2b).

We found an average rating score of 0.56 ± 0.20 (mean ± *SD*). Rating scores increased over time (Spearman correlation $r = 0.43$; $p < 0.001$) but were not related to journal impact factor ($r = 0.03$; $p = 0.720$) or to the NCII ($r = 0.15$; $p = 0.121$). Most studies (86.8%, out of 114) did not test for temporal changes in effect sizes (criterion 10). Also, most studies did not perform sensitivity analyses (criterion 9), evaluate the dependence among moderators (criterion 7) and did not provide heterogeneity statistics (criterion 5; Table 3). Nearly half of the studies did not check for publication bias (criterion 8) nor provided the data used for the meta-analysis (criterion 14). The issues of using multiple effect sizes per primary study (criterion 16) were also often ignored and approximately 36% of the studies did not provide details of the full literature search (criterion 1; Table 3). However, other important criteria were met more frequently; for example, most studies used standard metrics of effect size (criterion 15), explored the heterogeneity in effect sizes (criterion 6) and used the inverse–variance method to weight effect sizes (criterion 3; Table 3).

Most of the criteria were met in ecological and evolutionary studies (average rating score of 0.61 ± 0.21 *SD*) with a higher frequency than in limnological meta-analyses (Table 4). This was especially so for criteria 1, 3, 4, 8, 9, 11, 13, and 16 (Figure 3). In general, criteria 1, 7, 10, 11, and 16 were more frequently unmet (<50% of the studies) in both fields.

4 | DISCUSSION

Meta-analysis is considered a powerful method for summarising results of independent studies in different research areas (e.g. Lortie, 2014; Ellis, 2010; Gurevitch et al., 2019). However, the higher credibility of meta-analysis results may be a double-edged sword,

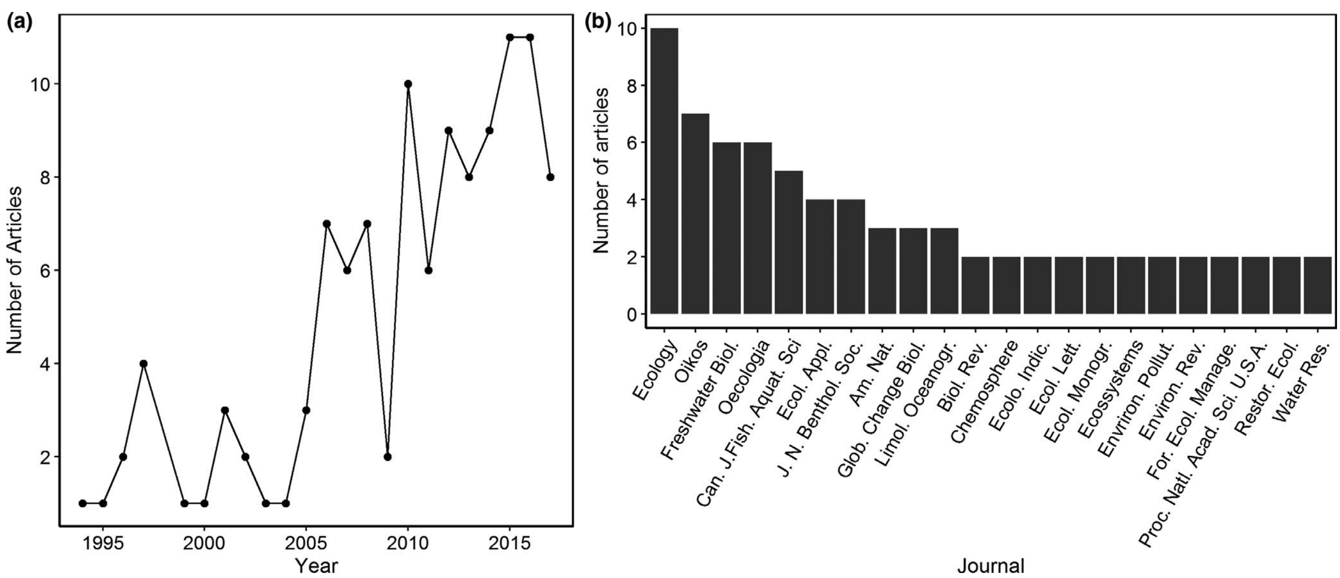


FIGURE 2 Annual number of studies on freshwater ecology using meta-analysis (a). Number of studies per journal (b). The dataset included a total of 114 studies published in 61 journals

	Quality criteria	NR (%)	Yes (%)	No (%)	Partial (%)
1	Searching details	7.02	28.95	35.96	28.07
2	Inclusion/exclusion criteria	7.02	57.02	19.3	16.67
3	Weighted effect sizes	–	75.44	24.56	–
4	Meta-analytical model	–	71.93	28.07	–
5	Heterogeneity in effect sizes	–	41.23	58.77	–
6	Causes of heterogeneity	0.88	82.46	16.67	–
7	Collinearity analysis	15.79	22.81	61.4	–
8	Publication bias	6.14	39.47	54.39	–
9	Sensitivity analysis	–	37.72	62.28	–
10	Changes in effect size	5.26	7.89	86.84	–
11	Controlling for phylogeny	71.05	3.51	25.44	–
12	Software	–	69.3	30.7	–
13	Bibliographic details	6.14	75.44	18.42	–
14	Data	–	52.63	47.37	–
15	Standard effect size	–	93.86	6.14	–
16	Multiple effect sizes	27.19	26.32	39.47	7.02

TABLE 3 Compliance with the criteria for reporting and following methodological standards assigned to freshwater ecology meta-analyses. This survey encompasses 114 studies published from 1991 to 2017. NR = not relevant for the study. A detailed description of the criteria is shown in Table 2

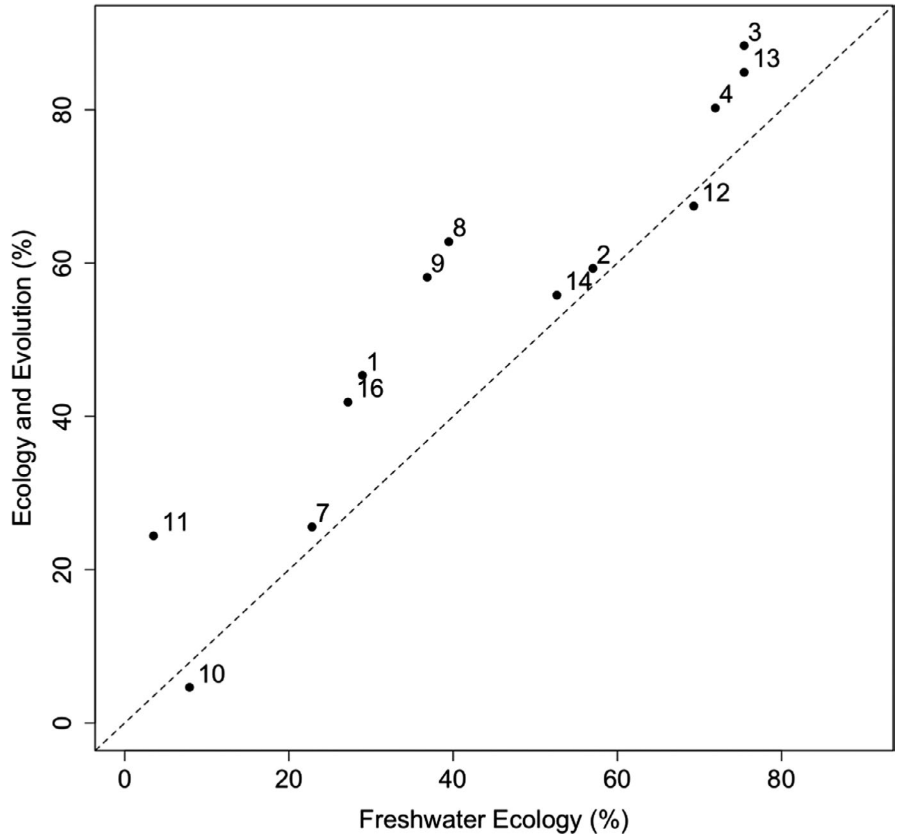
TABLE 4 Compliance with the criteria for reporting and following methodological standards assigned to 86 ecological and evolutionary meta-analyses reviewed by Senior et al. (2016). A detailed description of the criteria is shown in Table 2. *Indicates criteria that were disregarded for ecological and evolutionary meta-analyses

	Quality criteria	Yes (%)	No (%)	Partial (%)
1	Searching details	49.41	27.06	20.00
2	Inclusion/exclusion criteria	63.53	11.76	21.18
3	Weighted effect sizes	92.94	7.06	0.00
4	Meta-analytical model	84.71	15.29	0.00
5	Heterogeneity in effect sizes	*	*	*
6	Causes of heterogeneity	*	*	*
7	Collinearity analysis	27.06	63.53	0.00
8	Publication bias	65.88	30.59	0.00
9	Sensitivity analysis	61.18	38.82	0.00
10	Changes in effect size	5.88	90.59	0.00
11	Controlling for phylogeny	24.71	45.88	0.00
12	Software	71.76	28.24	0.00
13	Bibliographic details	89.41	7.06	0.00
14	Data	58.82	41.18	0.00
15	Standard effect size	*	*	*
16	Multiple effect sizes	42.35	35.29	4.71

especially if reliability is questionable (Ioannidis, 2010, 2016, 2017; Nakagawa et al., 2017). In this context, it was disturbing to find many studies stating that they conducted meta-analyses despite the use of different approaches (e.g. re-analysis of primary data, vote counting, scientometrics, qualitative reviews). Other studies failed to mention the criteria used to select studies, disregarded publication bias analyses (e.g. Koricheva & Gurevitch, 2014; Nakagawa et al., 2017 for plant ecology and ecology and evolution, respectivelly) or focused mostly on statistical significance, often neglecting to model heterogeneity (Higgins & Thompson, 2002; Koricheva & Gurevitch, 2014; Senior et al., 2016). Many of the studies failed to address non-independence issues with the use of multiple effect sizes from the same studies (e.g. Hedges et al., 2010), or at least to conduct a sensitivity analysis to demonstrate the robustness of their results (Gurevitch et al., 2018). The analysis of temporal stability in effect sizes was also disregarded in several meta-analyses, which is of concern as it may lead to misleading conclusions (e.g. Koricheva & Kulinskaya, 2019).

According to Koricheva and Gurevitch (2014), meta-analysis is a "set of statistical methods for combining outcomes (effect sizes) across different data sets addressing the same research question to examine patterns of response across these data sets and sources of heterogeneity in outcomes". Here, 224 studies were excluded solely for not being formal meta-analyses. Some of these exclusions occurred after reading the abstract (77) because the studies stated that their results were important for the development of future meta-analysis, while others cited the term meta-analysis in other contexts. The most striking result concerns the exclusion of 154 studies, after reading the full texts, for not using meta-analytic approaches (e.g. vote-counting and re-analysis of primary data), despite being described as meta-analyses by the authors. This erroneous use of the term meta-analysis has also been observed in other areas of ecology (Koricheva & Gurevitch, 2014; Senior et al., 2016; Vetter et al., 2013), despite many studies urging authors to "apply the

FIGURE 3 Relationship between percentages of meta-analytical studies complying with different criteria for both freshwater ecology (this study) and ecology and evolution (compiled by Senior et al., 2016). Numbers indicate the quality criteria (Table 2). Values above the line represent criteria that were met more frequently by ecological and evolutionary meta-analyses, while values below the line were met more frequently by freshwater ecology meta-analyses



term *meta-analysis* consistently and correctly and not to confuse it with other summary techniques" (e.g. Vetter et al., 2013). Indeed, we found that the number of studies erroneously self-classified as meta-analysis is still increasing (Spearman's $r = 0.74$; $p = 0.0002$; $n = 20$ years [1998–2017]).

We found that meta-analyses in freshwater ecology require improvements, in line with previous studies from different fields (Delaney et al., 2005; Ioannidis, 2016; Koricheva & Gurevitch, 2014). Meta-analysis quality was not correlated with journal impact factor, indicating that the most well-conducted meta-analyses are not necessarily published in journals with high impact factors (Brembs, 2018). The lack of correlation between the average number of citations an article received and its quality score, in turn, suggests that other factors (e.g. themes and specific results) are more important to *citability* than the quality of the meta-analysis itself (see also Padial et al., 2010 for a general analysis of citation frequency of ecological articles). However, we found that our rating scores were significantly correlated with publication year, suggesting that published guidelines to improve the use of meta-analyses (e.g. Koricheva & Gurevitch, 2014; Moher et al., 2009; Nakagawa et al., 2017) are being effective.

Reporting the search strings and databases (e.g. WoS and Scopus) used to find articles that provide data to a meta-analysis, is far from being a mere bureaucratic step (Nakagawa et al., 2007; Gusenbauer & Haddaway, 2020). For example, researchers may avoid wasting time through the repetition of an inclusive meta-analysis when search strings are reported or may detect the need for a new meta-analysis when/if the one published was too restrictive (see Babić et al., 2020). However, our results parallel those of Koricheva and Gurevitch, (2014) by showing that only a few studies included full details of bibliographic searches. The problem becomes more pronounced when the inclusion/exclusion criteria are not reported (for a discussion on this theme, see Lortie & Callaway, 2006). Even the easy-to-meet quality criterion of providing the list of primary studies included in the meta-analysis was not met by 21 studies, despite being essential for reproducibility testing. Specifically, if the list of primary studies used in a published meta-analysis differs substantially from a list of primary studies retrieved by using a different search string (e.g. due to the use of synonyms or different technical terms), then, all else being equal, a new meta-analysis, considering the different set of primary studies, may produce different results and new insights (e.g. compare Westgate et al., 2014 to de Morais et al., 2018 and Stein et al., 2014 to Ortega et al., 2018; see also Page & Moher, 2016 for a discussion about the problem of redundant meta-analyses). In addition, only nearly half of the meta-analyses in freshwater ecology provided the dataset used (60 studies). Proportionally, this figure was even lower in plant ecology (Koricheva & Gurevitch, 2014). Providing the dataset used is crucial for cumulative meta-analysis (Leimu & Koricheva, 2004) and to explore the same dataset using different moderators (e.g. Menegotto et al., 2019).

A critical step in a meta-analysis is to weight effect sizes by the inverse variance method (Borenstein et al., 2009; Koricheva & Gurevitch, 2013). Many of the studies we reviewed weighted the effect sizes (86 out of 114). Still, 28 of the studies included in our review did not use the inverse variance method. In some cases, the authors stated that the variance (or the data needed to estimate it) was not available in the primary studies. However, other weighting strategies are available, for example, estimating variance from partial data and weighting by *n* (Koricheva & Gurevitch, 2014). As the quality of reporting and transparency increases in primary studies, the frequency of these issues is expected to decrease (Gerstner et al., 2017).

Thirty-two studies (out of 114) did not report the meta-analytical model used to analyse the data. However, different models (fixed or random effects) have different assumptions, test different hypotheses, and may provide contrasting results (Borenstein et al., 2009; Koricheva & Gurevitch, 2013). In general, the use of a random-effects model is the best alternative because of the restrictive assumptions of a fixed-effect model (mainly that effect sizes are homogeneous across primary studies). Similarly, 35 studies did not report the software used, although this is important because it may influence results (Koricheva & Gurevitch, 2014).

Heterogeneity measures are crucial for a comprehensive interpretation of the weighted mean effect size (Forstmeier et al., 2017; Nakagawa et al., 2017; Senior et al., 2016). Interpreting the weighted mean effect size without the heterogeneity measures is comparable to drawing conclusions from the mean without showing the standard deviation. Few studies have quantified heterogeneity (47 out 114; heterogeneity quantified as T^2 and Q). Still, many explored possible causes of variation (94 out of 114), because most meta-analyses are more interested in exploring variation in effect sizes among studies and in the causes of this variation (Koricheva & Gurevitch, 2014). However, not reporting the heterogeneity or variation of effect sizes is concerning for ecological meta-analyses, given the high complexity of natural systems (Vetter et al., 2013).

The basic strategy to explore causes of heterogeneity is to conduct a meta-regression with several moderators, which corresponds to a weighted multiple regression. Thus, the same assumptions of multiple regression also apply to meta-regression (e.g. checking for collinearity among moderators; Borenstein et al., 2009; Koricheva & Gurevitch, 2013; Nakagawa et al., 2017). While many studies explored the causes of heterogeneity in effect sizes, less than a quarter checked for collinearity. Issues such as instability of parameter estimates, inflated standard errors and biased inferences (Dormann et al., 2013) arise when moderators are collinear. As in any other research approach, the issues with multicollinearity highlight the need for careful planning to conduct a meta-analysis (Berman & Parker, 2002). For example, many studies are screened to compile data on moderators hypothetically related to the effect size. However, without proper planning and theoretical reasoning, this hard work may be in vain if different pairs of moderators are highly correlated to each other or, even worse, if redundant moderators are compiled at the expense of ignoring other relevant moderators.

Publication bias is said to exist when the results of a study influence the decision of authors, reviewers, and editors to publish a manuscript. This subject has been extensively debated (Dwan et al., 2013; Møller & Jennions, 2001; Rothstein et al., 2006) due to its effects on systematic reviews and meta-analyses (and also on narrative reviews). Positive results are more likely to be published (Fanelli, 2012; Fanelli et al., 2017; Forstmeier et al., 2017). In this context, positive results may be over-represented in the pool of articles available for meta-analysis potentially leading to overestimation of the weighted mean effect size (Møller & Jennions, 2001). Although researchers highlight the need to study publication bias (e.g. Jennions, 2013; Møller & Jennions, 2001; Tomkins & Kotiaho, 2004), over half of the studies (62 out of 114) did not assess it. The lack of concern with publication bias is a recurrent issue in ecology (Koricheva & Gurevitch, 2014; Nakagawa & Santos, 2012; Roberts et al., 2006), as are requests to analyse this issue (Nakagawa & Poulin, 2012).

Twenty-seven meta-analytical studies, out of the 45 that tested for publication bias, used only one method. However, the use of multiple approaches is advisable (Lin et al., 2018). The funnel plot, which may indicate publication bias through asymmetric distribution of effect sizes, was the most used approach (28 papers). The funnel-plot has been criticised on several grounds (Jennions, 2013; Lin, 2019; Terrin et al., 2005). For instance, asymmetry may be the consequence of other factors when the number of primary studies (k) is small (Egger et al., 1997). Consequently, many studies also used the funnel plot associated with other tests (19 studies—fail-safe numbers, quantile-plot, and correlation). The *trim and fill* method is often used to *correct* for publication bias (Duval & Tweedie, 2000; Peters et al., 2007; and see Weinhandl & Duval, 2012 for an alternative method applied to meta-regression). However, this method tests how robust the results are to publication bias (Jennions & Møller, 2002a; Sutton, 2002) and, therefore, is more of a sensitivity analysis than a method to correct for publication bias (Rothstein et al., 2006).

Conducting sensitivity analyses is also important to check the robustness of results. Basically, it consists of double-checking decisions made throughout the study (Noble et al., 2017). The use of multiple effect sizes, models, publication bias methods and the evaluation of temporal changes in effect sizes are all issues that can be evaluated. However, the best strategy needs to be determined on a case-by-case basis.

A useful strategy to analyse the results is to evaluate whether the effect sizes vary over time (e.g. using a cumulative meta-analysis; Leimu & Koricheva, 2004; Ortega et al., 2018) or to use publication year as a moderator when exploring sources of heterogeneity (Lehmann et al., 2012; Wood & Eagly, 2009). Cumulative sum charts may also be used to detect possible outliers and temporal tendencies (Dogo et al., 2017; Kulinskaya & Koricheva, 2010). Causes of variation include the tendency of earlier studies to publish higher effect sizes (i.e. time-lag bias), publication bias, Proteus phenomenon (the attraction to contradictory results), true heterogeneity, or even chance (Ioannidis, 2008; Jennions & Møller, 2002b; Trikalinos & Ioannidis, 2006). A decrease in effect size over time can be also attributed to the test of a "valid theory beyond its domain of application" (Wilson et al., 2020). Despite evidence that effect sizes may change over time (see Koricheva et al., 2013 for detailed examples) and the practical implications of this change (Koricheva &

Kulinskaya, 2019), only nine meta-analytical studies in freshwater ecology explored this topic.

Independent data points are essential for statistical testing (Forstmeier et al., 2017). Non-independence of effect sizes occurs when the data exhibit a correlated or clustered structure, which may be driven by several factors (Lajeunesse, Rosenberg, & Jennions, 2013). For example, in a cross-species meta-analysis, one can find that the effect sizes are phylogenetically structured. Thus, by not accounting for the phylogenetic structure, we would violate the assumption of independence of effect sizes, which has been shown to change the results of meta-analyses (Chamberlain et al., 2012; Nakagawa & Santos, 2012). We found that 29 out of 33 studies for which this criterion was relevant did not control for phylogenetic structure. Nevertheless, methods to incorporate phylogeny have been extensively developed in evolutionary biology and are also available for different meta-analysis models (Adams, 2008; Lajeunesse, 2009; Lajeunesse et al., 2013; Nakagawa & Santos, 2012).

Primary studies often report more than one effect sizes. For example, a primary study may test the relationship between a response variable and different explanatory variables (across the same sampling units). However, if multiple effect sizes per study are used, the total sample size of a meta-analysis will be inflated (Jennions & Schmid, 2013), resulting in increased type I error rates (Forstmeier et al., 2017; Hedges et al., 2010; Noble et al., 2017). Only 30 freshwater ecology studies addressed this issue, primarily by comparing the weighted mean effect size obtained using multiple effect sizes to the weighted mean effect size obtained by using one (selected or averaged) effect size per study (e.g. following the approaches described in López-López et al., 2018). There are several statistical methods available to address effect size multiplicity (e.g. Jennions & Schmid, 2013; López-López et al., 2018; Tanner-Smith & Tipton, 2014). For example, Hedges et al. (2010) proposed a method called Robust Variance Estimate. Given the availability of software (e.g. Fisher & Tipton, 2015), we think that the issue of multiple effect sizes per study will be addressed in most future studies.

Our overall results (mean quality) may be generalisable among different fields in ecology (i.e. considering the results for freshwater ecology, ecology, and evolution [Senior et al., 2016] and plant ecology [Koricheva & Gurevitch, 2014]). However, ecological and evolutionary meta-analyses had a higher compliance with the quality criteria than freshwater ecology meta-analyses. It is noteworthy that few studies explored temporal dependence in effect sizes, checked for dependence among moderators, or controlled for phylogenetic dependence in either areas, suggesting this may be a recurrent issue. A similar pattern was observed for plant ecology (Koricheva & Gurevitch, 2014). Thus, we think that further studies should be especially aware of these issues, once lack of independence (temporal, phylogenetic, or in explanatory variables), temporal trends in effect sizes and collinearity among moderators are issues of primary concern.

Our findings point to the need for a greater adherence to guidelines for systematic reviews. Thus, we reiterate the importance of applying the checklists, which are widely available, to improve reporting of meta-analyses in different areas of ecology (e.g. Koricheva & Gurevitch, 2014; Moher et al., 2009, 2015; Parker et al., 2018; Roberts et al., 2006). We expect that the use of checklists will increase the quality of meta-analyses, as has been reported, for example, for biomedical research (Han et al., 2017). The need to enhance quality should not be taken lightly, given that their results are potentially used by decision makers as decisive evidence (Koricheva & Kulinskaya, 2019).

5 | CONCLUSION

We found that the quality of meta-analyses in freshwater ecology has increased over time; however, there is much room for improvement in complying with important quality criteria. Firstly, authors should not mislabel their studies as meta-analyses when, in fact, the method was not used. Secondly, compliance with checklists should be widely fostered as meta-analyses are increasingly being used to summarise findings in different areas of ecology. Thus, authors, reviewers, and editors should comprehensively use checklists to improve the quality of meta-analyses in freshwater ecology.

CONFLICT OF INTEREST STATEMENT

We declare there is no conflict of interest associated with any of the decisions made or components of this study.

ACKNOWLEDGEMENTS

We would like to thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brazil (CAPES) for the scholarship granted to S.L. (Code 001) and J.C.G.O., and L.M.B. received research fellowships from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, process no: 304314/2014-5). This work was also developed in the context of the National Institutes for Science and Technology (INCT) in Ecology, Evolution and Biodiversity Conservation, supported by MCTIC/CNPq (proc. the National Institutes for Science and Technology) and FAPEG. Two anonymous reviewers and the editor made helpful comments on earlier drafts of this paper.

DATA AVAILABILITY STATEMENT

All data collected to develop this study are available at Mendeley Data (https://doi.org/10.17632/pf9tnjfhp4.1).

ORCID

Sara Lodi https://orcid.org/0000-0001-8318-6362
Bruno S. Godoy https://orcid.org/0000-0001-9751-9885
Jean C. G. Ortega https://orcid.org/0000-0001-5097-9382
Luis M. Bini https://orcid.org/0000-0003-3398-9399

REFERENCES

Adams, D. C. (2008). Phylogenetic meta-analysis. *Evolution*, *62*, 567–572. https://doi.org/10.1111/j.1558-5646.2007.00314.x

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*, 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Babić, A., Pericic, T. P., Pieper, D., & Puljak, L. (2020). How to decide whether a systematic review is stable and not in need of updating: Analysis of Cochrane reviews. *Research Synthesis Methods*, *11*, 6. https://doi.org/10.1002/jrsm.1451

Berman, N. G., & Parker, R. A. (2002). Meta-analysis: Neither quick nor easy. *BMC Medical Research Methodology*, *2*, 10. https://doi.org/10.1186/1471-2288-2-10

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*, 1st. ed. John Wiley & Sons. https://doi.org/10.1002/9780470743386

Brembs, B. (2018). Prestigious science journals struggle to reach even average reliability. *Frontiers in Human Neuroscience*, *12*, 1–7. https://doi.org/10.3389/fnhum.2018.00376

Chamberlain, S. A., Hovick, S. M., Dibble, C. J., Rasmussen, N. L., Van Allen, B. G., Maitner, B. S., … Whitney, K. D. (2012). Does phylogeny matter? Assessing the impact of phylogenetic information in ecological meta-analysis. *Ecology Letters*, *15*, 627–636. https://doi.org/10.1111/j.1461-0248.2012.01776.x

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46. https://doi.org/10.1177/001316446002000104

Collins, J. A., & Fauser, B. C. J. M. (2005). Balancing the strengths of systematic and narrative reviews. *Human Reproduction Update*, *11*, 103–104. https://doi.org/10.1093/humupd/dmh058

Coursaris, C. K., & Van Osch, W. (2014). A scientometric analysis of social media research (2004–2011). *Scientometrics*, *101*, 357–380. https://doi.org/10.1007/s11192-014-1399-z

de Morais, G. F., Ribas, L. G. S., Ortega, J. C. G., Heino, J., & Bini, L. M. (2018). Biological surrogates: A word of caution. *Ecological Indicators*, *88*, 214–218. https://doi.org/10.1016/j.ecolind.2018.01.027

Delaney, A., Bagshaw, S. M., Ferland, A., Manns, B., Laupland, K. B., & Doig, C. J. (2005). A systematic evaluation of the quality of meta-analyses in the critical care literature. *Critical Care*, *9*, R575–R582. https://doi.org/10.1186/cc3803

Dogo, S. H., Clark, A., & Kulinskaya, E. (2017). Sequential change detection and monitoring of temporal trends in random-effects meta-analysis. *Research Synthesis Methods*, *8*, 220–235. https://doi.org/10.1002/jrsm.1222

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*, 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Dwan, K., Gamble, C., Williamson, P. R., & Kirkham, J. J. (2013). Systematic review of the empirical evidence of study publication bias and outcome reporting bias—an updated review. *PLoS One*, *8*, e66844. https://doi.org/10.1371/journal.pone.0066844

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*, 629–634. https://doi.org/10.1136/bmj.315.7109.629

Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511761676

Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics*, *90*, 891–904. https://doi.org/10.1007/s11192-011-0494-7

Fanelli, D., Costas, R., & Ioannidis, J. P. A. (2017). Meta-assessment of bias in science. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 3714–3719. https://doi.org/10.1073/pnas.1618569114

Felsenstein, J. (1985). Phylogenies and the comparative method. *The American Naturalist*, *125*, 1–15. https://doi.org/10.1086/284325

Felson, D. T. (1992). Bias in meta-analytic research. *Journal of Clinical Epidemiology*, *45*, 885–892. https://doi.org/10.1016/0895-4356(92)90072-U

Fisher, Z., & Tipton, E. (2015). robumeta: An R-package for robust variance estimation in meta-analysis. DOI: arXiv:1503.02220v1.

Fleming, P. S., Koletsi, D., Seehra, J., & Pandis, N. (2014). Systematic reviews published in higher impact clinical journals were of higher quality. *Journal of Clinical Epidemiology*, *67*, 754–759. https://doi.org/10.1016/j.jclinepi.2014.01.002

Forstmeier, W., Wagenmakers, E.-J., & Parker, T. H. (2017). Detecting and avoiding likely false-positive findings—a practical guide. *Biological Reviews*, *92*, 1941–1968. https://doi.org/10.1111/brv.12315

Gates, S. (2002). Review of methodology of quantitative reviews using meta-analysis in ecology. *Journal of Animal Ecology*, *71*, 547–557. https://doi.org/10.1046/j.1365-2656.2002.00634.x

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. https://doi.org/10.1111/2041-210X.12758

Gurevitch, J., Koricheva, J., Nakagawa, S., & Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature*, *555*, 175–182. https://doi.org/10.1038/nature25753

Gusenbauer, M., & Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*, *11*, 181–217. https://doi.org/10.1002/jrsm.1378

Han, S., Olonisakin, T. F., Pribis, J. P., Zupetic, J., Yoon, J. H., Holleran, K. M., … Lee, J. S. (2017). A checklist is associated with increased quality of reporting preclinical biomedical research: A systematic review. *PLoS One*, *12*, e0183591. https://doi.org/10.1371/journal.pone.0183591

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. https://doi.org/10.1002/jrsm.5

Higgins, J. P. T., & Green, S.. (2011). *Cochrane handbook for systematic reviews of interventions*. Version 5.1.0 [updated March 2011]. : The Cochrane Collaboration.

Higgins, J. P. T., Lane, P. W., Anagnostelis, B., Anzures-Cabrera, J., Baker, N. F., Cappelleri, J. C., … Whitehead, A. (2013). A tool to assess the quality of a meta-analysis. *Research Synthesis Methods*, *4*, 351–366. https://doi.org/10.1002/jrsm.1092

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*, 1539–1558. https://doi.org/10.1002/sim.1186

Hillebrand, H., & Cardinale, B. J. (2010). A critique for meta-analyses and the productivity–diversity relationship. *Ecology*, *91*, 2545–2549. https://doi.org/10.1890/09-0070.1

Houlahan, J. E., McKinney, S. T., Anderson, T. M., & McGill, B. J. (2017). The priority of prediction in ecological understanding. *Oikos*, *126*, 1–7. https://doi.org/10.1111/oik.03726

Ioannidis, J. P. A. (2008). Why most discovered true associations are inflated. *Epidemiology*, *19*, 640–648. https://doi.org/10.1097/EDE.0b013e31818131e7

Ioannidis, J. P. A. (2010). Meta-research: The art of getting it wrong. *Research Synthesis Methods*, *1*, 169–184. https://doi.org/10.1002/jrsm.19

Ioannidis, J. P. A. (2016). The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Quarterly*, *94*, 485–514. https://doi.org/10.1111/1468-0009.12210
Ioannidis, J. (2017). Next-generation systematic reviews: Prospective meta-analysis, individual-level data, networks and umbrella reviews. *British Journal of Sports Medicine*, *51*, 1456–1458. https://doi.org/10.1136/bjsports-2017-097621
Jennions, M. D. (2013). Publication and related biases. In: *The handbook of meta-analysis in ecology and evolution* (Eds: J. Koricheva, J. Gurevitch & K. Mengersen). Princeton University Press. Princeton.
Jennions, M. D., & Møller, A. P. (2002a). Publication bias in ecology and evolution: An empirical assessment using the "trim and fill" method. *Biological Reviews of the Cambridge Philosophical Society*, *77*, 211–222. https://doi.org/10.1017/s1464793101005875
Jennions, M. D., & Møller, A. P. (2002b). 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. https://doi.org/10.1098/rspb.2001.1832
Jennions, M. D., & Schmid, C. H. (2013). Statistical models for the meta-analysis of nonindependent data. In: *The handbook of meta-analysis in ecology and evolution* (Eds: J. Koricheva, J. Gurevitch & K. Mengersen). Princeton University Press. Princeton.
Koricheva, J., & Gurevitch, J. (2013). Place of meta-analysis among other methods of research synthesis. In: *The handbook of meta-analysis in ecology and evolution* (Eds: J. Koricheva, J. Gurevitch & K. Mengersen). Princeton University Press, Princeton.
Koricheva, J., & Gurevitch, J. (2014). Uses and misuses of meta-analysis in plant ecology. *Journal of Ecology*, *102*, 828–844. https://doi.org/10.1111/1365-2745.12224
Koricheva, J., Jennions, M., & Lau, J. (2013). Temporal trends in effect sizes: Causes, detection, and implications. In: *The handbook of meta-analysis in ecology and evolution* (Eds: J. Koricheva, J. Gurevitch & K. Mengersen). Princeton University Press, Princeton.
Koricheva, J., & Kulinskaya, E. (2019). Temporal instability of evidence base: A threat to policy making? *Trends in Ecology and Evolution*, *34*, 895–902. https://doi.org/10.1016/j.tree.2019.05.006
Kulinskaya, E., & Koricheva, J. (2010). Use of quality control charts for detection of outliers and temporal trends in cumulative meta-analysis. *Research Synthesis Methods*, *1*, 297–307. https://doi.org/10.1002/jrsm.29
Lajeunesse, M. J. (2009). Meta-analysis and the comparative phylogenetic method. *The American Naturalist*, *174*, 369–381. https://doi.org/10.1086/603628
Lajeunesse, M. J., Rosenberg, M. S., & Jennions, M. D. (2013). Phylogenetic nonindependence and meta-analysis. In: *The handbook of meta-analysis in ecology and evolution* (Eds: J. Koricheva, J. Gurevitch & K. Mengersen). Princeton University Press. Princeton.
Lehmann, A., Barto, E. K., Powell, J. R., & Rillig, M. C. (2012). Mycorrhizal responsiveness trends in annual crop plants and their wild relatives-a meta-analysis on studies from 1981 to 2010. *Plant and Soil*, *355*, 231–250. https://doi.org/10.1007/s11104-011-1095-1
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 B: Biological Sciences*, *271*, 1961–1966. https://doi.org/10.1098/rspb.2004.2828
Lin, L. (2019). Graphical augmentations to sample-size-based funnel plot in meta-analysis. *Research Synthesis Methods*, *10*, 376–388. https://doi.org/10.1002/jrsm.1340
Lin, L., Chu, H., Murad, M. H., Hong, C., Qu, Z., Cole, S. R., & Chen, Y. (2018). Empirical comparison of publication bias tests in meta-analysis. *Journal of General Internal Medicine*, *33*, 1260–1267. https://doi.org/10.1007/s11606-018-4425-7
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*, *9*, 336–351. https://doi.org/10.1002/jrsm.1310
Lortie, C. J. (2014). Formalized synthesis opportunities for ecology: Systematic reviews and meta-analyses. *Oikos*, *123*, 897–902. https://doi.org/10.1111/j.1600-0706.2013.00970.x
Lortie, C. J., & Callaway, R. M. (2006). Re-analysis of meta-analysis: Support for the stress-gradient hypothesis. *Journal of Ecology*, *94*, 7–16. https://doi.org/10.1111/j.1365-2745.2005.01066.x
Menegotto, A., Kurtz, M. N., da Lana, P., & Lana, P. C. (2019). Benthic habitats do show a significant latitudinal diversity gradient: A comment on Kinlock et al (2018). Global Ecology and Biogeography, 28, 1712–1717. DOI: 10.1111/geb.12970.
Mengersen, K., Jennions, M. D., & Schmid, C. H. (2013). Statistical models for the meta-analysis of non-independent data. In: *The handbook of meta-analysis in ecology and evolution* (Eds: J. Koricheva, J. Gurevitch & K. Mengersen). Princeton University Press. Princeton.
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 Med*, *6*, e1000097. https://doi.org/10.1371/journal.pmed.1000097
Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., … Stewart, L. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4*, 1. https://doi.org/10.1186/2046-4053-4-1
Møller, A. P., & Jennions, M. D. (2001). Testing and adjusting for publication bias. *Trends in Ecology and Evolution*, *16*, 580–586. https://doi.org/10.1016/S0169-5347(01)02235-2
Morrissey, M. B. (2016). Meta-analysis of magnitudes, differences and variation in evolutionary parameters. *Journal of Evolutionary Biology*, *29*, 1882–1904. https://doi.org/10.1111/jeb.12950
Murad, M. H., Asi, N., Alsawas, M., & Alahdab, F. (2016). New evidence pyramid. *Evidence-Based Medicine*, *21*, 125–127. https://doi.org/10.1136/ebmed-2016-110401
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*, 18. https://doi.org/10.1186/s12915-017-0357-7
Nakagawa, S., & Poulin, R. (2012). Meta-analytic insights into evolutionary ecology: An introduction and synthesis. *Evolutionary Ecology*, *26*, 1085–1099. https://doi.org/10.1007/s10682-012-9593-z
Nakagawa, S., & Santos, E. S. A. (2012). Methodological issues and advances in biological meta-analysis. *Evolutionary Ecology*, *26*, 1253–1274. https://doi.org/10.1007/s10682-012-9555-5
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. https://doi.org/10.1111/mec.14031
Olkin, I. (1992). Meta-analysis: Methods for combining independent studies. *Statistical Science*, *7*, 226–236.
Ortega, J. C. G., Thomaz, S. M., & Bini, L. M. (2018). Experiments reveal that environmental heterogeneity increases species richness, but they are rarely designed to detect the underlying mechanisms. *Oecologia*, *188*, 11–22. https://doi.org/10.1007/s00442-018-4150-2
Orwin, R. G., & Vevea, J. L. (2009). Evaluating coding decisions. In: The handbook of research synthesis & meta-analysis (Eds: H. Cooper, L. V. Hedges, & J. C. Valentine). Russell Sage Foundation, New York.
Padial, A. A., Nabout, J. C., Siqueira, T., Bini, L. M., & Diniz-Filho, J. A. F. (2010). Weak evidence for determinants of citation frequency in ecological articles. *Scientometrics*, *85*, 1–12. https://doi.org/10.1007/s11192-010-0231-7.
Page, M. J., & Moher, D. (2016). Mass production of systematic reviews and meta-analyses: An exercise in mega-silliness? *Milbank Quarterly*, *94*, 515–519. https://doi.org/10.1111/1468-0009.12211
Parker, T. H., Bowman, S. D., Nakagawa, S., Gurevitch, J., Mellor, D. T., Rosenblatt, R. P., & DeHaven, A. C. (2018). Tools for Transparency in Ecology and Evolution (TTEE). Available in <https://osf.io/g65cb/> DOI: 10.17605/OSF.IO/G65CB.

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2007). Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Statistics in Medicine*, *26*, 4544–4562. https://doi.org/10.1002/sim.2889

Philibert, A., Loyce, C., & Makowski, D. (2012). Assessment of the quality of meta-analysis in agronomy. *Agriculture, Ecosystems and Environment*, *148*, 72–82. https://doi.org/10.1016/j.agee.2011.12.003

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Roberts, P. D., Stewart, G. B., & Pullin, A. S. (2006). Are review articles a reliable source of evidence to support conservation and environmental management? A comparison with medicine. *Biological Conservation*, *132*, 409–423. https://doi.org/10.1016/j.biocon.2006.04.034

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2006). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. John Wiley & Sons.

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. https://doi.org/10.1002/ecy.1591

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., … 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. https://doi.org/10.1136/bmj.g7647

Song, S., Peacor, D. S., Osenberg, C. W., & Bence, J. R. (2020). An assessment of statistical methods for nonindependent data in ecological meta-analyses. *Ecology*, *101*(12), https://doi.org/10.1002/ecy.3184

Stein, A., Gerstner, K., & Kreft, H. (2014). Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters*, *17*, 866–880. https://doi.org/10.1111/ele.12277

Sutton, A. J. (2002). Empirical assessment of effect of publication bias on meta-analyses. *BMJ*, *320*, 1574. https://doi.org/10.1136/bmj.320.7249.1574

Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods*, *5*, 13–30. https://doi.org/10.1002/jrsm.1091

Terrin, N., Schmid, C. H., & Lau, J. (2005). In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *Journal of Clinical Epidemiology*, *58*, 894–901. https://doi.org/10.1016/j.jclinepi.2005.01.006

Tomkins, J. L., & Kotiaho, J. S. (2004). Publication bias in meta-analysis: Seeing the wood for the trees. *Oikos*, *104*, 194–196. https://doi.org/10.1111/j.0030-1299.2004.12357.x

Trikalinos, T. A., & Ioannidis, J. P. A. (2006). Assessing the evolution of effect sizes over time. In: Publication bias in meta-analysis: prevention, assessment and adjustments (Eds: H. R. Rothstein, A. J. Sutton & M. Borenstein). John Wiley & Sons, Chichester.

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2014). Meta-analysis of multiple outcomes: A multilevel approach. *Behavior Research Methods*, *47*, 1274–1294. https://doi.org/10.3758/s13428-014-0527-2

Vetter, D., Rucker, G., & Storch, I. (2013). Meta-analysis: A need for well-defined usage in ecology and conservation biology. *Ecosphere*, *4*, 1–24. https://doi.org/10.1890/ES13-00062.1

Weinhandl, E. D., & Duval, S. (2012). Generalization of trim and fill for application in meta-regression. *Research Synthesis Methods*, *3*, 51–67. https://doi.org/10.1002/jrsm.1042

Westgate, M. J., Barton, P. S., Lane, P. W., & Lindenmayer, D. B. (2014). Global meta-analysis reveals low consistency of biodiversity congruence relationships. *Nature Communications*, *5*, 3899. https://doi.org/10.1038/ncomms4899

Whittaker, R. J. (2010). Meta-analyses and mega-mistakes: Calling time on meta-analysis of the species richness-productivity relationship. *Ecology*, *91*, 2522–2533. https://doi.org/10.1890/08-0968.1

Willis, B. H., & Quigley, M. (2011). The assessment of the quality of reporting of meta-analyses in diagnostic research: A systematic review. *BMC Medical Research Methodology*, *11*, 163. https://doi.org/10.1186/1471-2288-11-163

Wilson, B. M., Harris, C. R., & Wixted, J. T. (2020). Science is not a signal detection problem. *Proceedings of the National Academy of Sciences of the United States of America*, *117*, 5559–5567. https://doi.org/10.1073/pnas.1914237117

Wood, W., & Eagly, A. H. (2009). Advantages of certainty and uncertainty. In: The handbook of research synthesis and meta-analysis (Eds: H. Cooper, L. V. Hedges & J. C. Valentine). Russell Sage Foundation, New York. DOI: 9780871541635.

Wooster, D. (1994). Predator impacts on stream benthic prey. *Oecologia*, *99*, 7–15. https://doi.org/10.1007/BF00317077

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Lodi S, Spacek Godoy B, Carlo Gonçalves Ortega J, Mauricio Bini L. Quality of meta-analyses in freshwater ecology: A systematic review. *Freshw Biol*. 2021;00:1–12. https://doi.org/10.1111/fwb.13695