doi:10.1038/nature25753

Meta-analysis and the science of research synthesis

Jessica Gurevitch[1], Julia Koricheva[2], Shinichi Nakagawa[3,4] & Gavin Stewart[5]

Meta-analysis is the quantitative, scientific synthesis of research results. Since the term and modern approaches to research synthesis were first introduced in the 1970s, meta-analysis has had a revolutionary effect in many scientific fields, helping to establish evidence-based practice and to resolve seemingly contradictory research outcomes. At the same time, its implementation has engendered criticism and controversy, in some cases general and others specific to particular disciplines. Here we take the opportunity provided by the recent fortieth anniversary of meta-analysis to reflect on the accomplishments, limitations, recent advances and directions for future developments in the field of research synthesis.

Synthesizing results across studies to reach an overall understanding of a problem and to identify sources of variation in outcomes is an essential part of the scientific process. Until recently, the results of scientific studies have been summarized in narrative reviews. However, this approach becomes inadequate when there are hundreds of studies on a given research question[1,2], and the difficulties of carrying out narrative reviews to identify and summarize evidence in a transparent and objective manner have become increasingly apparent as research results have mushroomed across scientific fields[3].

During the past few decades, scientifically rigorous systematic reviews and meta-analyses, carried out following formal protocols to ensure reproducibility and reduce bias, have become more prevalent in a range of fields[1] (Box 1). Systematic reviews aim to provide a robust overview of the efficacy of an intervention, or of a problem or field of research. They can be combined with quantitative meta-analyses to assess the magnitude of the outcome across relevant primary studies and to analyse the causes of variation among study outcomes (effect sizes). Narrative reviews remain useful for exploring the development of particular ideas (as we do here) and for advancing conceptual frameworks, but they cannot accurately summarize results across studies[4].

Four decades after its introduction, we are seeing widespread mainstream acceptance of meta-analysis as a research synthesis tool, but also the signs of what may be considered a 'midlife crisis' as it has begun the transition to a mature field. While the number of published meta-analyses has continued to increase rapidly, too many meta-analyses and systematic reviews are of low quality[5–7]. The publication of methodologically flawed meta-analyses indicates that peer reviewers, editors and authors are not fully aware of or are indifferent to the large body of well-developed meta-analytic methodology, and that reviewers might feel unqualified to address statistical issues. Low-quality meta-analyses have attracted strong criticism[5,8] and even calls for a halt in publication of all meta-analyses[9]. Although it is certainly both valid and valuable to criticize poor methodology and reporting, such criticism should result in a call for improved standards (as for pre-clinical trials[10]) rather than abandonment of the field[11]. We believe that the solution lies in the rigorous application of stricter methodological and reporting quality criteria for publishing meta-analyses (see, for example, Tools for Transparency in Ecology and Evolution; https://osf.io/g65cb), and in better training for practitioners and reviewers in the rationales and methodologies of meta-analyses and systematic reviews.

Here we highlight some of the main principles and characteristics of high-quality meta-analytic methodology and briefly summarize the development of the field. We also discuss the limitations, utility and achievements of meta-analysis in several fields and, as a case study, its role in advances in ecology, evolutionary biology and conservation (EEC). Finally, we address several recent criticisms of the meta-analytic approach and suggest ways in which future developments in research synthesis could facilitate the most rapid progress in the fields in which it is used.

Meta-analyses use well-documented methodologies

Systematic reviews aim to be transparent, reproducible and updatable, and to address well-defined questions. The systematic review process includes the use of formal methodological guidelines for the literature search, study screening (including critical appraisal of eligible studies according to pre-defined criteria), data extraction, coding and often statistical analysis (that is, meta-analysis), along with detailed, transparent documentation of each step. Software, protocols and reporting guidelines for systematic reviews and meta-analyses are well established in many fields; for example, PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses[12]; http://www.prisma-statement.org/) is "an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses" and includes a checklist of 27 items and a template flow chart for the presentation of a systematic review (a 'PRISMA flow diagram'; Fig. 1a). Guidelines for developing and preparing systematic review protocols are published in PRISMA-P (http://www.prisma-statement.org/Extensions/Protocols.aspx)[13].

If the systematic review reveals sufficient and appropriate quantitative data from the studies that are being summarized, then a meta-analysis can be conducted. In a meta-analysis, one or more outcomes in the form of effect sizes are extracted from each study. Effect sizes are designed to put the outcomes of the different studies being combined on the same scale, using a suite of metrics[14,15] that includes odds and risk ratios, standardized mean differences, z-transformed correlation coefficients and logarithmic ('log') response ratios. It is essential for the effect-size metric used to be

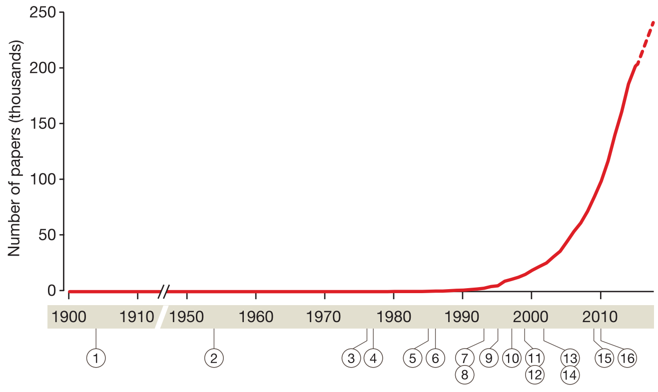
[1]Department of Ecology and Evolution, Stony Brook University, Stony Brook, New York 11794-5245, USA. [2]School of Biological Sciences, Royal Holloway University of London, Egham, Surrey, TW20 0EX, UK. [3]Evolution and Ecology Research Centre and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales 2052, Australia. [4]Diabetes and Metabolism Division, Garvan Institute of Medical Research, 384 Victoria Street, Darlinghurst, Sydney, New South Wales 2010, Australia. [5]School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK.

BOX 1

A brief history of meta-analysis

The first formal attempt to combine information from multiple sources (see figure) was made in 1904 by K. Pearson[83] with the aim of ascertaining the effectiveness of vaccination in preventing soldiers from contracting typhoid. R. A. Fisher, another important figure in the development of modern statistical science, subsequently introduced a method for combining probabilities from different studies[84]. In the late 1930s, W. Cochran and F. Yates described approaches that were essentially the same as modern fixed-effect and random-effects models[85], which were later formalized and generalized by Cochran[86]. However, it was not until the insight of psychologists G. Glass and M. Smith in 1977—that outcome measures from different experiments could be standardized and put on the same scale[87]—that meta-analysis began to affect scientific research to a large extent. Meta-analysis was initiated almost simultaneously in medicine and the social sciences[88] and was initially met in all fields with a combination of enthusiasm and condemnation[52,88]. Methodology was formalized and developed in the two decades following 1977 in multiple fields[16,89–91], with influential studies spreading from medical and social sciences to EEC in the early 1990s[23,92] (Table 1).

Rapid methodological and procedural developments have followed, with cross-disciplinary interactions being the key drivers of progress. The introduction of electronic literature databases and journal articles was central to the development of current practices; a lack of access in poorer institutions and countries hinders scientific progress. The highly interdisciplinary Society for Research Synthesis Methodology (http://www.srsm.org/) was established in 2005, after which it began publication of *Research Synthesis Methods*. The large collaborative networks the Cochrane Collaboration (established in 1993; now known as Cochrane; https://www.cochrane.org) and the Campbell Collaboration (established in 1999; https://www.campbellcollaboration.org) oversee systematic reviews in the medical and social sciences, respectively, bringing practitioners and methodologists together and setting standards for research-synthesis publications and evidence-based guidelines for practice and policy.



(1) **1904** First (medical) meta-analysis published (effect of innoculation against typhoid) (ref. 83)
(2) **1954** First meta-analytic methods formalized (fixed- and random-effects models) (ref. 86)
(3) **1976** Term 'meta-analysis' coined (ref. 95)
(4) **1977** First social science meta-analysis published (efficacy of psychotherapy) (ref. 87)
(5) **1985** Statistics textbook dedicated to meta-analytic methods released (ref. 16)
(6) **1986** Method for calculating between-study variance developed (ref. 96)
(7) **1993** Review of 302 social science meta-analyses on treatment efficacy published (ref. 97)
(8) **1993** Cochrane Collaboration established
(9) **1995** Term 'systematic review' introduced (ref. 98)
(10) **1997** Methods for assessing publication bias introduced (funnel plot and Egger's test) (ref. 19)
(11) **1999** QUOROM (Quality of Reporting of Meta-analyses) standards developed (ref. 99)
(12) **1999** Campbell Collaboration established
(13) **2002** Heterogeneity index I^2 proposed (ref. 100)
(14) **2002** Term 'network meta-analysis' coined (ref. 74)
(15) **2009** PRISMA guidelines established (ref. 12)
(16) **2010** *metafor* (free and comprehensive R package for meta-analysis) released (ref. 17)

Box 1 Figure | Milestones in the history of meta-analysis. The red line shows the number of papers from a Scopus search; the dashed component indicates the expected future trajectory. The milestone publications[12,16,17,19,74,83,86,87,95–100] are chosen on the basis of two main criteria—precedence and influence (for these criteria, we relied heavily on refs 93 and 94).

readily interpretable, scientifically meaningful and comparable among meta-analyses, and for its sampling distribution to be known, so that statistical models can be constructed appropriately.

The effect sizes are then entered into a statistical model with the goal of assessing overall effects and heterogeneity in outcomes. These models are based on an assumption of either a common effect ('fixed effect') or random effects (Fig. 1b)[16]. The common-effect (or fixed-effect) model assumes that variation in effect sizes among studies is due to within-study (sampling) variance and that all studies share a common 'true' effect. The random-effects model assumes that, in addition to sampling variance, the true effects from different studies also differ from one another, representing a random sample of a population of outcomes, and is analogous to a random-effects model in an analysis of variance (ANOVA). Thus, random-effects models include an extra variance component to account for between-study variance (heterogeneity) in addition to within-study variance. Common-effect models are based on the assumption that the results apply only to a given group of studies. Random-effects models apply more generally. In carrying out a meta-analysis, the central tendency (the mean) and its confidence limits are evaluated, as well as the heterogeneity in the effect across studies. To identify the magnitude and sources of variation in effect size among studies (Fig. 1c), earlier studies relied on simple heterogeneity tests[16], whereas more recent work often uses meta-regressions[17]. The 'main effect' or 'grand mean' can be of critical importance or largely irrelevant, depending on the goals of the meta-analysis and the magnitude and sources of heterogeneity (see sections 'Meta-analysis is essential for progress in science' and 'Meta-analysis in EEC as a case study'). Although these goals differ considerably among disciplines, quantifying heterogeneity is universally important.

Heterogeneity tests and meta-regressions both use weighting based on the precision of the estimate of the effect: larger studies with higher precision are weighted more heavily than smaller and/or more variable studies[18] (Fig. 1b, d). There are many issues to consider in constructing these statistical models, including appropriate weighting and how to account for non-independence (see sections 'Meta-analysis in EEC as a case study' and 'Limitations, controversies and challenges'). In addition, tools have been developed for evaluating publication bias and power and for conducting sensitivity analyses[19–21] (Fig. 1e, f).

Meta-analysis is essential for progress in science

Meta-analysis has generally been used with two different fundamental goals in mind, resulting in the use of contrasting approaches. The first of these goals is to assess the evidence for the effectiveness of specific interventions for a particular problem or hypothesized causal associations for a condition, often over a relatively small number of studies (fewer than about 25). The second, quite different, fundamental goal is to reach broad generalizations across larger numbers of study outcomes (dozens to hundreds) to provide a more comprehensive picture than can be attained from an individual primary study. The differences in approach

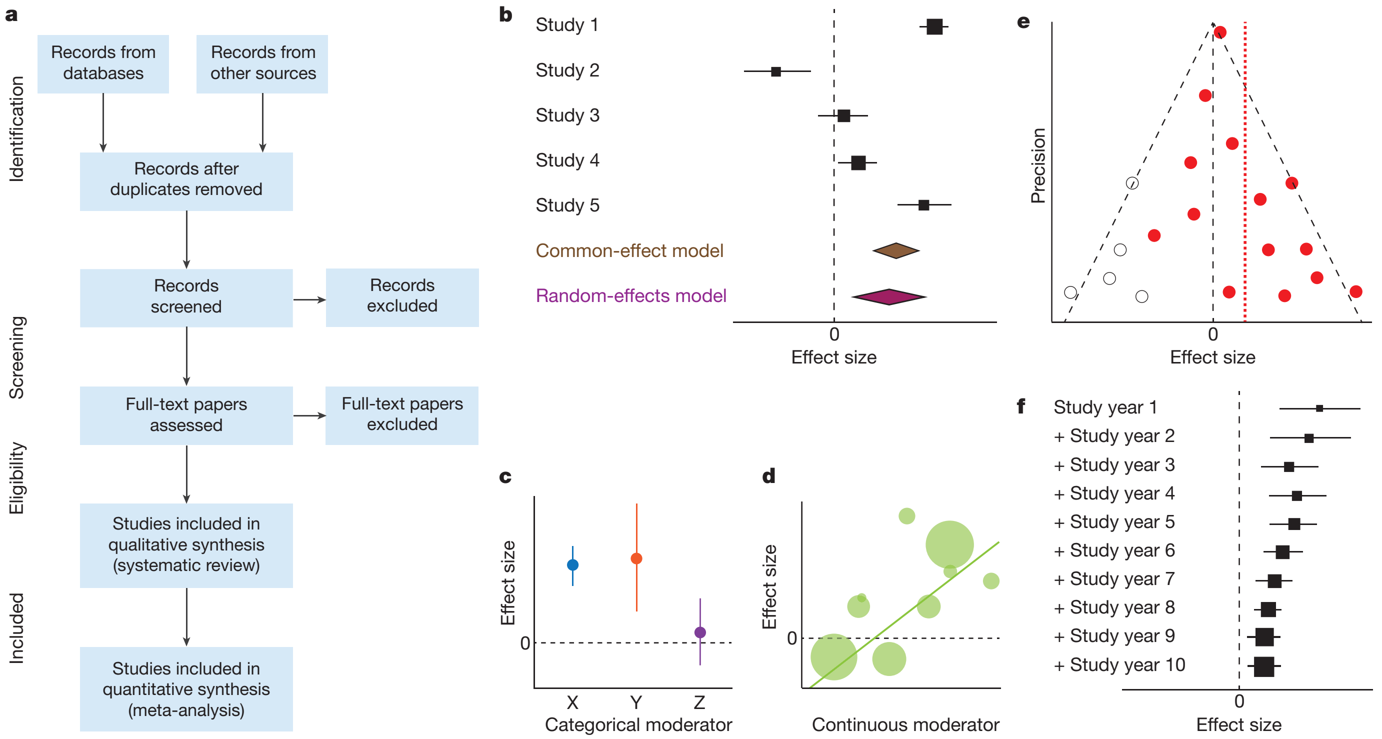


Figure 1 | Various charts and plots common to meta-analysis. **a**, A PRISMA flow diagram[12], which describes information flow (the number of relevant publications) at the four stages of the systematic review process ('identification', 'screening', 'eligibility' and 'included'). **b**, A 'forest' plot of the various means (symbol centres), confidence limits (95% confidence intervals; whiskers) and precision (indicated by the size or 'weight' of the symbols, with larger symbols indicating greater precision) of the effect-size determined from individual studies (black), and the overall means (symbol centres) and 95% confidence intervals (symbol widths) determined using meta-analysis with a common-effect (or fixed-effect) model (brown) and a random-effects model (purple). This type of plot is used to represent effect sizes and their confidence intervals graphically. **c**, A summary 'forest' plot of the mean effect sizes and 95% confidence intervals for different groups of studies. This type of plot may be used to assess categorical moderators (denoted X, Y and Z here) and are common in EEC and some social sciences. **d**, A 'bubble' plot showing a line predicted from a meta-regression analysis; the sizes of the bubbles reflect the sample sizes of the individual studies. This type of plot may be used to assess continuous predictors (such as publication year or length of a treatment). **e**, A 'funnel' plot displays the effect size against the precision with which it is estimated, which relates to its weight. Here we illustrate data (red points, with the dotted red line indicating an overall effect) that display 'funnel asymmetry', which could indicate publication bias, along with data (open circles) obtained after applying the trim-and-fill method, a sensitivity analysis that corrects for a potential publication bias. **f**, A 'forest' plot of a cumulative meta-analysis in which outcomes are added into the analysis in chronological order, demonstrating an increase in precision and a convergence of effect sizes as studies are added, and a temporal trend across studies. The dashed black lines in **b**–**f** indicate 'no effect' of an intervention on the outcome.

and goals affect not only the scale of meta-analyses, but every step of the research synthesis, from study inclusion criteria to the statistical models used. In both approaches, meta-analysis is used to synthesize evidence across studies to detect effects, to estimate their magnitudes and variation and to analyse the factors (covariates or moderators) that influence them.

When the goal is to assess evidence for specific interventions, the focus of meta-analyses is primarily on accurately estimating an overall mean effect, and may include identifying factors that modify that effect. This approach is exemplified by the PICO (population, intervention, comparator, outcome) framework (and its extensions) for formulating questions, in which specification of these elements is central to the purpose of the synthesis[22], as it is, for example, when assessing clinical effectiveness or the effectiveness of interventions in other disciplines. Question formulation using PICO has been adopted in a wide range of fields, including medicine and the social sciences. Although moderating factors might be important for understanding how the overall effect is influenced by study or population characteristics, meta-analyses for which the primary goal is to estimate the effects of a specific intervention accurately tend to emphasize the consequences of that intervention for a specific population. This type of meta-analysis must clearly and specifically delineate the population in question. Consequently, the results may apply only to that population; for example, the conclusions of a research synthesis of a medical intervention based on studies that included only middle-aged males might not apply to females or to younger males.

In the second case, when the goal is to reach broad generalizations, the population of studies may be large and heterogeneous and, although estimating the main effect of a particular phenomenon or experimental treatment may be important, identifying sources of heterogeneity in outcomes is often central to understanding the overall phenomenon[23]. Meta-analyses undertaken with the aim of reaching broad generalizations deliberately incorporate results from heterogeneous populations so that broad generalizations and the factors that modify them can be examined and tested. This approach is common in the fields of EEC and in some social sciences, in which meta-analyses have been used to address fundamental problems, to weigh the evidence for prominent theories or hypotheses and to consider the generality of common findings, observations or phenomena[23,24].

Of course, to some extent there is a continuum rather than an absolute dichotomy in meta-analytic approaches, with overlap between disciplines. A limitation of using broad inclusion criteria is the difficulty in adequately accounting for high heterogeneity. A limitation of a reductionist scope and narrow focus is the limited inference that is possible outside of a narrowly specified population or for factors that modify outcomes, whereas the inclusion of a broader definition of the population of interest and potential factors that could affect outcomes might be highly revealing. Both approaches can be limited or even biased. A collection of many narrowly focused reviews of what is essentially the same intervention can generate spurious results, as can the opposite approach of 'fishing' for significance among many hypothesized explanatory factors or covariates in an excessively broad study.

For both of these basic goals (evaluation of specific interventions or reaching a broad understanding of a general problem), meta-analysis provides a more powerful and less biased means for clarifying, quantifying and disproving (or confirming) assumed wisdom than do conventional approaches[25] including narrative reviews and flawed quantitative methods such as 'vote counts' (see section 'Limitations, controversies and challenges'). Meta-analytic methods have resolved apparently inconclusive data to arrive at a clearer picture, often more rapidly than other approaches. In medicine, meta-analyses can unambiguously assess the effectiveness of particular surgical or pharmaceutical interventions or the statistical significance of hypothesized causal associations. For example, a meta-analysis of 12 clinical studies was able to demonstrate conclusively a clear relationship between maternal obesity and risk of neural tube defects despite considerable variation in the effect sizes reported in individual studies (from a slightly greater incidence of these birth defects for overweight mothers compared to normal-weight mothers, to three times the risk (odds ratio of 3.11) for severely obese mothers compared to normal-weight mothers)[26]. Similarly, primary studies of the value of a family-based intervention approach for serious juvenile offenders called multi-systemic therapy were seemingly inconsistent; however, despite the logical and theoretical basis for multi-systemic therapy, a meta-analysis found no significant differences between it and conventional social services in the success of outcomes[27]. Both of these meta-analyses have had ramifications for evidence-based practice.

The most consequential effect of introducing formal research-synthesis methodology has been a profound change in the way scientists think about the outcomes of scientific research. An individual primary study may now be seen as a contribution towards the accumulation of evidence rather than revealing the conclusive answer to a scientific problem[25,28]. There are certainly cases where a single revelatory study has completely illuminated and resolved a major problem; however, in many cases syntheses can provide a more general and complete picture of the evidence than can any individual study. The results of initial studies are too often not confirmed by those of subsequent studies or by syntheses of a body of research. Additional major contributions of the introduction of meta-analysis have been increased attention to reporting standards in primary studies, including full and transparent reporting of data and the recognition that studies that report no significant effect are as potentially interesting and valuable as those that report low P values[29,30].

Meta-analysis in EEC as a case study

Meta-analysis was first adopted by ecologists and evolutionary biologists some 25 years ago (Table 1) and has had a considerable impact on this research field in both fundamental and applied areas. Meta-analytic approaches in ecology were introduced at around the same time as it became increasingly urgent to provide accurate quantitative assessments, predictions and practical solutions to pressing environmental issues such as biodiversity losses, the increase in invasive species and biotic responses to climate change. Meta-analysis has provided tools for summarizing evidence for these effects, their impacts and the effectiveness of interventions. The increased use of meta-analyses and systematic reviews in conservation and applied ecology has been facilitated by the promotion of evidence-based approaches in this field[31,32], especially through organizations such as the Centre for Evidence-Based Conservation (http://www.cebc.bangor.ac.uk) and the Collaboration for Environmental Evidence (https://www.environmentalevidence.org; Table 1).

Applications of meta-analyses and, more recently, systematic reviews in EEC have highlighted major gaps in research[33], provided assessments of the effects of major environmental drivers (such as climate change[34]) and of the effectiveness of conservation and management strategies[31], and enabled evaluations of the evidence for ecological and evolutionary theories[35]. Examples of influential ecological meta-analyses include quantifications of the effects of biodiversity on ecosystem functioning and services[36,37], which demonstrated that declines in species richness have negative effects on the functioning of ecosystems. It has been found[38] that ecological restoration can reverse environmental degradation and increase biodiversity and the provisioning of ecosystem services in a wide range of ecosystems globally, although not to full recovery compared to reference ecosystems.

Table 1 | Development of systematic reviews and meta-analyses in EEC

Year	Milestone
1991	First meta-analysis in ecology published[78]
1995	Seminal paper by Arnqvist and Wooster[79] published in *Trends in Ecology and Evolution*, introducing meta-analysis to many ecologists
1995	National Center for Ecological Analysis and Synthesis established in USA
1997	MetaWin, the first software for ecological meta-analysis created[46]
1999	Special feature on meta-analysis published in *Ecology*, including an influential paper on statistical issues in ecological meta-analysis[50] and the introduction of the logarithmic response ratio as a metric for effect size[80]
2001	First general review of meta-analysis in ecology published[81]
2003	Centre for Evidence-Based Conservation established in UK
2007	Collaboration for Environmental Evidence created
2008/2009	Seminal papers on phylogenetic meta-analysis published[43,45] and phyloMeta software for integrating phylogeny into meta-analyses released[82]
2011	*Environmental Evidence* (the official journal of the Collaboration for Environmental Evidence) established
2013	First handbook of meta-analysis in ecology and evolution published[73]
2014	OpenMEE, software for ecological and evolutionary meta-analysis, released[47]
2016	First international conference of the Collaboration for Environmental Evidence, in Stockholm

Similarly, meta-analytic techniques have provided evolutionary biologists the tools to test key hypotheses based on theories of natural selection, sexual selection and animal social behaviour at unprecedented scales[35]. Examples of prominent evolutionary meta-analyses include assessments of correlations between measures of genetic diversity, fitness and population size[39]. One conclusion is that a reduction in population size due to habitat fragmentation reduces genetic variation, which in turn has a negative impact on fitness in the affected populations.

In EEC, meta-analytic techniques have greatly expanded the ability to construct large-scale overviews of study outcomes—over larger spatial scales, different time periods, multiple systems and a diversity of organisms that are beyond the scope of any one researcher or research group. For example, a global meta-analysis[40] of almost 600 latitudinal gradients in species diversity verified the high degree of generality of the decline in diversity with latitude, but also identified important factors that modify this pattern. Meta-analysis has also been a valuable tool for practitioners in EEC involved in collaborative research who wish to combine original results from experiments carried out across multiple study sites[41,42].

Unlike clinical medicine and the social sciences, fields in which research focuses on a single species, the multi-species nature of much of EEC research and therefore of meta-analyses has led practitioners to integrate phylogenetic comparative methods with meta-analytic models to take into account potential non-independence among lineages due to shared evolutionary history[43–45]. Non-independence among outcomes due to the variation among sources may be more obvious in EEC than in other fields because of the large size and complex data structure of many meta-analyses in EEC. However, non-independence is a ubiquitous problem for research synthesis in most research fields, and much work remains to be done to better model and account for sources of non-independence.

The structural characteristics of data in EEC and the goals of generality typically result in high heterogeneity. Rather than seeking to explain all of the heterogeneity among studies, the goal is often to identify key factors of commonality—to detect the signals amid the noise when gaining information about these hypothesized key factors is more important than achieving a clean accounting of all sources of variability. This is a different perspective from that of meta-analyses that focus narrowly on, for example, detecting the efficacy of a specific intervention.

Advances in meta-analysis in EEC have been stimulated by many factors, including learning from practitioners in other disciplines, effective and widespread short courses for students and practising scientists, and the development of software that is tailored specifically to this field[46,47]. Methodological innovations in meta-analytic techniques that have been incorporated or developed in EEC, in addition to phylogenetic approaches, include the meta-analysis of factorial experiments[48], the introduction and wide acceptance of randomization (permutation) tests in meta-analysis[49], the early embrace of random-effects and mixed-effects models when they were still highly controversial in other disciplines[50], and methods for the inclusion of qualitative information such as expert opinions[51].

The introduction and incorporation of meta-analysis in ecological research have raised similar objections to those raised in other disciplines, and these criticisms and others have been similarly refuted across disciplines[11]. For instance, critics have claimed that the potential for publication bias in the literature (that is, the under-reporting of non-significant results or disconfirming evidence[21]) invalidates the use of meta-analysis. This objection has been refuted by research synthesists in many fields, who point out that when publication bias exists, it presents problems that are not unique to meta-analyses, but affect any attempt to summarize the results of the literature or to reach valid conclusions from it. In another instance, as in the early criticisms of meta-analysis in social sciences[52], some ecologists have claimed that ecological studies are too heterogeneous to be combined statistically in a meaningful way[9] and that ecology is best served by accumulating a catalogue of case studies[53]. Analogously, the basis for the early objections to introducing statistics to ecology in the mid-twentieth century was the inability to fully account for the uniqueness of individual organisms and the micro-site environmental variation using means and statistical tests. Despite the criticism, the introduction of meta-analysis in EEC has been embraced enthusiastically by the majority of scientists in these disciplines as a 'remote sensing tool' that helps scientists to generalize the findings of individual studies to reach a broader understanding[11], and the number of meta-analyses published in EEC has increased exponentially over time[54].

Limitations, controversies and challenges

Despite its current utility and future potential, meta-analysis has various limitations as a tool for research synthesis and for informing decisions. Meta-analyses and systematic reviews can highlight areas in which evidence is deficient, but they cannot overcome these deficiencies—they are statistical and scientific techniques, not magical ones. For example, in a systematic review of the literature on hypotheses for explaining biological invasions, a major gap was found[33] in published studies on invasive species in the tropics, highlighting not only what is known but also what is unknown globally about this problem. Although the existence of such knowledge gaps limits the generality of conclusions that can be drawn from the existing literature, the ability of systematic reviews and meta-analyses to identify these gaps is a strength of these approaches because it directs future primary studies to the areas for which evidence is most needed. Other challenges for meta-analyses and systematic reviews include publication bias and research bias[50], the latter describing the over- or under-representation of populations, species or systems in the literature, which results in a biased view of the totality. The presence of these issues can be strongly suspected by scientists, but although their magnitude can sometimes be estimated in a meta-analysis[19,20], it cannot be truly corrected in research syntheses[55,56]. Similarly, a synthesis may be constrained by either selective or incomplete data reporting in primary publications[30].

One undesirable consequence of the growing recognition and high impact of meta-analysis is an increase in less-than-rigorous applications of these methods and in the application of arbitrary and less-well-justified methodologies that are sometimes inaccurately referred to as meta-analyses. The use of statistically flawed approaches can lead to erroneous and misleading results that masquerade as serious research syntheses. The term meta-analysis should be applied only to studies that use well-established statistical procedures, such as appropriate effect-size calculation, weighting and heterogeneity analysis[57], and statistical models that take into account the distinct hierarchical structure of meta-analytic data, or to studies that develop rigorously justified methodological advances of these methods. Unfortunately, the term is often misapplied to any study that uses data from several primary publications, regardless of the rigour of the methodology. Statistically flawed procedures such as vote-counting, which provide only limited information about study outcomes, can be very misleading and have long been discredited, are still used in published papers[6,50]. Vote-counting is a deceptively plausible and appealingly convenient procedure whereby the generality of findings in a group of studies is assessed by counting up the number of significant and non-significant results in individual studies (or by elaborations on this approach). Although it is vulnerable to erroneous inferences and provides unreliable information on the magnitudes or heterogeneity of effects, it persists, zombie-like, returning by the efforts of the naive or determinedly ignorant to haunt the scientific literature. Vote-counting is not a meta-analytic technique, and is not an acceptable basis for meaningfully summarizing research results in published papers.

Meta-analyses that are not weighted by inverse variances are common and often poorly justified, and present different problems. Unlike vote-counts, unweighted meta-analyses can be unbiased and may provide information on the magnitude of the effects[8]. However, in an unweighted analysis, within- and between-study variation cannot be readily separated, and so common- and random-effects models cannot be used and heterogeneity may be difficult to assess properly. Unweighted meta-analysis also increases the influence of small studies[29], which have often been found to report larger and more variable effects than those reported for larger studies (as a result of the smaller studies being more likely to suffer from random noise, and possibly publication bias). An alternative when variances are unavailable from primary studies is weighting by sample size or other metric, but this method does not incorporate the information that an inverse-variance-weighted analysis provides and can introduce unknown biases. These problems are particularly acute with small sample sizes. One argument that is often made in support of unweighted meta-analysis is that the variances needed for a weighted meta-analysis are frequently unavailable owing to poor data reporting in the primary studies, and it is undesirable to leave studies with missing data out of the meta-analysis. One possible solution is to use one of the various methods that have been developed for imputing or otherwise modelling missing data. And, although data reporting practices are being improved slowly, it may be that many older studies are simply inadequate for accurate quantitative reviews. Another argument for unweighted meta-analysis is that the meta-analysis simplifies to an essentially unweighted analysis when between-study variation is much larger than within-study variation[58]. However, a weighted meta-analysis is required to assess the two types of variation in the first place, and we submit that it would be preferable to report the weighted and unweighted results in such cases.

Another unfortunate outcome of the high impact and growing prestige of meta-analysis[59], coupled with the use of metrics such as citation numbers and *h*-indices in evaluations of research accomplishments, is an unease among some primary researchers about the fairness and rewards of the scientific process[8,60]. Some have decried reviews as "the black-market of scientific currency"[61], with calls to replace citations to reviews and meta-analyses with citations of primary studies[61]. Worse, research synthesists in medicine have recently been described as "research parasites"[62] of primary studies and the researchers who conduct them. On the other hand, it could be argued that primary studies without context, comparison or summary are ultimately of limited value. Moreover, methods for research synthesis are not the exclusive province of any one group, but can be used by primary researchers in their own areas of expertise. The introduction of more explicit guidelines and standards for conducting and reporting meta-analyses could address some of these grievances, and we agree that better methods for citing primary studies in meta-analyses should be implemented to give full credit for the original studies. 'Research parasites' can also serve to increase scientific diversity by adding another 'trophic level', thus improving the functioning of the scientific 'ecosystem'.

Advances, developments and future promise

Meta-analysis is the grandmother of the 'big data' and 'open science' movements. For hundreds of years, scientists have collected data in individual studies, based on observations and experimentation[63]. The introduction and implementation of meta-analytic techniques was the first large-scale, coordinated effort to collect and synthesize pre-existing data to determine patterns, make predictions, reach generalizations and make evidence-based decisions. Discoveries that have resulted from the analysis of big data, in parallel with the development of open-science practices, transparency and the importance of replication of research, are transforming many research areas. 'Big data' refers to large, complex datasets that may be mined for patterns or for making predictions, and has been influential in a broad range of areas (for example, genomics, climatology and advertising). The processes involved in the searching, curation and evaluation of data, and in quality control, are essential components of big-data practice, all of which have been the subject of conceptual exploration and formal methodological development in meta-analysis for many years[64]. However, the approach has been different from that taken for meta-analyses. Meta-analysis is inherently statistical, whereas big data has been framed within the field of computer science. Greater cross-disciplinary interactions should prove productive for both fields. Although formal systematic reviews and meta-analyses have long been established in many disciplines, they are only recently making inroads in fields such as molecular biology and genomics. Rapid gains in scientific progress stand to be made when these methods are more fully implemented throughout the biological sciences, and throughout science more generally.

Open-science practices have emphasized full and unbiased access to scientific data[65], which is of longstanding importance and central to future progress in meta-analysis. Pre-registration (called 'registration' in some fields) of planned studies can reduce selective reporting of outcomes; publication of 'registered reports' in which the methods and proposed analyses for a study are peer-reviewed and published before the research is conducted can reduce publication bias. Limitations on accessing information are serious impediments for best practices in meta-analysis. By minimizing selective and poor reporting and advocating full access to the data and code associated with each analysis, open-science standards, including guidelines such as those in the Equator Network (https://www.equator-network.org)[30,66] can alleviate many problems in research synthesis and propel more rapid scientific advances.

In addition to the benefits that have been accrued from the increased availability of unbiased information, advances in meta-analytic techniques are being driven by methodological developments. Advances include: the use of machine learning and artificial intelligence (AI) to screen studies for inclusion in systematic reviews and meta-analyses[67]; increasingly sophisticated software and models for complex meta-regression[17,47]; robust variance estimation in studies with small sample sizes[68]; meta-analysis of individual participant data; and integration of meta-analysis and decision support in medicine and other fields[69]. Bayesian meta-analysis has been implemented in many fields and is a particularly useful approach when external sources of information can provide valid priors[70] or when a dataset is of sufficient quality and size that distributions can be fitted to it instead of attempting to fit it to familiar distributions. Meta-analytic approaches have been used to synthesize data to address methodological issues such as heterogeneity and its interpretation[71] and the implications of the inclusion or exclusion of unpublished literature[72]. Better integration of big data, AI and meta-analysis will depend on both conceptual and methodological developments, and is reliant on greater trans-disciplinary links between statistics, computer science, the biological and social sciences, and other scientific fields. It is not impossible to envisage automated systems whereby AI aids not only in the real-time acquisition but also in the critical appraisal and meta-analysis of data, potentially integrating different information streams to inform tailored decisions in many areas of applied science.

The statistical methodologies that underpin and support meta-analysis have been undergoing continual development. Areas of particular current interest include multiple imputation to model missing data, advanced use of meta-regression and model selection to evaluate the influence of more complex data structures and multiple covariates, and hierarchical modelling of multi-level data, including that from individual 'participant' data in medicine[22] and in EEC[73]. Network meta-analyses seek to provide comparisons of multiple interventions, including indirect comparisons[74]. These methods are particularly useful when a set of randomized control trials with pairwise comparisons of interventions has been carried out with common interventions among the studies, but when not all studies include all interventions. Developments in and applications of this powerful approach have advanced considerably in clinical medicine over the past ten years[75], providing better information about which treatment is most effective when there are multiple treatment options and pathways. 'Living' reviews, which are constantly updated, can prevent stale information from being cemented into belief or practice and have the potential to change the fundamental understanding of a problem or approach, because knowledge is being updated and new papers are being published continuously[76]. Rather than summarizing information in many individual reviews, living reviews and living cumulative network meta-analyses may also help to reduce waste in research by using the available primary studies more efficiently, by identifying gaps in research and by determining when the evidence is sufficient for decision and policy making[77]. However, their full implementation might require a reward shift both for primary researchers and synthesists.

Perhaps the most important foundation for advances in meta-analytic techniques is education in high-quality research-synthesis methods. Training in meta-analytic methods and concepts should be part of the basic training for higher-degree candidates in basic and applied scientific fields, including research post-graduates, medical doctors and other professional science practitioners (such as environmental consultants). This would formally embed their work in the context of existing evidence and facilitate learning of both statistical and critical appraisal skills. Those involved in primary research also need a better understanding of meta-analysis to exploit the revolution of open data fully. Most importantly, a new generation of scientists, peer reviewers, editors and science-policy practitioners would benefit from an increased understanding of the methodologies and interpretation of evidence synthesis.

Meta-analysis can be a key tool for facilitating rapid progress in science by quantifying what is known and identifying what is not yet known. Evidence synthesis should become a regular companion to primary scientific research to maximize the effectiveness of scientific inquiry. An evidence-based approach is important for progress in science, policy, and medical and conservation practice. This will require collaboration between statisticians, primary researchers and research synthesists, between meta-analysts and stakeholders, and among research synthesists across different disciplines. We are confident that, provided such collaborations are successful, meta-analysis will survive its 'midlife crisis' and emerge stronger and with a new-found purpose.

Received 4 March 2017; accepted 12 January 2018.

1. Jennions, M. D., Lortie, C. J. & Koricheva, J. in *The Handbook of Meta-analysis in Ecology and Evolution* (eds Koricheva, J. *et al.*) Ch. 23, 364–380 (Princeton Univ. Press, 2013).
2. Roberts, P. D., Stewart, G. B. & Pullin, A. S. Are review articles a reliable source of evidence to support conservation and environmental management? A comparison with medicine. *Biol. Conserv.* **132,** 409–423 (2006).
3. Bastian, H., Glasziou, P. & Chalmers, I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med.* **7,** e1000326 (2010).

4. Borman, G. D. & Grigg, J. A. in *The Handbook of Research Synthesis and Meta-analysis* 2nd edn (eds Cooper, H. M. *et al.*) 497–519 (Russell Sage Foundation, 2009).
5. Ioannidis, J. P. A. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q.* **94,** 485–514 (2016).
6. Koricheva, J. & Gurevitch, J. Uses and misuses of meta-analysis in plant ecology. *J. Ecol.* **102,** 828–844 (2014).
7. Littell, J. H. & Shlonsky, A. Making sense of meta-analysis: a critique of "effectiveness of long-term psychodynamic psychotherapy". *Clin. Soc. Work J.* **39,** 340–346 (2011).
8. Morrissey, M. B. Meta-analysis of magnitudes, differences and variation in evolutionary parameters. *J. Evol. Biol.* **29,** 1882–1904 (2016).
9. Whittaker, R. J. Meta-analyses and mega-mistakes: calling time on meta-analysis of the species richness-productivity relationship. *Ecology* **91,** 2522–2533 (2010).
10. Begley, C. G. & Ellis, L. M. Drug development: Raise standards for preclinical cancer research. *Nature* **483,** 531–533 (2012); clarification **485,** 41 (2012).
11. Hillebrand, H. & Cardinale, B. J. A critique for meta-analyses and the productivity-diversity relationship. *Ecology* **91,** 2545–2549 (2010).
12. Moher, D. *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. **6,** e1000097 (2009).
This paper provides a consensus regarding the reporting requirements for medical meta-analysis and has been highly influential in ensuring good reporting practice and standardizing language in evidence-based medicine, with further guidance for protocols, individual patient data meta-analyses and animal studies.
13. Moher, D. *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst. Rev*. **4,** 1 (2015).
14. Nakagawa, S. & Santos, E. S. A. Methodological issues and advances in biological meta-analysis. *Evol. Ecol.* **26,** 1253–1274 (2012).
15. Nakagawa, S., Noble, D. W. A., Senior, A. M. & Lagisz, M. Meta-evaluation of meta-analysis: ten appraisal questions for biologists. *BMC Biol.* **15,** 18 (2017).
16. Hedges, L. & Olkin, I. *Statistical Methods for Meta-analysis* (Academic Press, 1985).
17. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **36,** 1–48 (2010).
18. Anzures-Cabrera, J. & Higgins, J. P. T. Graphical displays for meta-analysis: an overview with suggestions for practice. *Res. Synth. Methods* **1,** 66–80 (2010).
19. Egger, M., Davey Smith, G., Schneider, M. & Minder, C. Bias in meta-analysis detected by a simple, graphical test. *Br. Med. J*. **315,** 629–634 (1997).
20. Duval, S. & Tweedie, R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* **56,** 455–463 (2000).
21. Leimu, R. & Koricheva, J. Cumulative meta-analysis: a new tool for detection of temporal trends and publication bias in ecology. *Proc. R. Soc. Lond. B* **271,** 1961–1966 (2004).
22. Higgins, J. P. T. & Green, S. (eds) *Cochrane Handbook for Systematic Reviews of Interventions: Version 5.1.0* (Wiley, 2011).
This large collaborative work provides definitive guidance for the production of systematic reviews in medicine and is of broad interest for methods development outside the medical field.
23. Lau, J., Rothstein, H. R. & Stewart, G. B. in *The Handbook of Meta-analysis in Ecology and Evolution* (eds Koricheva, J. *et al.*) Ch. 25, 407–419 (Princeton Univ. Press, 2013).
24. Lortie, C. J., Stewart, G., Rothstein, H. & Lau, J. How to critically read ecological meta-analyses. *Res. Synth. Methods* **6,** 124–133 (2015).
25. Murad, M. H. & Montori, V. M. Synthesizing evidence: shifting the focus from individual studies to the body of evidence. *J. Am. Med. Assoc.* **309,** 2217–2218 (2013).
26. Rasmussen, S. A., Chu, S. Y., Kim, S. Y., Schmid, C. H. & Lau, J. Maternal obesity and risk of neural tube defects: a meta-analysis. *Am. J. Obstet. Gynecol.* **198,** 611–619 (2008).
27. Littell, J. H., Campbell, M., Green, S. & Toews, B. Multisystemic therapy for social, emotional, and behavioral problems in youth aged 10–17. *Cochrane Database Syst. Rev*. https://doi.org/10.1002/14651858.CD004797.pub4 (2005).
28. Schmidt, F. L. What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *Am. Psychol*. **47,** 1173–1181 (1992).
29. Button, K. S. *et al.* Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* **14,** 365–376 (2013); erratum **14,** 451 (2013).
30. Parker, T. H. *et al.* Transparency in ecology and evolution: real problems, real solutions. *Trends Ecol. Evol.* **31,** 711–719 (2016).
31. Stewart, G. Meta-analysis in applied ecology. *Biol. Lett.* **6,** 78–81 (2010).
32. Sutherland, W. J., Pullin, A. S., Dolman, P. M. & Knight, T. M. The need for evidence-based conservation. *Trends Ecol. Evol.* **19,** 305–308 (2004).
33. Lowry, E. *et al.* Biological invasions: a field synopsis, systematic review, and database of the literature. *Ecol. Evol.* **3,** 182–196 (2013).
34. Parmesan, C. & Yohe, G. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421,** 37–42 (2003).
35. Jennions, M. D., Lortie, C. J. & Koricheva, J. in *The Handbook of Meta-analysis in Ecology and Evolution* (eds Koricheva, J. *et al.*) Ch. 24, 381–403 (Princeton Univ. Press, 2013).
36. Balvanera, P. *et al.* Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecol. Lett.* **9,** 1146–1156 (2006).
37. Cardinale, B. J. *et al.* Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* **443,** 989–992 (2006).
38. Rey Benayas, J. M., Newton, A. C., Diaz, A. & Bullock, J. M. Enhancement of biodiversity and ecosystem services by ecological restoration: a meta-analysis. *Science* **325,** 1121–1124 (2009).
39. Leimu, R., Mutikainen, P. I. A., Koricheva, J. & Fischer, M. How general are positive relationships between plant population size, fitness and genetic variation? *J. Ecol.* **94,** 942–952 (2006).
40. Hillebrand, H. On the generality of the latitudinal diversity gradient. *Am. Nat.* **163,** 192–211 (2004).
41. Gurevitch, J. in *The Handbook of Meta-analysis in Ecology and Evolution* (eds Koricheva, J. *et al.*) Ch. 19, 313–320 (Princeton Univ. Press, 2013).
42. Rustad, L. *et al.* A meta-analysis of the response of soil respiration, net nitrogen mineralization, and aboveground plant growth to experimental ecosystem warming. *Oecologia* **126,** 543–562 (2001).
43. Adams, D. C. Phylogenetic meta-analysis. *Evolution* **62,** 567–572 (2008).
44. Hadfield, J. D. & Nakagawa, S. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *J. Evol. Biol.* **23,** 494–508 (2010).
45. Lajeunesse, M. J. Meta-analysis and the comparative phylogenetic method. *Am. Nat.* **174,** 369–381 (2009).
46. Rosenberg, M. S., Adams, D. C. & Gurevitch, J. *MetaWin: Statistical Software for Meta-Analysis with Resampling Tests* Version 1 (Sinauer Associates, 1997).
47. Wallace, B. C. *et al.* OpenMEE: intuitive, open-source software for meta-analysis in ecology and evolutionary biology. *Methods Ecol. Evol.* **8,** 941–947 (2016).
48. Gurevitch, J., Morrison, J. A. & Hedges, L. V. The interaction between competition and predation: a meta-analysis of field experiments. *Am. Nat.* **155,** 435–453 (2000).
49. Adams, D. C., Gurevitch, J. & Rosenberg, M. S. Resampling tests for meta-analysis of ecological data. *Ecology* **78,** 1277–1283 (1997).
50. Gurevitch, J. & Hedges, L. V. Statistical issues in ecological meta-analyses. *Ecology* **80,** 1142–1149 (1999).
51. Schmid, C. H. & Mengersen, K. in *The Handbook of Meta-analysis in Ecology and Evolution* (eds Koricheva, J. *et al.*) Ch. 11, 145–173 (Princeton Univ. Press, 2013).
52. Eysenck, H. J. Exercise in mega-silliness. *Am. Psychol*. **33,** 517 (1978).
53. Simberloff, D. Rejoinder to: Don't calculate effect sizes; study ecological effects. *Ecol. Lett.* **9,** 921–922 (2006).
54. Cadotte, M. W., Mehrkens, L. R. & Menge, D. N. L. Gauging the impact of meta-analysis on ecology. *Evol. Ecol.* **26,** 1153–1167 (2012).
55. Koricheva, J., Jennions, M. D. & Lau, J. in *The Handbook of Meta-analysis in Ecology and Evolution* (eds Koricheva, J. *et al.*) Ch. 15, 237–254 (Princeton Univ. Press, 2013).
56. Lau, J., Ioannidis, J. P. A., Terrin, N., Schmid, C. H. & Olkin, I. The case of the misleading funnel plot. *Br. Med. J.* **333,** 597–600 (2006).
57. Vetter, D., Rucker, G. & Storch, I. Meta-analysis: a need for well-defined usage in ecology and conservation biology. *Ecosphere* **4,** 1–24 (2013).
58. Mengersen, K., Jennions, M. D. & Schmid, C. H. in *The Handbook of Meta-analysis in Ecology and Evolution* (eds Koricheva, J. *et al.*) Ch. 16, 255–283 (Princeton Univ. Press, 2013).
59. Patsopoulos, N. A., Analatos, A. A. & Ioannidis, J. P. A. Relative citation impact of various study designs in the health sciences. *J. Am. Med. Assoc.* **293,** 2362–2366 (2005).
60. Kueffer, C. *et al.* Fame, glory and neglect in meta-analyses. *Trends Ecol. Evol.* **26,** 493–494 (2011).
61. Cohnstaedt, L. W. & Poland, J. Review Articles: The black-market of scientific currency. *Ann. Entomol. Soc. Am.* **110,** 90 (2017).
62. Longo, D. L. & Drazen, J. M. Data sharing. *N. Engl. J. Med.* **374,** 276–277 (2016).
63. Gauch, H. G. *Scientific Method in Practice* (Cambridge Univ. Press, 2003).
64. Science Staff. Dealing with data: introduction. Challenges and opportunities. *Science* **331,** 692–693 (2011).
65. Nosek, B. A. *et al.* Promoting an open research culture. *Science* **348,** 1422–1425 (2015).
66. Stewart, L. A. *et al.* Preferred reporting items for a systematic review and meta-analysis of individual participant data: the PRISMA-IPD statement. *J. Am. Med. Assoc.* **313,** 1657–1665 (2015).
67. Saldanha, I. J. *et al.* Evaluating Data Abstraction Assistant, a novel software application for data abstraction during systematic reviews: protocol for a randomized controlled trial. *Syst. Rev.* **5,** 196 (2016).
68. Tipton, E. & Pustejovsky, J. E. Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *J. Educ. Behav. Stat.* **40,** 604–634 (2015).
69. Mengersen, K., MacNeil, M. A. & Caley, M. J. The potential for meta-analysis to support decision analysis in ecology. *Res. Synth. Methods* **6,** 111–121 (2015).
70. Ashby, D. Bayesian statistics in medicine: a 25 year review. *Stat. Med.* **25,** 3589–3631 (2006).
71. Senior, A. M. *et al.* Heterogeneity in ecological and evolutionary meta-analyses: its magnitude and implications. *Ecology* **97,** 3293–3299 (2016).
72. McAuley, L., Pham, B., Tugwell, P. & Moher, D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *Lancet* **356,** 1228–1231 (2000).
73. Koricheva, J., Gurevitch, J. & Mengersen, K. (eds) *The Handbook of Meta-Analysis in Ecology and Evolution* (Princeton Univ. Press, 2013).

This book provides the first comprehensive guide to undertaking meta-analyses in ecology and evolution and is also relevant to other fields where heterogeneity is expected, incorporating explicit consideration of the different approaches used in different domains.
74. Lumley, T. Network meta-analysis for indirect treatment comparisons. *Stat. Med.* **21,** 2313–2324 (2002).
75. Zarin, W. *et al.* Characteristics and knowledge synthesis approach for 456 network meta-analyses: a scoping review. *BMC Med.* **15,** 3 (2017).
76. Elliott, J. H. *et al.* Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. *PLoS Med.* **11,** e1001603 (2014).
77. Vandvik, P. O., Brignardello-Petersen, R. & Guyatt, G. H. Living cumulative network meta-analysis to reduce waste in research: a paradigmatic shift for systematic reviews? *BMC Med.* **14,** 59 (2016).
78. Jarvinen, A. 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 (1991).
79. Arnqvist, G. & Wooster, D. Meta-analysis: synthesizing research findings in ecology and evolution. *Trends Ecol. Evol.* **10,** 236–240 (1995).
80. Hedges, L. V., Gurevitch, J. & Curtis, P. S. The meta-analysis of response ratios in experimental ecology. *Ecology* **80,** 1150–1156 (1999).
81. Gurevitch, J., Curtis, P. S. & Jones, M. H. Meta-analysis in ecology. *Adv. Ecol. Res* **32,** 199–247 (2001).
82. Lajeunesse, M. J. phyloMeta: a program for phylogenetic comparative analyses with meta-analysis. *Bioinformatics* **27,** 2603–2604 (2011).
83. Pearson, K. Report on certain enteric fever inoculation statistics. *Br. Med. J.* **2,** 1243–1246 (1904).
84. Fisher, R. A. *Statistical Methods for Research Workers* (Oliver and Boyd, 1925).
85. Yates, F. & Cochran, W. G. The analysis of groups of experiments. *J. Agric. Sci*. **28,** 556–580 (1938).
86. Cochran, W. G. The combination of estimates from different experiments. *Biometrics* **10,** 101–129 (1954).
87. Smith, M. L. & Glass, G. V. Meta-analysis of psychotherapy outcome studies. *Am. Psychol.* **32,** 752–760 (1977).
88. Glass, G. V. Meta-analysis at middle age: a personal history. *Res. Synth. Methods* **6,** 221–231 (2015).
89. Cooper, H. M., Hedges, L. V. & Valentine, J. C. (eds) *The Handbook of Research Synthesis and Meta-analysis* 2nd edn (Russell Sage Foundation, 2009).
This book is an important compilation that builds on the ground-breaking first edition to set the standard for best practice in meta-analysis, primarily in the social sciences but with applications to medicine and other fields.
90. Rosenthal, R. *Meta-analytic Procedures for Social Research* (Sage, 1991).
91. Hunter, J. E., Schmidt, F. L. & Jackson, G. B. *Meta-analysis: Cumulating Research Findings Across Studies* (Sage, 1982).
92. Gurevitch, J., Morrow, L. L., Wallace, A. & Walsh, J. S. A meta-analysis of competition in field experiments. *Am. Nat.* **140,** 539–572 (1992).
This influential early ecological meta-analysis reports multiple experimental outcomes on a longstanding and controversial topic that introduced a wide range of ecologists to research synthesis methods.
93. O'Rourke, K. An historical perspective on meta-analysis: dealing quantitatively with varying study results. *J. R. Soc. Med.* **100,** 579–582 (2007).
94. Shadish, W. R. & Lecy, J. D. The meta-analytic big bang. *Res. Synth. Methods* **6,** 246–264 (2015).
95. Glass, G. V. Primary, secondary, and meta-analysis of research. *Educ. Res.* **5,** 3–8 (1976).
96. DerSimonian, R. & Laird, N. Meta-analysis in clinical trials. *Control. Clin. Trials* **7,** 177–188 (1986).
97. Lipsey, M. W. & Wilson, D. B. The efficacy of psychological, educational, and behavioral treatment. Confirmation from meta-analysis. *Am. Psychol.* **48,** 1181–1209 (1993).
98. Chalmers, I. & Altman, D. G. *Systematic Reviews* (BMJ Publishing Group, 1995).
99. Moher, D. *et al.* Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of reporting of meta-analyses. *Lancet* **354,** 1896–1900 (1999).
100. Higgins, J. P. & Thompson, S. G. Quantifying heterogeneity in a meta-analysis. *Stat. Med.* **21,** 1539–1558 (2002).

Acknowledgements We dedicate this Review to the memory of Ingram Olkin and William Shadish, founding members of the Society for Research Synthesis Methodology who made tremendous contributions to the development of meta-analysis and research synthesis and to the supervision of generations of students. We thank L. Lagisz for help in preparing the figures. We are grateful to the Center for Open Science and the Laura and John Arnold Foundation for hosting and funding a workshop, which was the origination of this article. S.N. is supported by Australian Research Council Future Fellowship (FT130100268). J.G. acknowledges funding from the US National Science Foundation (ABI 1262402).

Author Contributions All authors contributed equally in designing the study and writing the manuscript, and so are listed alphabetically.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to J.G. (jessica.gurevitch@stonybrook.edu), J.K. (julia.koricheva@rhul.ac.uk), S.N. (s.nakagawa@unsw.edu.au) or G.S. (gavin.stewart@newcastle.ac.uk).

Reviewer Information *Nature* thanks D. Altman, M. Lajeunesse, D. Moher and G. Romero for their contribution to the peer review of this work.