Meta-analysis and the science of research synthesis

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

1 Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY 11794-5245, USA

2 School of Biological Sciences, Royal Holloway University of London, Egham, Surrey, TW20 0EX, UK

3 Evolution & Ecology Research Centre and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia

4 Diabetes and Metabolism Division, Garvan Institute of Medical Research, 384 Victoria Street, Darlinghurst, Sydney, NSW 2010, Australia

5 School of Agriculture, Food and Rural Development, Newcastle University, Newcastle upon Tyne

NE1 7RU, UK

*all the authors contributed equally to this work and the authors are listed alphabetically

Author Contributions: All authors designed the study and wrote the manuscript.

Author Information: 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)

Preface

Meta-analysis, the quantitative synthesis of research results first introduced in the 1970s, has helped to establish evidence-based practice, resolve seemingly contradictory results, and has had a revolutionary impact on many scientific fields. At the same time, its implementation has engendered various criticisms and controversies, including mixing apples and oranges, and the proper relationship between primary researchers and research synthesists. The recent 40th anniversary of meta-analysis provides a timely opportunity to reflect on the accomplishments, limitations, recent advances, and the direction of future developments in the field of research synthesis.

(Introduction)

Synthesizing results across studies to reach an overall understanding of a problem and identify sources of variation in outcomes is an essential part of the scientific process. For over a century narrative reviews have been the major approach to summarizing the results of scientific findings. However, as the number of publications and research results has mushroomed across scientific fields, the difficulties encountered in carrying out narrative reviews to identify and summarize evidence in a transparent manner has become increasingly apparent[1]. Finding, selecting and summarizing the results of a dozen studies to provide evidence for the efficacy of a specific intervention or the generality of a particular phenomenon may (or may not) be possible informally and verbally. When the number of studies is an order of magnitude greater than that, the inability of narrative processes to adequately summarize the outcomes increases exponentially[2,3]. Narrative reviews are also subject to biases as a consequence of subjective and opaque decision processes, and conclusions can be subtly distorted by issues of statistical power (as is the case with vote counts, discussed below).

Over the last few decades, systematic reviews with transparent protocols to ensure reproducibility and reduce bias have become more prevalent[2]. Systematic reviews include Methods and Results sections, similar to primary scientific studies. Systematic reviews, field synopses and scoping reviews can robustly provide an overview of a field, for example, to compare where studies have been conducted, how many were published per year, study duration, and other study characteristics, but without being combined with formal meta-analysis cannot summarize study outcomes accurately. Narrative reviews remain useful for exploring the development of particular ideas (such as we do in this paper) or conceptual frameworks, but they fall short as a tool for summarizing results across studies, although in some cases they can be combined with systematic reviews and meta-analyses[4]. However, neither systematic reviews (alone, without meta-analyses) nor narrative reviews can provide unbiased information on the efficacy or magnitude of a technique or the generality of a phenomenon, in contrast to meta-analyses.

Formal statistical meta-analysis methods have been developed to put the results of different studies on the same scale and combine them using various statistical models[5-8]. The term “meta-analysis” was first introduced by Gene Glass in 1976[9]. The proposed approach was met with a mixture of great enthusiasm and condemnation[10,11]. Considerable methodological advances have been made in meta-analysis over the past four decades, and the approach has rapidly spread across and within many disciplines, leading to large changes in the way research results are viewed and evaluated in medicine, epidemiology, social sciences, ecology and evolutionary biology, and other disciplines. While there are general similarities in the application of these approaches across disciplines, the methodology, emphases, and terminology vary somewhat with the research field and the goals of the synthesis[12].

Forty years after its introduction, we are seeing both widespread mainstream acceptance of meta-analysis as a research tool, and also what may be considered a 'meta-analytic midlife crisis.' While the number of published meta-analyses has continued to increase rapidly within and among disciplines, too many meta-analyses are of low quality[13-15]. The publication of meta-analyses that are poorly carried out indicates that peer reviewers, editors, and authors are not fully aware of or are indifferent to the large body of well-developed meta-analysis methodology, or feel unqualified to address methodological issues in their reviews. Inadequate meta-analyses have attracted strong criticism[13,16] and even calls for a halt in publication of meta-analyses[17]. By contrast, criticism of flawed methodology and applications in other fields typically prompts calls for improvement rather than demolition of the field (e.g., papers noting serious problems with pre-clinical cancer studies have not called for an end to pre-cancer clinical research but rather for its improvement[18]). While it is certainly both valid and valuable to criticise poor methodology and studies that are poorly carried out, abandonment of a field due to flaws in some published papers is truly throwing the baby out with the bathwater[19]. We believe that the solution lies instead in rigorous application of stricter quality criteria for published meta-analyses (e.g., Tools for Transparency in Ecology and Evolution, TTEE: osf.io/g65cb). We highlight some of the main principles and characteristics of meta-analytic methodology in this review. We also demonstrate both the limitations and the utility and achievements made by applications of meta-analysis in several fields, and highlight its role in advances in ecology, evolutionary biology and conservation (EEC) as a case study. Finally, we address several recent criticisms of the meta-analytic approach and suggest ways that future developments in the field of research synthesis can facilitate the most rapid progress in the fields in which it is employed.

Systematic reviews and meta-analyses employ well-documented methodologies

The systematic review process includes formal protocols for the literature search, study screening according to pre-determined criteria, data extraction and coding, and often statistical analysis (i.e. meta-analysis) along with detailed, transparent documentation of each step. Procedural protocols for systematic reviews are well established especially in the medical and social sciences (Preferred Reporting Items for Systematic Reviews, or PRISMA; www.prisma-statement.org)[20]. The PRISMA statement includes a checklist of 27 items for presentation and a flow chart (often referred to as PRISMA diagram; Fig. 1a), which aims to promote the transparency of each procedural step of a systematic review. A major goal for systematic reviews is to carry out reviews that are reproducible and updatable using scientific methodology.

Well-established methods for carrying out meta-analyses generally depend on accounting for sampling variation within studies and on sources of variation among studies. These methods include statistical tools for determining appropriate measures of outcome that are comparable among studies, weighting by study precision, and modelling variation among studies (Box 1; Fig1b,c,d). In addition, tools have been developed for quantifying heterogeneity (between-study variance; Box 1), evaluating publication bias and power, and conducting sensitivity analyses (Fig1e,f).

Systematic reviews and meta-analysis are essential for scientific progress

Meta-analysis has generally been used for two different fundamental goals, employing 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, usually over a relatively small number (<25) of studies. The second, quite different, fundamental goal is to reach broad generalizations across larger numbers (dozens to hundreds) of studies to provide a more comprehensive picture than is possible in an individual primary study. Both goals begin with the same tools and diverge in interesting ways, and are increasingly central in different disciplines. The differences in approach and goals affect not only the size 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, estimate their magnitudes, and analyse the factors influencing those effects.

Where the goal is to assess evidence for specific interventions, the focus is 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) in question formulation, where specification of these elements is central to the purpose of the synthesis[21,22]. This has been essential for assessment of the effectiveness of interventions, and for informing evidence-based clinical decisions. While moderating factors may be important to understanding how this overall effect is influenced by study or population characteristics, the goal is generally to focus on the consequences of implementing a specific intervention for a specific population. This implies clearly delineating that population, very specifically and often narrowly.

In the second case, where 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 a major goal, identifying sources of heterogeneity in outcomes is often central to understanding the overall phenomenon[12]. Such meta-analyses deliberately incorporate results on heterogeneous populations so that broad generalizations and the factors modifying them can be examined and tested. This approach is common in the fields of EEC, and in some social sciences, where meta-analysis has been used to address fundamental problems, weigh the evidence for prominent theories or hypotheses, and document the generality of important phenomena[12,23]. Of course, to some extent there is a continuum rather than an absolute dichotomy in approaches, with overlap between disciplines.

For both of these two basic goals, meta-analysis has been a more powerful and less biased means for clarifying, quantifying and disproving (or confirming) assumed wisdom than have previous conventional approaches[24], such as narrative reviews and flawed quantitative methods such as ‘vote counts’ (discussed below). Meta-analytic methods have resolved apparently inconclusive data to arrive at a clearer picture, often sooner than other approaches. In medicine, meta-analyses can provide unambiguous evidence for the effectiveness—or lack thereof—of particular surgical or pharmaceutical interventions or the significance of hypothesized causal associations. For example, a number of individual studies had indicated that maternal obesity was associated with neural tube defects (NTDs) in newborns, but studies ranged from showing no effect to a 3-fold increase in the risk of NTDs with obesity[25]. This lack of concordance is not surprising for a relatively rare birth defect, making it difficult to come to a conclusion from just examining the primary studies individually. A meta-analysis of 20 clinical studies was, however, able to conclusively demonstrate a clear relationship between maternal obesity and an increased risk, with a 1.7-fold rise for moderately obese and over a 3-fold risk for severely obese women[25] . This meta-analysis has important public health implications, particularly given the increasing prevalence of obesity in pregnancy.

An example of the effectiveness of a meta-analysis in resolving an ambiguous set of evidence in the social science literature concerns the value of a family-based intervention approach for serious juvenile offenders called multi-systemic therapy (MST). Despite the logical and theoretical basis for MST and strong advocacy in its favour, a meta-analysis of 35 studies found that there were no significant differences between MST and conventional social services in the success of outcomes (e.g. subsequent arrests or convictions) when the results of the different studies were synthesized, with no support for either clinically significant advantages or harmful outcomes of this approach[26]. Like the clinical meta-analysis of NTD risk, this meta-analysis has public policy ramifications if evidence-based policy is implemented.

The most consequential impact of the introduction of formal research synthesis methodology has been a profound change in the way scientists think about the outcome of scientific research. Instead of looking at ground breaking individual studies as the basis for revealing the conclusive understanding of a scientific problem, individual studies are viewed by meta-analysts as contributions toward the accumulation of evidence[24,27,28]. Clearly there are cases where a single revelatory study completely illuminates and resolves a major problem. However, in many cases of the resolution of scientific problems, when the results of an individual study are combined with those of other studies, the resulting synthesis provides a more general and complete picture of the questions being addressed. Sometimes, the striking results of initial studies are not confirmed by those of subsequent studies or by syntheses of a body of research. This is the basis of recent support for broad implementation of systematic review methodology as, for example, in chemical risk assessment where potentially biased selection of single studies remains the norm[29].

A brief history of meta-analysis

The first formalized attempt to combine information from multiple sources (Fig. 2) was made over 100 years ago by Karl Pearson[30] to ascertain the effectiveness of vaccination in preventing soldiers contracting typhoid. Ronald Fisher, another major figure in the development of modern statistical science, described a method to combine probabilities from different studies[31]. In the late 1930s, William Cochran and his colleague, Frank Yates described approaches that were essentially the same as modern fixed-effect and random-effects models[32] (see Box 1), which were formalized and generalized by Cochran[33]. However, it was not until the insight of psychologists Gene Glass and Mary Smith, that outcome measures from different experiments could be standardized and put on the same scale[34] that meta-analysis began to really impact scientific research. It may be that the time was right, because modern meta-analysis was initiated almost simultaneously in the medical and social sciences in the late 1970s[11]. Methodology was formalized and developed in the following two decades in multiple fields[5-8].

In the early 1990s these methods were taken up by and adapted by ecologists[35], and spread from there to evolutionary biology, conservation, and related disciplines (Table 1).

This resurgence of meta-analysis was followed by rapid methodological and procedural developments where truly cross-disciplinary interactions and fertilization has been notable. The highly interdisciplinary nature of meta-analysis led to the establishment of the *Society for Research Synthesis Methodology* (www.srsm.org) in 2005, whose membership includes statisticians, clinical researchers, epidemiologists, economists, criminologists, social work researchers, educational psychologists, evolutionary biologists and ecologists, among others. Major collaborative networks, the *Cochrane Collaboration* (now known as *Cochrane*; (www.cochrane.org) and *Campbell Collaboration* (www.campbellcollaboration.org) oversee systematic reviews in medical and social sciences, respectively, and bring practitioners and methodologists together and set standards for research synthesis publications and establish evidence-based guidelines for practice and policy.

A case study of a discipline: meta-analysis in ecology, evolution and conservation

Meta-analysis has been implemented to address a wide diversity of problems over a great range of spatial scales and conceptual hierarchical levels in EEC, and these fields provide a case study for exploring its potential for addressing questions with different conceptual foci. Meta-analysis was first adopted by ecologists and evolutionary biologists just over 20 years ago (Table 1), but it has already made a considerable impact on this research field. Meta-analytic approaches in ecology have been introduced at the same time as pressure has increased on ecologists to provide accurate quantitative assessments, predictions and practical solutions to pressing environmental issues such as biodiversity losses, invasive species and biotic responses to climate change, with meta-analysis providing powerful tools for summarizing evidence for these effects. An 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 [36,37], especially through organizations such as the *Centre for Evidence Based Conservation* (www.cebc.bangor.ac.uk) and the *Collaboration for Environmental Evidence* (www.environmentalevidence.org/; Table 1).

Applications of meta-analysis and systematic reviews in EEC have allowed for assessment of impacts of major environmental drivers (e.g., climate change[38]), evaluation of the evidence for ecological and evolutionary theories[39], assessment of effectiveness of conservation and management strategies[36], and highlighting major research gaps[40]. Examples of influential ecological meta-analyses include, for instance, assessment of effects of biodiversity on ecosystem functioning and services[41,42], which have demonstrated that losses of species richness have negative impact on functioning of different ecosystems. From an applied angle, Benayas and colleagues[43] considered the degree to which ecological restoration reverses environmental degradation caused by human activities. Ecological restoration was found to increase provision of biodiversity and ecosystem services in a wide range of ecosystems across the globe, although values of both remained lower in restored versus intact reference ecosystems.

Similarly, evolutionary biologists have tested major hypotheses based on theories of natural selection, sexual selection and social behaviour at unprecedented scales as meta-analysis have allowed them to synthesize large bodies of work in respective topics[39]. Examples of highly cited evolutionary meta-analyses include assessments of correlations between measures of genetic diversity, fitness and population size, that have demonstrated that reduction in population size due to habitat fragmentation reduces genetic variation, and that these losses of genetic diversity have a negative impact on fitness in affected populations[44].

Meta-analysis has also been important in EEC for greatly expanding the capability to evaluate 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, Parmesan and Yohe showed significant range shifts and advancement of spring events in a meta-analysis of climate change impact using data on over 1,700 species from diverse geographic regions[38]. Meta-analysis has also been a valuable tool for practitioners in EEC involved in collaborative research who needed to combine results from experiments carried out across multiple study sites[45,46].

Because meta-analysis was introduced later to EEC than in some other disciplines, ecologists and evolutionary and conservation biologists learned from scientists in other fields and collaborated in developing meta-analysis for use in EEC. While basic methods are the same as in other fields, different data structures, different kind of research questions, and different ways practitioners interpret data necessitated development of different approaches, including implementation of different metrics of effect size, including response ratios[47] and metrics based on measures of variance[48].

Another major difference in the focus of meta-analysis between the medical and social sciences and EEC is the magnitude and nature of both heterogeneity and non-independence of outcomes across as well as within studies. For example, unlike clinical medicine and social sciences where the research is on a single species, the multi-species nature of much of EEC research and therefore of meta-analyses has led biologists to combine phylogenetic comparative methods with meta-analytic models to take into account potential non-independence among lineages due to shared evolutionary history[49-51]. Ecological meta-analyses often use broad study inclusion criteria and address broad questions with a focus on analyses of sources of heterogeneity along with main effects. A limitation of using broad inclusion criteria and a wide diversity of studies is maintaining a clear definition of the problems being addressed in the face of high heterogeneity. A limitation of the reductionist scope and narrow focus exemplified by meta-analysis in clinical medicine is the limited inference possible for factors modifying outcomes, where inclusion of a broader definition of the population of interest might be highly revealing.

Advances in meta-analyses in EEC have been stimulated by software specifically tailored for this field[52,53]. Methodological innovations incorporated or developed in meta-analysis in EEC include the meta-analysis of factorial experiments[54], introduction of randomization (permutation) tests in meta-analysis[55], early embrace of random-effects and mixed models when these were still highly controversial in other disciplines[56], and methods for inclusion of qualitative information such as expert opinions[57].

The introduction and incorporation of meta-analysis in ecological research has raised similar objections to those raised in other disciplines. For instance, as in the criticism of first applications of meta-analysis in social sciences[10], opposition to meta-analysis has been expressed by some ecologists on the basis that ecological studies are too heterogeneous to be meaningfully combined statistically[17] and that ecology is best served by accumulating a catalogue of case studies[58]. Overall, however, introduction of meta-analysis in EEC has been enthusiastically embraced by majority of scientists in these disciplines with the number of meta-analyses published increasing exponentially[59,60]. Practitioners have pointed out that early objection to the introduction of statistics to ecology was analogously based on the uniqueness of individual organisms and micro-site environmental variation. According to Hillebrand and Cardinale, ecological meta-analysis is a remote sensing tool helping scientists to see the forest for the trees, *i.e.* to generalize the findings of individual studies to reach a broader understanding[19].

Limitations, controversies and challenges

Despite both its current utility and future potential, meta-analysis has certain limitations as a tool for research synthesis and for informing decisions. Meta-analysis and systematic reviews can highlight areas where evidence is deficient but cannot overcome these deficiencies. For example, in a systematic review of the literature on hypotheses explaining biological invasions, Lowry and colleagues found a major gap in studies on invasive species in the tropics, highlighting not only what is known but also what is unknown globally about this problem[40]. Other challenges for meta-analysis and systematic reviews include publication bias, i.e. the underreporting of non-significant results or disconfirming evidence[61,62], and research bias[56], where populations, species, or systems are over- or under-represented in the literature, giving a biased view of the totality. These issues may be strongly suspected and their magnitude can sometimes be estimated[63], but cannot truly be corrected[64,65]. Similarly, a synthesis may be constrained by either selective or incomplete reporting [66,67]. The open data movement has considerable scope for remedying these problems, with pre-registration minimising selective reporting and open data minimising publication biases and potentially allowing direct calculation of effect sizes[68].

One unfortunate consequence of the growing recognition and high impact of meta-analyses is an increase in the application of arbitrary and less-well-justified methodology, inaccurately termed 'meta-analysis,' as well as less-than-rigorous applications of these methods. This is not simply a matter of semantics, as the use of statistically flawed approaches can lead to erroneous and misleading results that masquerade as serious research syntheses. The term "meta-analysis" has frequently been misapplied to any study using data from a number of primary publications, regardless of whether appropriate statistical procedures such as effect size calculation and heterogeneity analysis have been employed[69] or appropriate statistical models used that take into account the distinct structure of meta-analytic data. Statistically flawed procedures such as vote-counting[14], a biased statistical technique that provides only limited information about study outcomes, have been long discredited but are still being employed in published papers. Vote-counting is a deceptively convenient procedure in which 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 (and by elaborations on this approach). Although it is vulnerable to misleading and erroneous inferences, it persists zombie-like, returning like the undead to haunt the naïve or determinedly uninformed. For example, a recent paper on species richness in *Science* used a variant of vote counting methods[70]. Vote-counting is not an accurate or unbiased methodology for research synthesis, is not meta-analysis, and is not an acceptable basis for summarizing research results in published papers or reports that are intended to provide meaningful research syntheses.

Another challenge is an unfortunate outcome of the high impact and growing prestige of meta-analyses[71], coupled with use of metrics such as citation numbers and *h*-indices in evaluations of research accomplishments, resulting in unease among some primary researchers about the fairness and rewards of the scientific process[72]. For instance, review articles have been recently decried "the black-market of scientific currency" with calls to replace citations to reviews and meta-analyses by citations of primary studies[73]. Concerns have been raised that "the surge in popularity of meta-analysis may be occurring at the cost of qualitative synthesis" [16]. A recent review of published ecological meta-analyses revealed that the majority of them are produced in a handful of countries around the world, exacerbating academic inequality[74]. This is not dissimilar to recently expressed views in medicine that research synthesists are "research parasites"[75] of primary studies and the researchers who conduct them. However, research synthesis methods are also available to primary researchers, who can carry out meta-analyses in their own areas of expertise, and the two categories of scientists need not be exclusive. The introduction of more explicit guidelines and standards for conducting and reporting meta-analysis could address some of these grievances, and we agree that better methods for citing primary studies in meta-analysis should be implemented to give full credit for the original studies. Perhaps "research parasites" can even serve to increase scientific diversity by the addition of another "trophic level" and can improve the quality of scientific ecosystem functioning and multi-functionality. This controversy serves as an illustration of the problem whereby common criticisms of meta-analysis are propagated across domains but the refutations are not[76]. More nuanced critiques would allow methodological improvement without discounting the advantages and coherence provided by meta-analysis.

Advances, developments and future promises

For hundreds of years, scientists collected data in individual studies, based on observations and experimentation[77]. The implementation of meta-analysis to solve fundamental and applied problems was the first large-scale, coordinated, sustained effort to collect and quantitatively synthesize a body of pre-existing data to determine patterns, to make predictions, to reach generalizations, and to make evidence-based decisions on a range of practices. Currently, there is a great deal of interest in discoveries resulting from the analysis of 'Big Data;' and in parallel, a move to open science practices, transparency, and replication of research to determine the generality and reproducibility of results. We suggest that meta-analysis may be the grandmother of both the new field of Big Data and of the recent emphasis on open science practices. Big Data is a recent term describing large, complex data sets that may be mined for patterns or for making predictions; the term and the issues have been taken up in fields from genomics to climatology to advertising. Data searching, curating, evaluation and quality control are essential aspects of collecting and using Big Data. All of these elements have been the subject of conceptual exploration and formal methodological development in meta-analysis for many years[78]. However, the approach in these two fields is somewhat different. This difference may be in part because meta-analysis is based on the application of statistics and of statisticians working with scientists, while the analysis of Big Data has typically been the province of computer scientists working in coordination with subject-area practitioners.

Open science practices have been strongly advocated particularly in the past few years to facilitate full access to scientific data, particularly in an unbiased manner[79]. Data access and transparency have long been central issues in meta-analysis, because limitations placed on accessing information are serious impediments for best practices in research synthesis, and some of the current limitations of meta-analysis may be overcome by the open data movement. Developing standards of open-science practice have propelled advances in data-sharing, reporting and indexing through initiatives such as *Cochrane*, *Campbell Collaboration*, and *Equator Network* (http://www.equator-network.org). These align with the more generalized open data movement which has considerable scope for remedying publication biases with pre-registration, minimising selective reporting with open data and facilitating direct generation of effects or multi-level analysis based on raw data[68,80,81].

In addition to the benefits accruing from the increased availability of unbiased information, advances in meta-analysis are being propelled by methodological developments, and include the use of machine learning to screen studies[82] for inclusion in systematic reviews and meta-analyses for large reviews, increasingly sophisticated software and models for complex meta-regression[53,83], and integration of meta-analysis with decision support in medicine and other domains[84]. The coupling of Big Data, artificial intelligence and meta-analysis is an important conceptual as well as methodological development reliant on larger trans-disciplinary linkages between statistics, computer science, and biological sciences, social sciences and other scientific fields.

The statistical methodologies underpinning and supporting meta-analysis have been undergoing nearly constant methodological development. Areas of particular current interest include multiple imputations 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 increasingly generalized network meta-analysis and hierarchical modelling of multi-level data, including that from individual participant data in medicine[21,85,86] and in EEC. Network meta-analyses seek to provide comparisons of multiple interventions, including indirect comparisons[87]. These methods are useful where a set of randomized control trials in which pairwise comparisons of interventions have been carried out, and there are common interventions among the studies, but not all studies include all interventions. Developments in and applications of this powerful approach has increased dramatically in clinical medicine over the last 10 years[88].

Perhaps the most important foundation for advances in meta-analysis is support for education in high quality research synthesis methods. Training in meta-analysis should become 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 (e.g. 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 better understanding of meta-analysis to fully exploit the revolution in open data. Most importantly, a new generation of scientists, peer-reviewers, editors, and science-policy practitioners would benefit from increased understanding of evidence synthesis and interpretation.

Meta-analysis can be a key tool in 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. It requires collaboration between statisticians, primary researchers and research synthesists as well as collaboration of meta-analysts across different disciplines. If such collaboration is successful, we are confident that meta-analysis will survive its 'midlife crisis' and will emerge stronger and with a new-found purpose.

Acknowledgements

We dedicate this review to the memory of Ingram Olkin and William Shadish, founding members of the Society of Research Synthesis Methodology, who made tremendous contributions to the development of meta-analysis and research synthesis through their long and distinguished careers and generous help to and collaboration with many other researchers across a wide spectrum of disciplines. We also extend apologies to the many researchers whose relevant work we could not cite here due to space constraints. We also thank Losia Lagisz for help in preparing figures.

We are grateful for 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). JG gratefully acknowledges funding from the U.S. National Science Foundation (ABI 1262402).

References

1 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).

2 Jennions, M. D., Lorite, C. J. & Koricheva, J. in *The handbook of meta-analysis in ecology and evolution* (eds J Koricheva, J Gurevitch, & K Mengersen) Ch. 23, 364-380 (Princeton University Press, 2013).

3 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).

4 Borman, G. D. & Grigg, J. A. in *The handbook of research synthesis and meta-analysis* (eds Harris M. Cooper, Larry V. Hedges, & Jeff C. Valentine) 497-519 (Russell Sage Foundation, 2009).

5 Cooper, H., Hedges, L. V. & Valentine, J. C. *The handbook of research synthesis and meta-analysis* 2nd edn, (Russell Sage Foundation, 2009).

The first edition of this book was a major early compilation that set the standard for best practice in meta-analysis primarily in the social sciences but with applications to medicine and other fields.

6 Hedges, L. & Olkin, I. *Statistical methods for meta-analysis*. (Academic Press, 1985).

7 Rosenthal, R. *Meta-analytic procedures for social research*. Rev. edn, (Sage, 1991).

8 Hunter, J. E., Schmidt, F. L. & Jackson, G. B. *Meta-analysis: Cumulating Research Findings Across Studies*. (Sage, 1982).

9 Glass, G. V. Primary, secondary, and meta-analysis of research. *Edu. Res.* **5**, 3-8, (1976).

10 Eysenck, H. J. Exercise in mega-silliness. *Am. Psychol.* **33**, 517-517, (1978).

This is the first critical response to meta-analysis, raising the issue of heterogeneity and the value (or otherwise) of pooling results from disparate types of studies.

11 Glass, G. V. Meta-analysis at middle age: a personal history. *Res. Synth. Methods* **6**, 221-231, (2015).

12 Lau, J., Rothstein, H. R. & Stewart, G. B. in *The handbook of meta-analysis in ecology and evolution* (eds J Koricheva, J Gurevitch, & K Mengersen) Ch. 25, 407-419 (Princeton University Press, 2013).

13 Ioannidis, J. P. A. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q* **94**, 485-514, (2016).

A critique of poor quality research synthesis by a proponent of evidence-based medicine highlighting the weakness of linkages between exploratory, primary and secondary research and identifying the need for focused high quality meta-analysis

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

15 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).

16 Morrissey, M. B. Meta-analysis of magnitudes, differences and variation in evolutionary parameters. *J. Evol. Biol.* **29**, 1882-1904, (2016).

17 Whittaker, R. J. Meta-analyses and mega-mistakes: calling time on meta-analysis of the species richness-productivity relationship. *Ecology* **91**, 2522-2533, (2010).

18 Begley, C. G. & Ellis, L. M. Drug development: Raise standards for preclinical cancer research. *Nature* **483**, 531-533, (2012).

19 Hillebrand, H. & Cardinale, B. J. A critique for meta-analyses and the productivity-diversity relationship. *Ecology* **91**, 2545-2549, (2010).

20 Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. & Grp, P. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PloS Med.* **6**, e1000097, (2009).
A consensus regarding the reporting requirements for medical meta-analysis. This has been highly influential in ensuring good reporting practice and standardising language in evidence-based medicine. Further PRISMA guidance has been developed for protocols, individual patient data meta-analyses and animal studies.
21 Higgins, J. P. T. & Green, S. *Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0. The Cochrane Collaboration*. (2013).
This large collaborative work provides definitive guidance for the production of systematic reviews in medicine. It is of broader interest beyond medicine as it provides linkages to methods work underpinning the guidance that is less well developed outside of medical fields.
22 Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R. & Cheraghi-Sohi, S. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Serv. Res.* **14**, 579, (2014).
23 Lortie, C. J., Stewart, G., Rothstein, H. & Lau, J. How to critically read ecological meta-analyses. *Research Synthesis Methods* **6**, 124-133, (2015).
24 Murad, M. H. & Montori, V. M. Synthesizing evidence: shifting the focus from individual studies to the body of evidence. *JAMA* **309**, 2217-2218, (2013).
25 Rasmussen, S. A., Chu, S. Y., Kim, S. Y., Schmid, C. H. & Lau, J. Maternal obesity and risk of neural tube defects: a metaanalysis. *Am. J. Obstet. Gynecol.* **198**, 611-619, (2008).
26 Littell, J. H., Popa, M. & Forsythe, B. Multisystemic Therapy for social, emotional, and behavioral problems in youth aged 10-17. *The Cochrane database of systematic reviews*, CD004797, (2005).
27 Schmidt, F. L. What do data really mean?: research findings, meta-analysis, and cumulative knowledge in psychology. *Am Psychol* **47**, 1173-1181, (1992).
28 Ioannidis, J. P. Meta-research: the art of getting it wrong. *Res. Synth. Methods* **3**, 169-184, (2010).
29 Whaley, P. *et al.* Implementing systematic review techniques in chemical risk assessment: Challenges, opportunities and recommendations. *Environment International* **92–93**, 556-564, (2016).
30 Pearson, K. Report on certain enteric fever inoculation statistics. *Brit. Med. J.* **1904**, 1243-1246, (1904).
The earliest published application combining results from multiple studies, setting the stage for the development of modern meta-analytical methods
31 Fisher, R. A. *Statistical Methods for Research Workers*. (Oliver and Boyd, 1925).
32 Yates, F. & Cochran, W. G. The analysis of groups of experiments. *J. Agr. Sci.* **28**, 556-580, (1938).
33 Cochran, W. G. The Combination of Estimates from Different Experiments. *Biometrics* **10**, 101-129, (1954).
34 Smith, M. L. & Glass, G. V. Meta-analysis of psychotherapy outcome studies. *Am. Psychol.* **32**, 752-760, (1977).
35 Gurevitch, J., Morrow, L. L., Wallace, A. & Walsh, J. S. A meta-analysis of competition in field experiments. *Am. Nat.* **140**, 539-572, (1992).
An influential early ecological meta-analysis combining multiple experimental outcomes on a longstanding and controversial topic that introduced a wide range of ecologists to research synthesis methods.
36 Stewart, G. Meta-analysis in applied ecology. *Biol Letters* **6**, 78-81, (2010).
37 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).

38 Parmesan, C. & Yohe, G. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421**, 37-42, (2003).
39 Jennions, M. D., Lorite, C. J. & Koricheva, J. in *The handbook of meta-analysis in ecology and evolution* (eds J Koricheva, J Gurevitch, & K Mengersen) Ch. 24, 381-403 (Princeton University Press, 2013).
40 Lowry, E. *et al.* Biological invasions: a field synopsis, systematic review, and database of the literature. *Ecol. Evol.* **3**, 182-196, (2012).
41 Balvanera, P. *et al.* Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecol. Lett.* **9**, 1146-1156, (2006).
42 Cardinale, B. J. *et al.* Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* **443**, 989-992, (2006).
43 Benayas, J. M. R., 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).
44 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).
45 Gurevitch, J. in *The handbook of meta-analysis in ecology and evolution* (eds J Koricheva, J Gurevitch, & K Mengersen) Ch. 19, 313-320 (Princeton University Press, 2013).
46 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).
47 Hedges, L. V., Gurevitch, J. & Curtis, P. S. The Meta-Analysis of Response Ratios in Experimental Ecology. *Ecology* **80**, 1150-1156, (1999).
48 Nakagawa, S. *et al.* Meta-analysis of variation: ecological and evolutionary applications and beyond. *Methods Ecol. Evol.* **6**, 143-152, (2015).
49 Adams, D. C. Phylogenetic meta-analysis. *Evolution* **62**, 567-572, (2008).
50 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).
51 Lajeunesse, M. J. Meta-Analysis and the comparative phylogenetic method. *Am. Nat.* **174**, 369-381, (2009).
52 Rosenberg, M. S., Adams, D. C. & Gurevitch, J. *MetaWin: Statistical Software for Meta-Analysis with Resampling Tests. Version 1* (Sinauer Associates, 1997).
53 Wallace, B. C. *et al.* OpenMEE: Intuitive, open-source software for meta-analysis in ecology and evolutionary biology. *Methods Ecol. Evol.*, DOI: 10.1111/2041-210X.12708 (2016).
54 Gurevitch, J., Morrison, J. A., Hedges, L. V. & Associate Editor: Peter, J. M. The Interaction between competition and predation: a meta-analysis of field experiments. *Am. Nat.* **155**, 435-453, (2000).
55 Adams, D. C., Gurevitch, J. & Rosenberg, M. S. Resampling tests for meta-analysis of ecological data. *Ecology* **78**, 1277-1283, (1997).
56 Gurevitch, J. & Hedges, L. V. Statistical issues in ecological meta-analyses. *Ecology* **80**, 1142-1149, (1999).
57 Schmid, C. H. & Mengersen, K. in *The handbook of meta-analysis in ecology and evolution* (eds J Koricheva, J Gurevitch, & K Mengersen) Ch. 11, 145-173 (Princeton University Press, 2013).
An important gentle introduction to Bayesian methods for meta-analysis. These approaches are important in analyses where multiple related but different studies are combined as in ecology and evolution. This chapter explains how to undertake analyses that express uncertainty realistically and allow what is known about one thing to inform another unknown thing where they are related.

58 Simberloff, D. Rejoinder to Simberloff (2006): Don't calculate effect sizes; study ecological effects. *Ecol. Lett.* **9**, 921-922, (2006).
59 Cadotte, M. W., Mehrkens, L. R. & Menge, D. N. L. Gauging the impact of meta-analysis on ecology. *Evol. Ecol.* **26**, 1153-1167, (2012).
60 Gurevitch, J., Curtis, P. S. & Jones, M. H. Meta-analysis in ecology. *Adv. Ecol. Res.* **32**, 199-247, (2001).
61 Jennions, M. D., Lorite, C. J., Rosenberg, M. S. & Rothstein, H. R. in *The handbook of meta-analysis in ecology and evolution* (eds J Koricheva, J Gurevitch, & K Mengersen) Ch. 14, 207-236 (Princeton University Press, 2013).
62 Leimu, R. & Koricheva, J. Cumulative meta-analysis: a new tool for detection of temporal trends and publication bias in ecology. *P Roy Soc Lond B Bio* **271**, 1961-1966, (2004).
63 Egger, M., Smith, G., Schneider, M. & Minder, C. Bias in metaanalysis detected by a simple, graphical test. *Br Med J* **315**, 629 - 634, (1997).
64 Koricheva, J., Jennions, M. D. & Lau, J. (eds J Koricheva, J Gurevitch, & K Mengersen) Ch. 15, 237-254 (Princeton University Press, 2013).
65 Lau, J., Ioannidis, J. P. A., Terrin, N., Schmid, C. H. & Olkin, I. Evidence based medicine - The case of the misleading funnel plot. *Br Med J* **333**, 597-600, (2006).
66 Chan, A. W., Hrobjartsson, A., Haahr, M. T., Gotzsche, P. C. & Altman, D. G. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* **291**, 2457-2465, (2004).
67 Lajeunesse, M. J. in *The handbook of meta-analysis in ecology and evolution* (eds J Koricheva, J Gurevitch, & K Mengersen) Ch. 13, 195-206 (Princeton University Press, 2013).
68 Parker, T. H. *et al.* Transparency in ecology and evolution: real problems, real solutions. *Trends Ecol. Evol.* **31**, 711-719, (2016).
69 Vetter, D., Rucker, G. & Storch, I. Meta-analysis: A need for well-defined usage in ecology and conservation biology. *Ecosphere* **4**, (2013).
70 Adler, P. B. *et al.* Productivity is a poor predictor of plant species richness. *Science* **333**, 1750-1753, (2011).
71 Patsopoulos, N. A., Analatos, A. A. & Ioannidis, J. P. A. Relative citation impact of various study designs in the health sciences. *JAMA* **293**, 2362-2366, (2005).
72 Kueffer, C. *et al.* Fame, glory and neglect in meta-analyses. *Trends Ecol. Evol.* **26**, 493-494, (2011).
73 Cohnstaedt, L. W. & Poland, J. Review Articles: The Black-Market of Scientific Currency. *Ann. Entomol. Soc. Am.*, (2016).
74 Mori, A. S., Qian, S. & Tatsumi, S. Academic inequality through the lens of community ecology: a meta-analysis. *PeerJ* **3**, e1457, (2015).
75 Longo, D. L. & Drazen, J. M. Data Sharing. *N. Engl. J. Med.* **374**, 276-277, (2016).
76 Brannan, M. J., Fleetwood, S., O'Mahoney, J. & Vincent, S. Critical Essay: Meta-analysis: A critical realist critique and alternative. *Hum. Relat.* **70**, 11-39, (2016).
77 Gauch, H. G. *Scientific Method in Practice*. (Cambridge University Press, 2003).
78 Staff, S. Challenges and Opportunities. *Science* **331**, 692-692, (2011).
79 Nosek, B. A. *et al.* Promoting an open research culture. *Science* **348**, 1422-1425, (2015).
80 Riley, R. D., Lambert, P. C. & Abo-Zaid, G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *Br. Med. J.* **340**, (2010).
81 Stewart, L. A., Clarke, M., Rovers, M. & et al. Preferred reporting items for a systematic review and meta-analysis of individual participant data: The prisma-ipd statement. *JAMA* **313**, 1657-1665, (2015).

82 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).
83 Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Soft.* **36**, 1-48, (2010).
84 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).
85 Nakagawa, S. & Santos, E. S. A. Methodological issues and advances in biological meta-analysis. *Evol. Ecol.* **26**, 1253-1274, (2012).
86 Koricheva, J., Gurevitch, J. & Mengersen, K. *The handbook of meta-analysis in ecology and evolution*. (Princeton University Press, 2013).
The first comprehensive guidance on undertaking meta-analysis in the fields of ecology and evolution. The guidance is likely to be relevant to the broader environmental field and domains where heterogeneity is expected. In contrast to other books on research synthesis there is explicit consideration of the different approaches adopted in different domains.
87 Lumley, T. Network meta-analysis for indirect treatment comparisons. *Stat. Med.* **21**, 2313-2324, (2002).
88 Zarin, W. *et al.* Characteristics and knowledge synthesis approach for 456 network meta-analyses: a scoping review. *BMC Med.* **15**, (2017).
89 Higgins, J. & Thompson, S. Quantifying heterogeneity in a meta-analysis. *Stat. Med.* **21**, 1539 - 1558, (2002).
90 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**, 62-66, (1991).
91 Arnqvist, G. & Wooster, D. Meta-analysis: synthesizing research findings in ecology and evolution. *Trends Ecol. Evol.* **10**, 236-240, (1995).
92 Lajeunesse, M. J. phyloMeta: a program for phylogenetic comparative analyses with meta-analysis. *Bioinformatics* **27**, 2603-2604, (2011).
93 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).
94 Sutton, A. J. in *The handbook of research synthesis and meta-analysis* (eds H Cooper, L V Hedges, & J C Valentine) 435-452 (Russell Sage Foundation, 2009).
95 Duval, S. & Tweedie, R. A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *J. Am. Stat. Assoc.* **95**, 89-98, (2000).
96 Trikalinos, T. A. & Ioannidis, J. P. in *Publication bias in meta-analysis : prevention, assessment and adjustments* (eds H Rothstein, A J Sutton, & M Borenstein) 241–259 (John Wiley, 2005).
97 O'Rourke, K. An historical perspective on meta-analysis: dealing quantitatively with varying study results. *J. R. Soc. Med.* **100**, 579-582, (2007).
98 Shadish, W. R. & Lecy, J. D. The meta-analytic big bang. *Res. Synth. Methods* **6**, 246-264, (2015).
99 DerSimonian, R. & Laird, N. Meta-analysis in clinical trials. *Control. Clin. Trials* **7**, 177-188, (1986).
100 Moher, D. *et al.* Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet* **354**, 1896-1900, (1999).

Box 1: Meta-analytic models and heterogeneity

Two basic models of meta-analysis are the fixed-effect model (Fig 1a) and the random-effects model (Fig 1b)[6]. The fixed-effect model assumes all effect sizes have a common (true) overall effect and thus it is also referred to as the common-effect model. The model can be written as:

$$y_i = b_0 + \varepsilon_i,$$

where y_i is the effect size from the ith study (black squares), b_0 is the overall true mean effect (Fig. 1a, brown diamond and dotted brown line) and ε_i is the deviation from the overall mean for the *i*th study (solid black lines). The study deviation ε_i is distributed according to sampling (error) variance, σ_i^2 (the variance of beige distributions). The inverse of this sampling variance is generally used as the study weight in a meta-analysis; note that the inverse variance can also be called 'precision' but in the meta-analytic literature, precision is often reserved for the square root of the inverse variance. The assumption of a shared true effect in the fixed-effect model is often too restrictive, because in many research areas, effect sizes from different studies would not be believed to share a common true effect size.

The random-effects model, on the other hand, assumes that the true effect sizes from different studies differ from one another by random variation, written as:

$$y_i = b_0 + \alpha_i + \varepsilon_i,$$

where b_0 is the overall mean across studies (Fig. 1b, purple diamond and dotted purple line), α_i is the deviation of the true effect of study *i* from the overall mean across studies, $b_0 + \alpha_i$ is the true mean for the *i*th study (dotted orange lines) and ε_i is the within-study sampling error (solid black lines), the deviation from the study-specific mean. The study-specific effect, or the deviation from the overall mean, (α_i; solid orange lines) has the variance τ^2 (i.e., the between-study variance; blue distributions), and like the fixed-effect mode, the sampling error (ε_i) has a variance of σ_i^2. The typical (average) value of the sampling variances from all of the studies is the within-study variance, $\bar{\sigma}^2$ [89]. In meta-analysis, heterogeneity is the amount of the variation across studies after removing within-study sampling errors. Therefore, the absolute value of heterogeneity is τ^2 while the relative value of heterogeneity, known as I^2, is the ratio between τ^2 and the sum of the τ^2 and $\bar{\sigma}^2$.

Meta-regression models have become a major way to quantify observed heterogeneity across studies (older studies used the heterogeneity statistic, *Q*, to quantify heterogeneity[6]). Meta-regression models include covariates (also known as moderators or explanatory variables), to explain portions of the total heterogeneity. Meta-regressions can be modelled as fixed-effect (Fig Ic) or random-effects models (Fig Id). The simplest form of the fixed-effect meta-regression with one (either continuous or 2-level categorical) moderator can be written as:

$$y_i = b_0 + b_1 x_i + \varepsilon_i,$$

where $b_0$is the intercept and b_1 is the slope (regression coefficient; dotted brown line) for the moderator x, x_i is the moderator value for *i*th study, and ε_i is the same within-study sampling error (solid black lines) as in the simple fixed-effect model. The simplest form of the random-effects meta-regression can be written as:

$$y_i = b_0 + \alpha_i + b_1 x_i + \varepsilon_i,$$

where b_1 is the common slope (dotted purple and dotted orange lines), b_0 is the overall intercept (where the dotted purple line intersect the zero value of x), $b_0 + \alpha_i$ are the study specific intercepts for each of the *i* studies (the interaction between the dotted orange lines and $x = 0$), and α_i (solid orange lines) and ε_i (solid black lines) are the same as in the random-effects model. Moderators can be continuous or categorical variables (when categorical, they are treated using dummy variables.) The assumption of the fixed-effect meta-regression is that a moderator(s) can explain all between-study heterogeneity whereas that of the random-effects meta-regression is that a part of heterogeneity is due to a moderator(s), but that there is also (unexplained) heterogeneity between studies in the true effect. Examples of moderators are differences in dosage of a drug treatment across studies, differences in study duration, or level of CO_2 used in studying responses of plants to elevated CO_2. More complex meta-regression models can include two or more moderators. Further, meta-analytic models can include more random factors (hierarchical levels) to accommodate more complexity of meta-analytic data typical to ecology and evolution [85].

Figure I. Visualization of two major meta-analytic and meta-regression models. a. a fixed-effect meta-analysis, **b.** a random-effects meta-analysis, **c.** a fixed-effect meta-regression and **d.** a random-effect meta-regression. The black squares are observations (effect sizes) with square size representing sample size (panels a-d), the dotted brown line is the overall mean or the slope from fixed-effect models with the brown diamond showing the 95% confidence interval of the mean (panels a,c), the dotted purple line is the overall mean or the (common) slope from random-effects models with the brown diamond showing the 95% confidence interval (panels b,d), the solid black lines are deviations from the overall mean (panels a,b) or slope (panels c,d), the dotted orange lines are study-specific means (panel b) or slopes with study specific intercepts (panel d), the solid orange lines are deviations from the study-specific means (panel b) or slopes (panel d) and the beige distributions represent the sampling error variance (within-study variance; panels a,c) while the blue distributions the between-study variance (panels a-d).

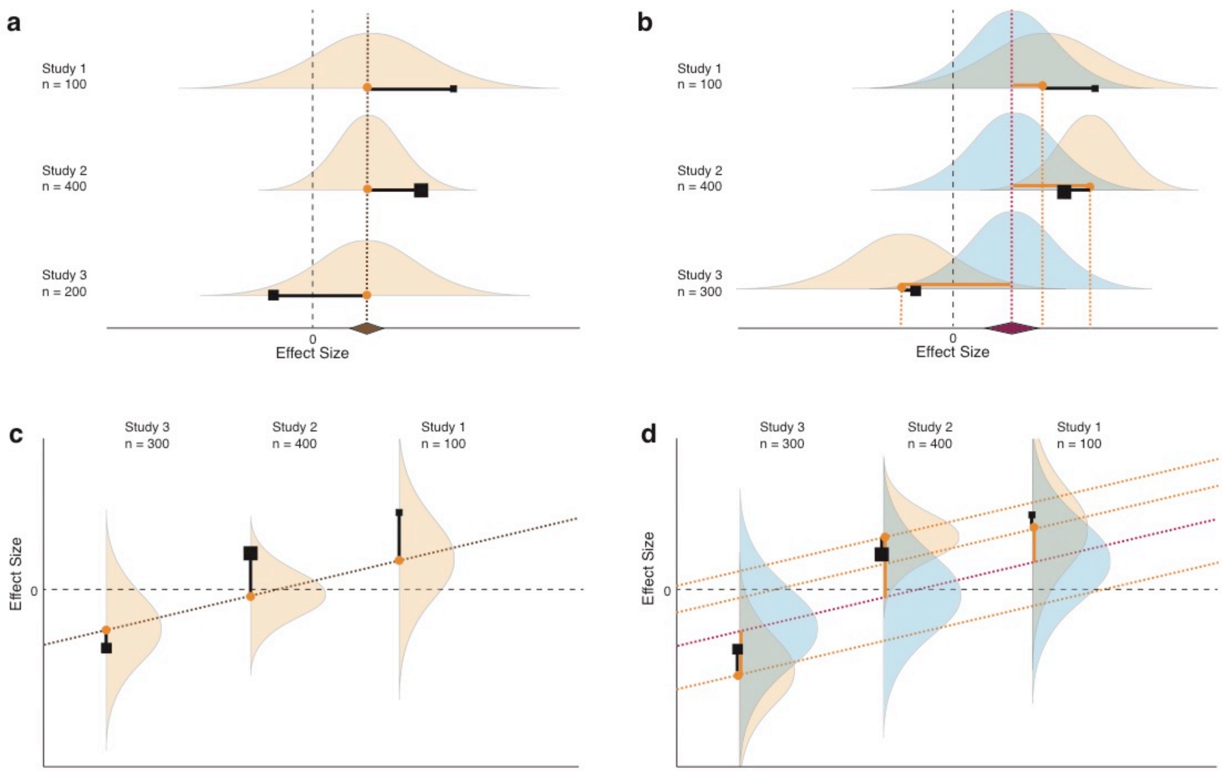
a
Study 1
n = 100
Study 2
n = 400
Study 3
n = 200
0
Effect Size
b
Study 1
n = 100
Study 2
n = 400
Study 3
n = 300
0
Effect Size
c
Study 3
n = 300
Study 2
n = 400
Study 1
n = 100
Effect Size
0
d
Study 3
n = 300
Study 2
n = 400
Study 1
n = 100
Effect Size
0

Table 1. Milestones of systematic review and meta-analytic development in ecology, evolution and conservation.

Year	Milestone
1991	First meta-analysis in ecology is published[90]
1995	Seminal paper by Arnqvist and Wooster is published in *Trends in Ecology and Evolution* introducing meta-analysis to many readers in ecology[91]
1995	National Center for Ecological Analysis and Synthesis is established in USA
1997	MetaWin, 1st software for ecological meta-analysis is created[52]
1999	Special feature on meta-analysis is published in the journal *Ecology*, including a paper by Hedges and colleagues[47] introducing log response ratio as a new ecological metric of effect size, and an influential paper on statistical issues in ecological meta-analysis[56]
2001	First review of meta-analysis in ecology is published[60]
2003	Centre for Evidence-Based Conservation (CEBC) is established in UK
2007	Collaboration for Environmental Evidence is created
2008/9	Seminal papers on phylogenetic meta-analysis are published [49,51] and phylometa software for integrating phylogeny into meta-analysis is created[92]
2011	Environmental Evidence (the official journal of the Collaboration for Environmental Evidence) is established
2013	First Handbook of meta-analysis in ecology and evolution is published [86]
2014	OpenMEE, software for ecological and evolutionary meta-analysis, is released[53]
2016	1st International Conference of the Collaboration for Environmental Evidence takes place in Stockholm

Figure 1. A variety of charts and plots common in meta-analysis. a. PRISMA diagram[20], **b.** a forest plot showing overall means based on a fixed-effect model and random-effects model, common in medical meta-analyses (see Box 1), **c.** box and whisker plots of mean effect sizes, common in ecological and evolutionary meta-analysis, **d.** a bubble plot to show a predicted line from a meta-regression analysis where the size of the bubble reflect study sample size[93], **e.** a funnel plot of original data (red points) showing some funnel asymmetry, which may indicate publication bias [63,94], with augmented data (white points) from the trim-and-fill method, which is a sensitivity analysis correcting for a potential publication bias[95] and, **f.** a forest plot of a cumulative meta-analysis where effect sizes are meta-analyzed cumulatively in a chronological order, thus, showing a temporal trend and convergence of effect sizes across studies [96].

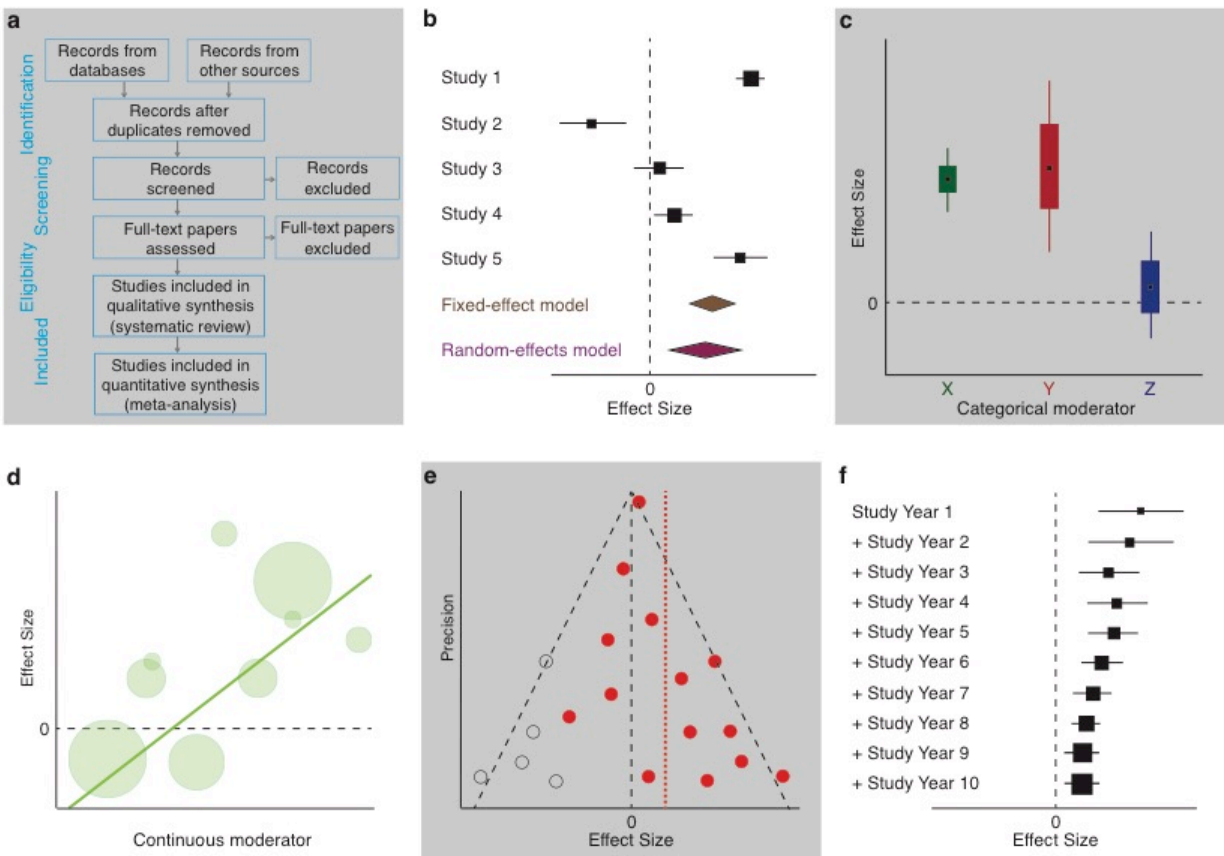

Figure 2. Milestones in meta-analytic history. Red line shows the number of papers from a Scopus search. These historical milestone publications are chosen based two main criteria, precedence and influence (we replied heavily on these references[97,98]).