Introduction to meta-analysis

Outline for today

- Meta analysis compared with traditional review article
- Quantitative summaries vs. vote-counting
- How to carry out a meta-analysis
- Effect size
- Fixed and mixed-effects
- Associating effect sizes with relevant variables
- Publication bias
- Make your results accessible to meta-analysis

Scientific studies on a topic are often repeated

New studies improve/expand on previous studies, or examine the same issue in a different study system, or using different methods

- Schoener et al. (1983) found 164 published field experiments on interspecific competition.
- Gardner et al. (2003) obtained results from 51 separate studies reporting coral cover from 294 sites from across the Caribbean.
- Bell et al. (2009) found 759 published estimates of the repeatability of behavior, from 114 studies of 98 species.
- Vilà et al (2011) reviewed 199 articles reporting 1041 field studies describing the ecological impacts of 135 alien plant taxa.

A method is needed to summarize results from multiple studies

- Dr. Benjamin Spock sold 50 million copies of *Baby and Child Care* 1950s—1990s. In it he wrote "*I think it is preferable to accustom a baby to sleeping on his stomach from the beginning if he is willing*". Other pediatricians made similar recommendations.
- From the 1950s into the 1990s, more than 100,000 babies died of sudden infant death syndrome (SIDS).
- In the early 1990s, researchers realized that the risk of SIDS decreased by at least 50% when babies were put to sleep on their backs rather than face down.
- Subsequent education campaigns led to a dramatic drop in the number of SIDS deaths.
- However, research was available from 1970 that sleeping on the stomach was hazardous to babies. An earlier synthesis of the data could have got the answer much sooner.

The review article was the traditional approach

An expert in the field assembles the studies published on a topic, thinks about them carefully and (hopefully) fairly, and then writes a review article summarizing the overall conclusions reached.

A first-rate review article advances a field far beyond a mere summary.

It reviews and comments on the current state of thought and knowledge about a particular topic.

Such a review will propose new hypotheses, uncover previously unnoticed relationships, and point to new paths of research.

The traditional review lacks a quantitative method

This might lead to two problems

- *Bias*.
  In his 1986 book *How to Live Longer and Feel Better*, Linus Pauling (the only person to be awarded two unshared Nobel Prizes) cited 30 studies supporting his idea that large daily doses of vitamin C reduces the risk of contracting the common cold, but cited no studies opposing the idea, even though a number had been published. Not all reviews are **so** biased, but there are few rules regarding selection of studies for review.
- *Lack of a quantitative summary of research findings*.
  Reviews don't tell us about how large the effect is.

Vote-counting was a step in the right direction

Divide studies into two categories: those that yielded a statistically significant result supporting the research hypothesis, and those that did not. The proportions of studies 'voting' for or against the hypothesis are then counted.

Vol. 122, No. 2 The American Naturalist August 1983

FIELD EXPERIMENTS ON INTERSPECIFIC COMPETITION

THOMAS W. SCHOENER

Department of Zoology, University of California, Davis, California 95616

EXISTENCE OF COMPETITION

An overwhelming fraction of experimental attempts to detect interspecific competition in the field did so: 148 of 164 studies, or 90%, demonstrate some competition. One-hundred ten of the 148 studies record changes in numbers through local births and deaths or migration.

Limitations of vote-counting

- By counting only the statistically significant studies vote-counting ignores all the quantitative information about the magnitudes of effects.
- Too conservative. “Votes” are affected by the power of individual studies, which may be weak.
- Significance level by itself doesn’t indicate whether two or more studies obtained the same outcome.
- The magnitude of the effect is downplayed.
- It is difficult to quantify the effects of publication bias.
- Method is unable to weigh the effect of studies differing in sample size, and therefore power.

Limitations of vote-counting

- The Antiplatelet Trialists' Collaboration (1994) conducted a meta-analysis of 142 randomized experiments testing whether taking aspirin or other antiplatelet medication following a stroke or myocardial infarction ("heart attack") reduced the risk of future stroke. Total $N > 70{,}000$.
- The vote: 19 of 142 studies showed a statistically significantly better result for patients on antiplatelet therapy than for the control patients. Two of the 142 studies showed a significantly worse rate of vascular events with aspirin treatment.
- Yet 14.7% (5400/36,711) of patients in the control groups had subsequent vascular events, compared with 11.4% (4183/36,536) in the treated group. Small effect but real, according to meta-analysis methods. This conclusion saved many lives.

Meta-analysis, the “analysis of analyses”

Meta-analysis refers to the statistical synthesis of results from a series of studies (Borenstein et al 2009).

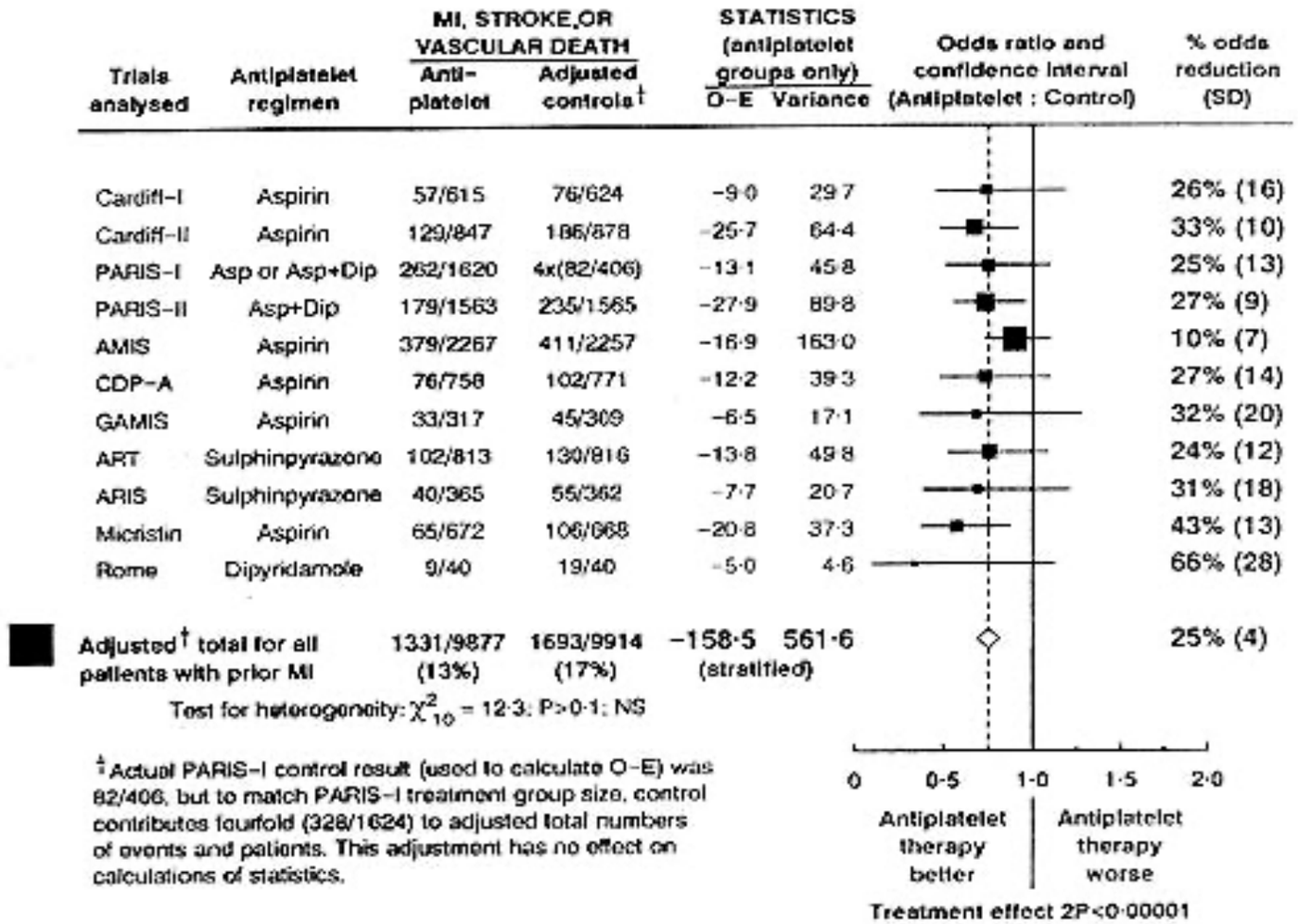
The method involves compiling all known scientific studies estimating an effect (“systematic review”) and quantitatively combining them to give an overall estimate of the effect.

Meta-analysis allows us to generalize. It lets us determine how frequent, how important, and how consistent effects are across a variety of systems.

Meta-analysis gets past the occasional sensational result (the one you read about in the newspaper) to an objective assessment of all the evidence.

Meta-analysis, the “analysis of analyses”

Came from medical research, in which all studies are all of the same species (humans). Here is a “forest plot” from the Antiplatelet Trialists’ Collaboration.



Trials analysed	Antiplatelet regimen	MI, STROKE, OR VASCULAR DEATH: Anti-platelet	MI, STROKE, OR VASCULAR DEATH: Adjusted controls†	STATISTICS (antiplatelet groups only): O–E	STATISTICS (antiplatelet groups only): Variance	Odds ratio and confidence interval (Antiplatelet : Control)	% odds reduction (SD)
Cardiff–I	Aspirin	57/615	76/624	−9·0	29·7		26% (16)
Cardiff–II	Aspirin	129/847	186/878	−25·7	64·4		33% (10)
PARIS–I	Asp or Asp+Dip	262/1620	4x(82/406)	−13·1	45·8		25% (13)
PARIS–II	Asp+Dip	179/1563	235/1565	−27·9	89·8		27% (9)
AMIS	Aspirin	379/2267	411/2257	−16·9	163·0		10% (7)
CDP–A	Aspirin	76/758	102/771	−12·2	39·3		27% (14)
GAMIS	Aspirin	33/317	45/309	−6·5	17·1		32% (20)
ART	Sulphinpyrazone	102/813	130/816	−13·8	49·8		24% (12)
ARIS	Sulphinpyrazone	40/365	55/362	−7·7	20·7		31% (18)
Micristin	Aspirin	65/672	106/668	−20·8	37·3		43% (13)
Rome	Dipyridamole	9/40	19/40	−5·0	4·6		66% (28)
Adjusted† total for all patients with prior MI		1331/9877 (13%)	1693/9914 (17%)	−158·5 (stratified)	561·6		25% (4)

Test for heterogeneity: $\chi^2_{10} = 12{\cdot}3$; P>0·1; NS

†Actual PARIS–I control result (used to calculate O–E) was 82/406, but to match PARIS–I treatment group size, control contributes fourfold (328/1624) to adjusted total numbers of events and patients. This adjustment has no effect on calculations of statistics.

Meta-analysis, the “analysis of analyses”

Ecologists and evolutionary biologists attempt to generalize across a much wider range of species and systems.

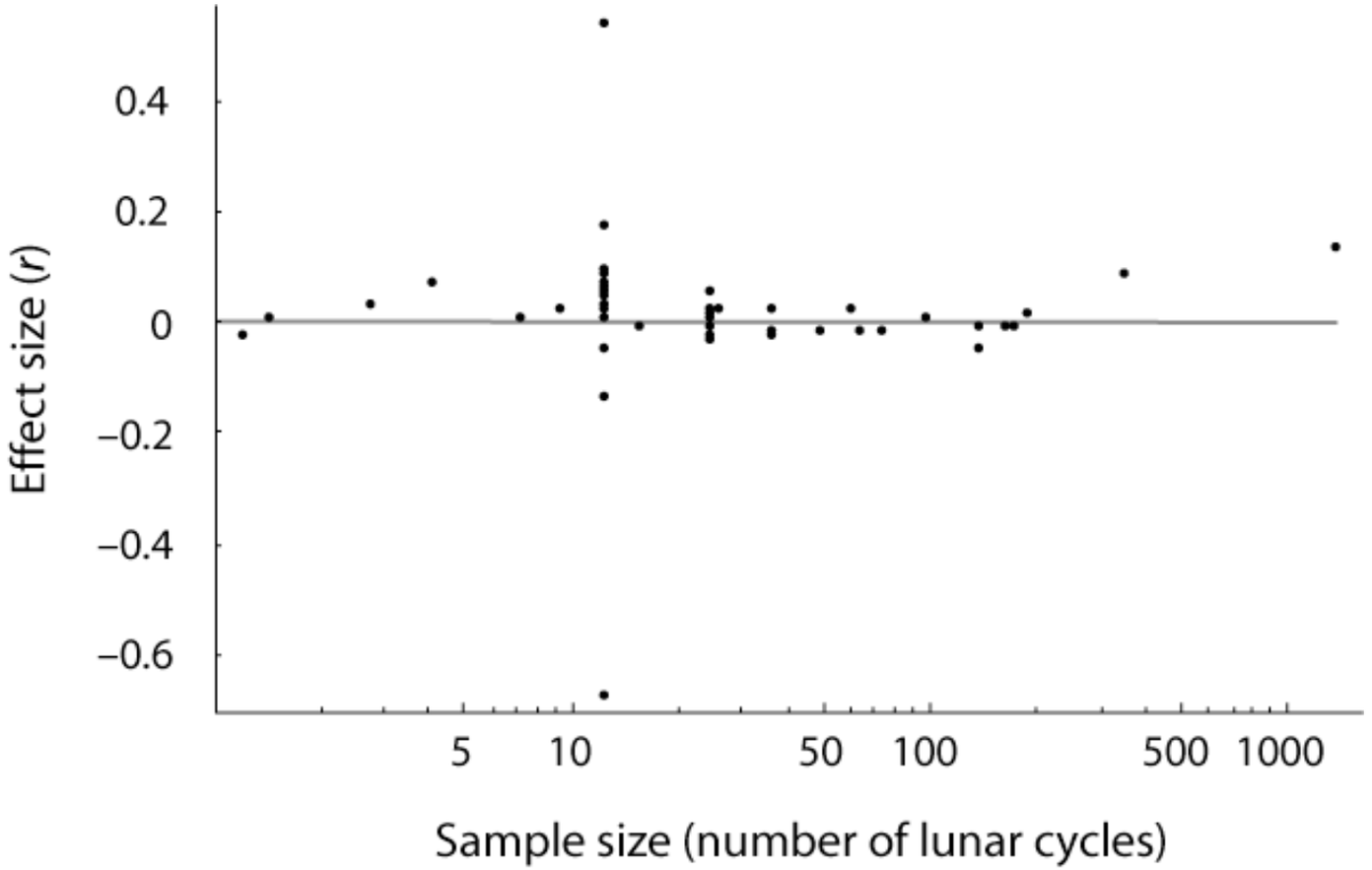
This is more challenging than studies carried out on a single species (e.g., humans).

Example 1: Meta-analysis of the Transylvania effect

- Many people believe that a full moon can affect human behavior. The word lunacy is derived from the Latin luna, moon.
- Legends of strange happenings, such as werewolves and vampires, have been connected to full moons for centuries.
- Lord Blackstone, an 18th-century English jurist, was the first to define a condition of madness exacerbated by the lunar cycle: "*A lunatic, or non compos mentis, is properly one who hath lucid intervals, sometimes enjoying his senses and sometimes not and that frequently depending upon the changes of the moon*."
- Rotton and Kelly (1985) showed that 50% of university students believed that people act strangely during a full moon.
- Vance (1995) reported that as many as 81% of mental health professionals believed that the full moon alters individual behaviour.

Example 1: Meta-analysis of the Transylvania effect

Rotton and Kelly (1985) carried out a meta-analysis of studies correlating homicide rates, psychiatric hospital admissions, suicide rates, crisis calls, etc. The average effect size *r* was smaller than 0.01.



Steps of a meta-analysis

1. Define the question and scope.

- A narrow question applied to a homogeneous group?
  “Does aspirin reduce incidence of myocardial infarction?”
- Or a heterogeneous set of studies or variables?
  “How much genetic variation exists in populations for behavioral traits?”
- Only experiments with controls and randomization? Only replicated experiments? Only experiments with blinding?
- It may be best to adopt a reasonably wide scope and investigate later whether differences between methods lead to different effects overall.

Steps of a meta-analysis

2. Literature search, systematic review, gather data.

- Make it exhaustive to avoid bias.
- Easily-found studies are *different* from those that we cannot find easily. Studies finding large, statistically significant effects are more likely to be published, more likely to be in “first-rate” journals, and more likely to be referenced in other articles.
- Statistical techniques exist to account partially for publication bias (funnel plots) but they do not replace an exhaustive survey.
- Decide whether to (hold your nose and) include studies of apparently poor quality. Failure to have well-defined criteria can lead to bias (we are more likely to discard a poor study if it disagrees with our pet hypothesis).

Steps of a meta-analysis

- Ideally, the data obtained should all be independent, but non-independence of various sorts creeps in (e.g. multiple studies by the same lab).
- A single study may provide measurements on multiple species, or measurements of multiple responses on the same species. Include them all or take a summary measure?
- One or a small number of species (e.g., great tit) or systems (e.g., intertidal zone) may be overrepresented in the literature. Treat them all as independent?
- It may be worse to leave data out, or take summary measures, than to throw every data point into the analysis.

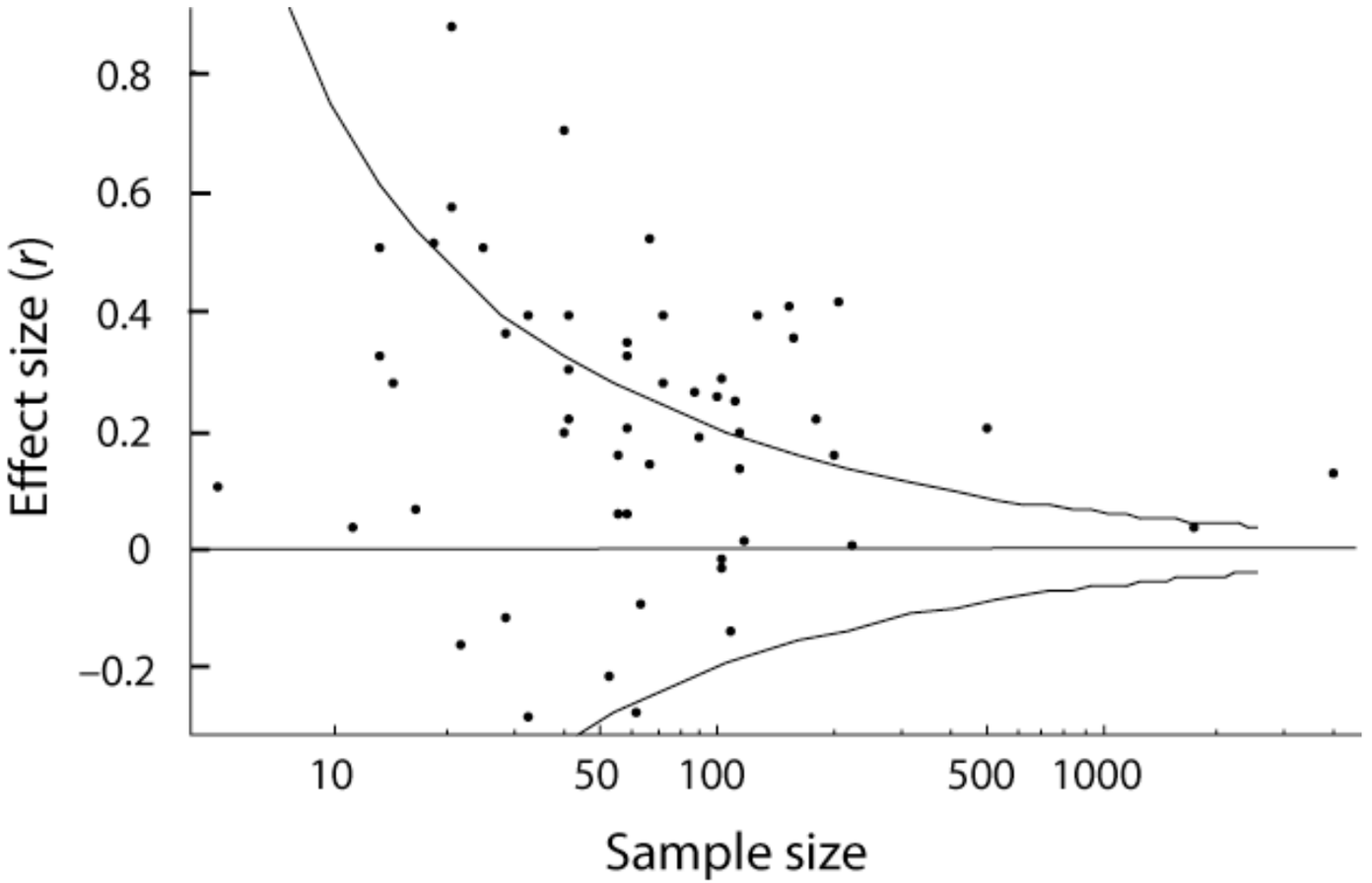
Example 2: Testosterone vs aggression

Book et al. (2001) asked “Are testosterone levels and aggression correlated in human males?” It included a huge diversity of types of studies:

- levels of testosterone in prisoners convicted of violent crimes compared to those of prisoners convicted of property crimes.
- levels of testosterone in university students compared with their answers to questionnaires that asked them for levels of agreement to statements like “If somebody hits me, I hit back.”
- levels of aggression in !Kung San males as determined by counting “their scars and sometimes still open wounds in the head region.”
- drunken Finnish spouse-abusers compared to drunken Finns drinking quietly in a bar.
- members of “rambunctious” fraternities compared to “responsible” fraternities.

Example 2: Testosterone vs aggression

Below is the "funnel plot" of studies comparing human aggression to levels of testosterone. The curves show the approximate boundaries of the critical regions that would reject the null hypothesis in any one study with $\alpha = 0.05$.



Steps of a meta-analysis

3. Calculate an effect size that can be combined across studies to produce a quantitative summary of the findings.

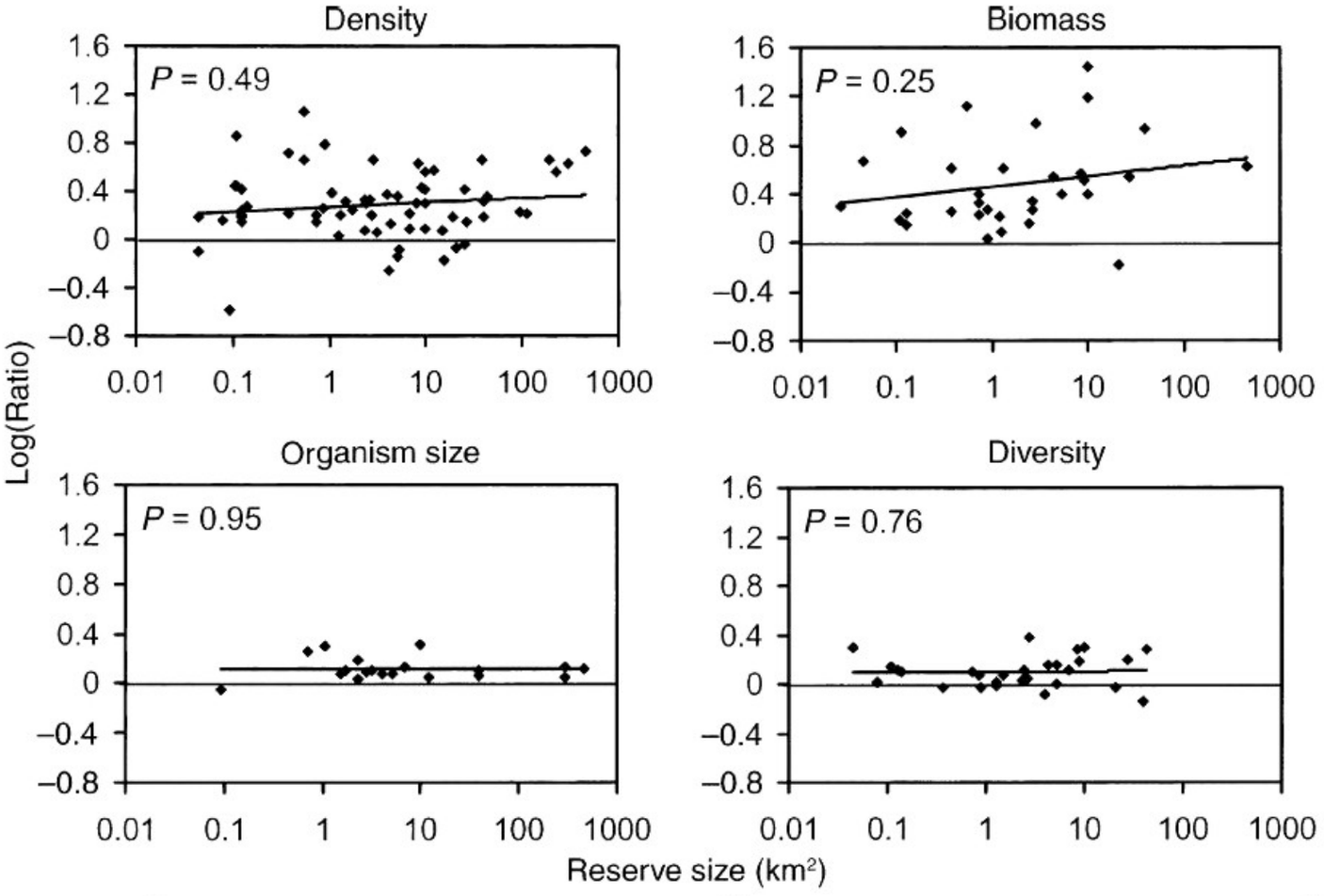
   - Correlation coefficient *r* is commonly used though not always ideal, because effect size depends on the range of the data.
   - Odds ratio – used in highly homogeneous studies (e.g., in tests of aspirin and myocardial infarctions).
   - Response ratio: $R = \bar{Y}_E / \bar{Y}_C$ or log of response ratio: $\ln(R)$
   - Standardized mean difference, Cohen's d or Hedges' g:

$$g = \frac{\bar{Y}_E - \bar{Y}_C}{s_{\text{pooled}}} J(m)$$

s is the pooled sample variance and $J(m)$ is a small-sample bias correction.

Example 3: Effectiveness of marine reserves

Halpern (2003) used the log of response ratio to compare marine reserves to comparison areas (or the same area before reserve establishment) in abundance and diversity of fish and/or invertebrates



Steps of a meta-analysis

4. Statistical inference on average effect size.

Fixed effects models

- Most commonly used in medical studies.
- Assumes that the multiple studies have the same mean, differing only because of sampling error. If every study were infinitely large, every study would yield an identical result. No heterogeneity among the studies.
- Perhaps never justified unless all studies conducted similarly and on the same species. This is rarely the case in ecology and evolution.

Steps of a meta-analysis

Random (mixed) effects models

- Random variation is present among means of studies in addition to sampling error.
- Individual studies are therefore estimating different treatment effects.
- Most interest is focused on the central value, or mean, of the distribution of effects.
- The idea of a random effects meta-analysis is also to understand the distribution of effects across different studies.

Steps of a meta-analysis

Fixed effect model

Effect size of each study *i* is $Y_i = \Theta + \varepsilon_i$

where Θ is the one "true" effect size, common to all studies.

Random effect model

Effect size of study *i* is $Y_i = \mu + \zeta_i + \varepsilon_i$

where μ is the grand mean and ζ_i is the deviation of the "true" effect size of study *i* from the grand mean.

The difference affects how each study is weighed when calculating the average effect size over all studies. We'll do this in the workshop.

`lmer()` in `lme4` package can't be used for random effects meta-analysis in R, because it won't calculate the necessary weights. Use dedicated packages available (e.g., `metafor`).

Steps of a meta-analysis

5. Look for effects of study quality. For example, are effect sizes different on average between studies that included blinding and those that did not?

6. Look for associations with variables that might explain heterogeneity of effect sizes among studies. For example, does the average effect size differ between studies carried out on women subjects and those on male subjects?

Example 4: Meta-analysis of competition in field experiments

Gurevitch et al (1992) study of inter- and intra-specific competition, looking only at studies published in 1980's

Vol. 140, No. 4 The American Naturalist October 1992

A META-ANALYSIS OF COMPETITION IN FIELD EXPERIMENTS

JESSICA GUREVITCH, LAURA L. MORROW, ALISON WALLACE, AND JOSEPH S. WALSH*

Department of Ecology and Evolution, State University of New York, Stony Brook, New York 11794-5245

Example 4: Meta-analysis of competition in field experiments

They looked for effects of study quality

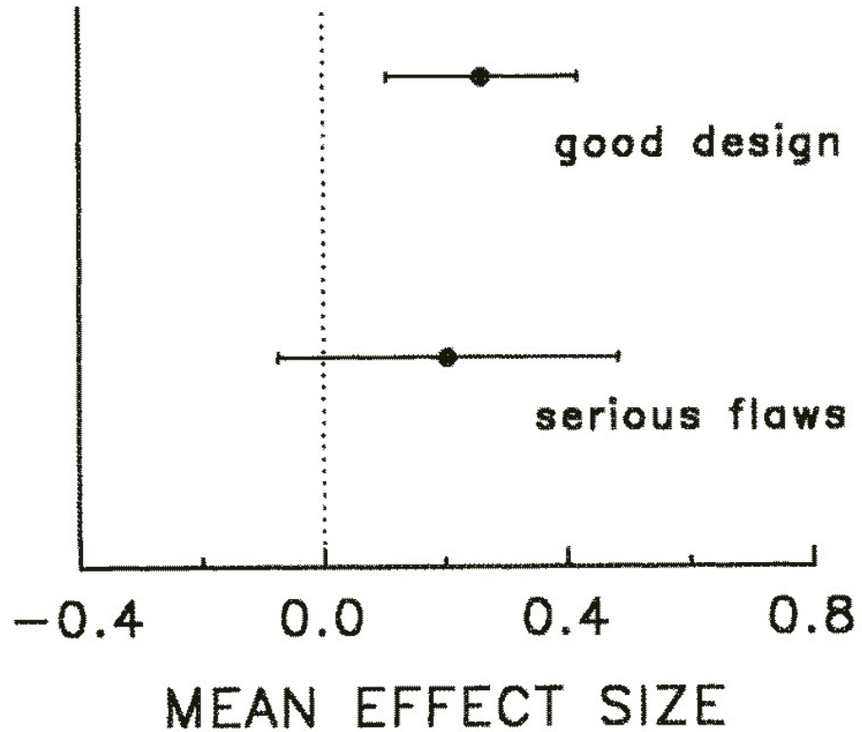


FIG. 12.—Mean effect size (d_+) and 95% CI for carnivores in experiments with good experimental designs or only minor design problems in contrast with those in experiments with serious problems in experimental design.

Example 4: Meta-analysis of competition in field experiments

They looked for associations with variables that might explain variation in effect size

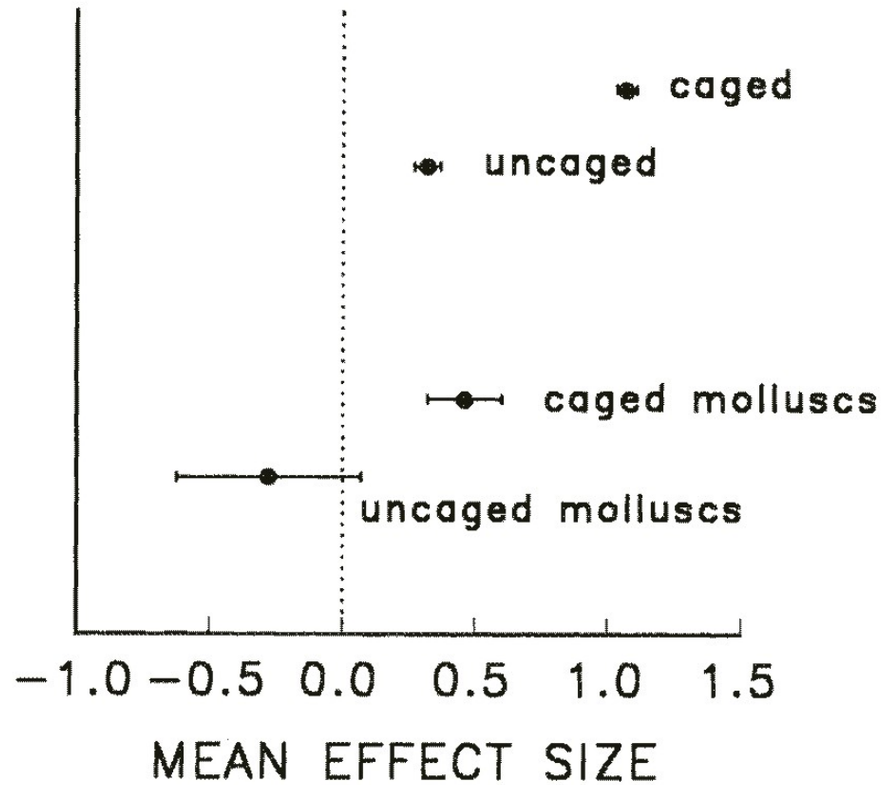


FIG. 7.—Mean effect size (d_+) and 95% CI of competition for all caged or enclosed organisms in contrast with all uncaged organisms (*top*) and mean effect size (d_+) of interspecific competition for marine mollusks in caged vs. uncaged trials (*bottom*).

File-drawer problem

In meta-analysis, the difficulties caused by publication bias are called the *file-drawer problem*, in reference to the unknown studies sitting unavailable in researchers' file drawers or hidden in obscure journals.

The *file-drawer problem* is the possible bias in estimates and tests caused by publication bias.

Funnel plot

Funnel plots can give an indication of the bias resulting from small studies.

Soma and Garamszegi (2011) used the Trimfill algorithm to fill in hypothetical missing studies in the funnel plot to achieve theoretical symmetry.

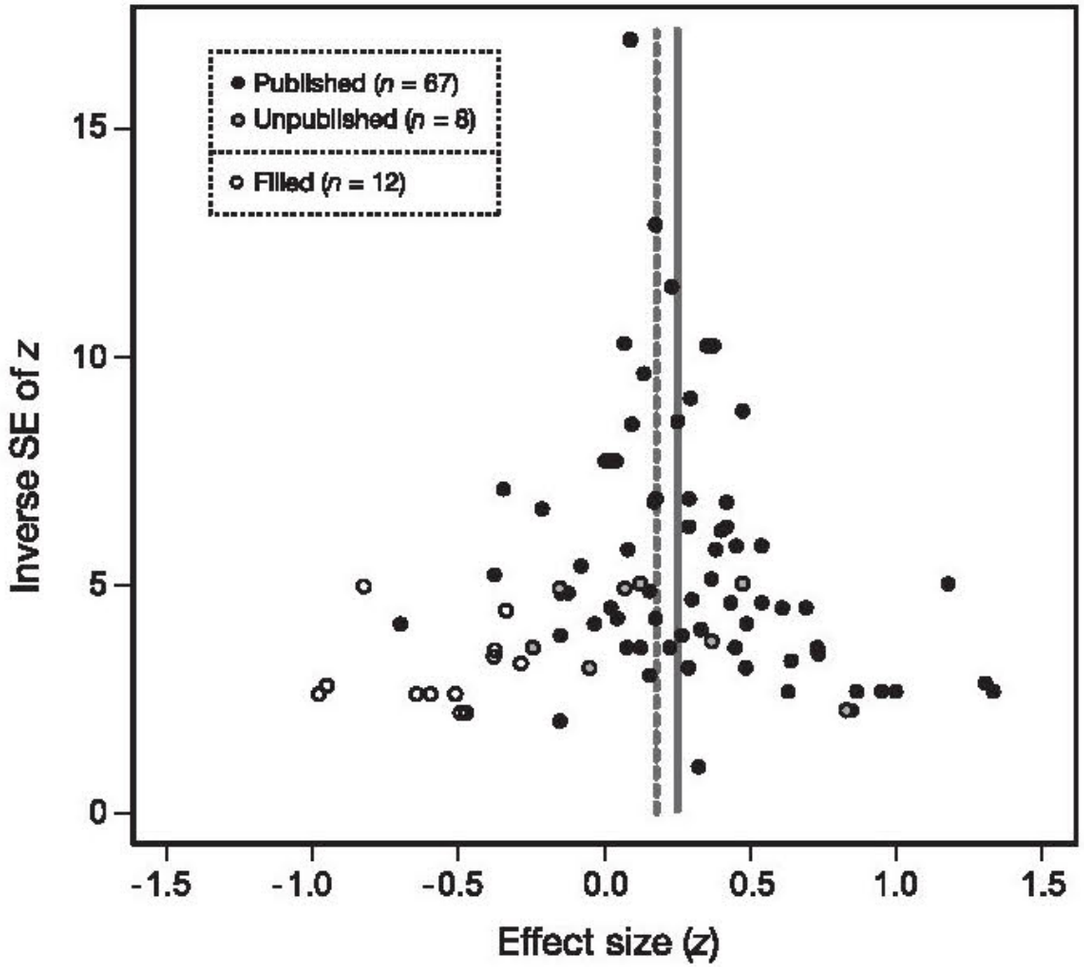


Figure 1
Funnel plot of effect sizes for the relationship between song complexity and reproductive success when using the multiple effect size data set (data set A). Black and gray circles show published and unpublished effect sizes, respectively. Solid and dotted lines show mean effect sizes before and after controlling for publication bias, in which theoretical missing data points (open circles, $n = 12$) were added to adjust funnel plot asymmetry.

Fail-safe number

The *fail-safe number* calculates how many missing studies would be needed to change the overall result of the meta-analysis.

Vilà et al (2011) estimated the number of studies that would have to be added to change the results of their invasive plant meta-analysis from significant to non-significant as 37,689. This was too implausible, so they concluded that their estimates were reliable.

(Fig 1a: top line refers to total plant production; other lines are effects on native plants and animals (1b))

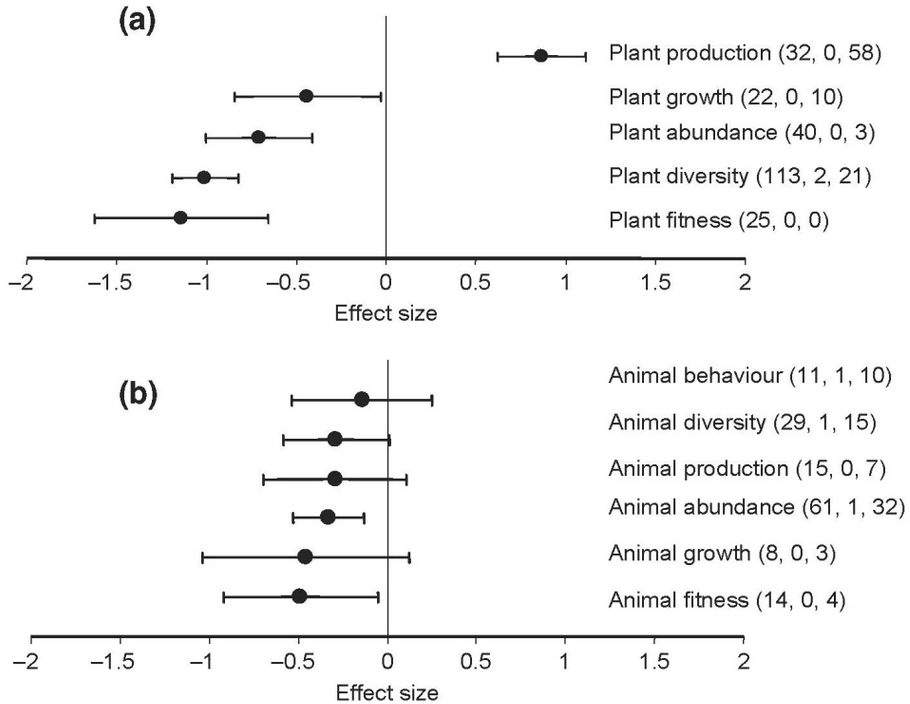


Figure 1 Mean effect size (Hedges' *d*) of differences between alien plant species impacts to (a) plant species and communities and (b) animal species and communities. The bars around the means denote bias-corrected 95%-bootstrap confidence intervals. A mean effect size is significantly different from zero when its 95% confidence interval do not bracket zero. Positive mean effect sizes indicate that the invaded plots had on average greater values for variables describing a particular impact type. The sample sizes with Hedges' $d < 0$, Hedges' $d = 0$ and Hedges' $d > 0$ are given next to the bars.

Make <u>your</u> results accessible to meta-analysis

Many published papers do not report enough information for meta-analysts to extract the numbers that they need. As a result, many otherwise relevant papers have to be discarded. Don't let this happen to your work.

- Always give sizes of effects and their standard errors. A *P*-value by itself is useless.
- Give estimates of the means and standard deviations of the important variables.
- Always indicate your sample sizes or degrees of freedom.
- Make the data accessible. Publish the raw data in the paper or deposit to an online archive such as Dryad.

Consider a meta-analysis for your first thesis chapter

Often, the first chapter of a thesis is a review of the literature. If your review is a systematic review, and you kept track of the important quantities and feature of each study, you may have enough for a quantitative component – a meta-analysis.

Best practices for conducting systematic review and meta-analysis

OPEN ACCESS Freely available online

PLoS MEDICINE

Guidelines and Guidance

Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement

David Moher[1,2*], Alessandro Liberati[3,4], Jennifer Tetzlaff[1], Douglas G. Altman[5], The PRISMA Group[¶]

1 Ottawa Methods Centre, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada, 2 Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada, 3 Università di Modena e Reggio Emilia, Modena, Italy, 4 Centro Cochrane Italiano, Istituto Ricerche Farmacologiche Mario Negri, Milan, Italy, 5 Centre for Statistics in Medicine, University of Oxford, Oxford, United Kingdom

Introduction

Systematic reviews and meta-analyses have become increasingly important in health care. Clinicians read them to keep up to date with their field [1,2], and they are often used as a starting point for developing clinical practice guidelines. Granting agencies may require a systematic review to ensure there is justification for further research [3], and some health care journals are moving in this direction [4]. As with all research, the value of a systematic review depends on what was done, what was found, and the clarity of reporting. As with other publications, the reporting quality of systematic reviews varies, limiting readers' ability to assess the strengths and weaknesses of those reviews.

Several early studies evaluated the quality of review reports. In 1987, Mulrow examined 50 review articles published in four leading medical journals in 1985 and 1986 and found that none met all eight explicit scientific criteria, such as a quality assessment of included studies [5]. In 1987, Sacks and colleagues [6] evaluated the adequacy of reporting of 83 meta-analyses on 23 characteristics in six domains. Reporting was generally poor; between one and 14 characteristics were adequately reported (mean = 7.7; standard deviation = 2.7). A 1996 update of this study found little improvement [7].

In 1996, to address the suboptimal reporting of meta-analyses, an international group developed a guidance called the

clinicians, medical editors, and a consumer. The objective of the Ottawa meeting was to revise and expand the QUOROM checklist and flow diagram, as needed.

The executive committee completed the following tasks, prior to the meeting: a systematic review of studies examining the quality of reporting of systematic reviews, and a comprehensive literature search to identify methodological and other articles that might inform the meeting, especially in relation to modifying checklist items. An international survey of review authors, consumers, and groups commissioning or using systematic reviews and meta-analyses was completed, including the International Network of Agencies for Health Technology Assessment (INAHTA) and the Guidelines International Network (GIN). The survey aimed to ascertain views of QUOROM, including the merits of the existing checklist items. The results of these activities were presented during the meeting and are summarized on the PRISMA Web site (http://www.prisma-statement.org/).

Only items deemed essential were retained or added to the checklist. Some additional items are nevertheless desirable, and review authors should include these, if relevant [10]. For example, it is useful to indicate whether the systematic review is an update [11] of a previous review, and to describe any changes in procedures from those described in the original protocol.

Best practices for conducting systematic review and meta-analysis

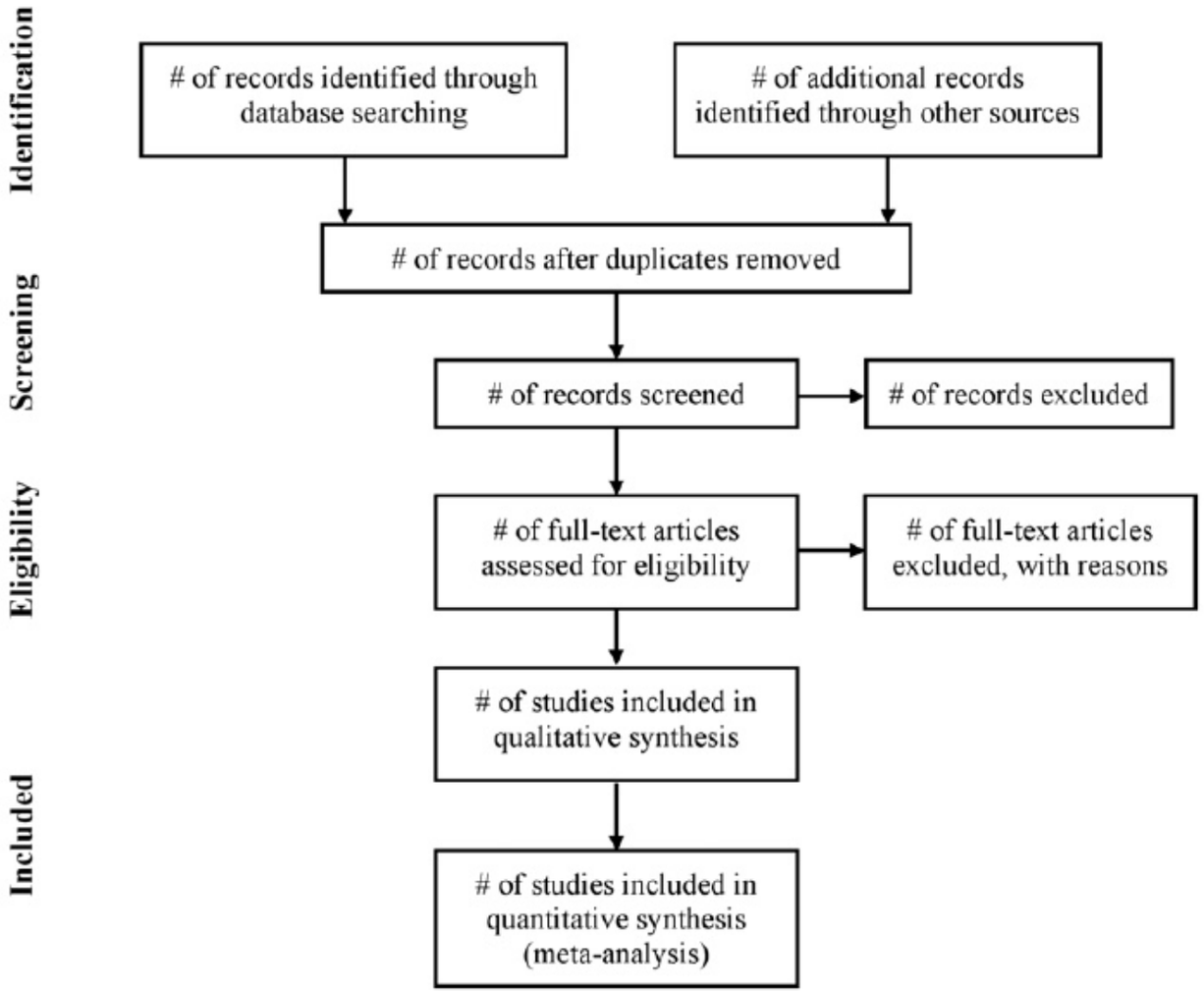


Figure 1. Flow of information through the different phases of a systematic review.
doi:10.1371/journal.pmed.1000097.g001

Best practices for conducting systematic review and meta-analysis

PRISMA detailed checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	

INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	

Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	
Risk of bias in	12	Describe methods used for assessing risk of bias	

individual studies		of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	

RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	

Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	

FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	

R toolkit for meta-analysis?

Methods in Ecology and Evolution

British Ecological Society

Methods in Ecology and Evolution 2016, 7, 323–330 doi: 10.1111/2041-210X.12472

APPLICATION

Facilitating systematic reviews, data extraction and meta-analysis with the METAGEAR package for R

Marc J. Lajeunesse*

Department of Integrative Biology, University of South Florida, 4202 East Fowler Avenue, Tampa, FL 33620, USA

Summary

1. The R package ecosystem is rich in tools for the statistics of meta-analysis. However, there are few resources available to facilitate research synthesis as a whole.
2. Here, I present the METAGEAR package for R. It is a comprehensive, multifunctional toolbox with capabilities aimed to cover much of the research synthesis taxonomy: from applying a systematic review approach to objectively assemble and screen the literature, to extracting data from studies, and to finally summarize and analyse these data with the statistics of meta-analysis.
3. Current functionalities of METAGEAR include the following: an abstract screener GUI to efficiently sieve bibliographic information from large numbers of candidate studies; tools to assign screening effort across multiple collaborators/reviewers and to assess inter-reviewer reliability using kappa statistics; PDF downloader to automate the retrieval of journal articles from online data bases; automated data extractions from scatter-plots, box-plots and bar-plots; PRISMA flow diagrams; simple imputation tools to fill gaps in incomplete or missing study parameters; generation of random-effects sizes for Hedges' *d*, log response ratio, odds ratio and correlation coefficients for Monte Carlo experiments; covariance equations for modelling dependencies among multiple effect

comment

How to do meta-analysis of open datasets

The amount of open data in ecology and evolution is increasing rapidly, yet this resource remains underused. Here, we introduce a new framework and case study for conducting meta-analyses of open datasets, and discuss its benefits and current limitations.

Antica Culina, Thomas W. Crowther, Jip J. C. Ramakers, Phillip Gienapp and Marcel E. Visser

In recent decades, the meta-analysis approach has emerged as the most valuable avenue for scientific progress, along with empirical studies and theoretical models[1,2]. Traditional meta-analysis combines results from a number of studies (ideally all) conducted on the same research question, to statistically summarize findings, evaluate discrepancies and detect generalizable effects[2]. The ability to detect overarching patterns makes meta-analyses extremely relevant to evolutionary ecology, which is characterized by highly complex systems, heterogeneous environments and variable methodologies[3,4].

Systematic advances in the meta-analysis approach over the past decade have been intended to improve the transparency, replicability, reliability and impact of data synthesis efforts[2,5–7]. However, despite these advances, the major outstanding limitation of any synthesis remains the challenge and gain a comprehensive understanding of the available information has never been greater[9]. Yet, unlike other scientific fields, this resource remains relatively unexploited in the field of ecology and evolution[10,11].

Data retrieval for meta-analysis

Here, we describe how to transparently retrieve and select data, when the information retrieval starts from published (open) datasets, rather than from published studies. Our standard is based on existing guidelines for the information retrieval in ecological/evolutionary meta-analysis[5,6,12,13], but adapted specifically for open data. The retrieval and selection process should be highly transparent — we provide a checklist of the information that needs to be recorded (Table 1). This information should ideally be supported by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses[14] diagram (Supplementary Fig. 1).

data needed to answer the meta-analysis question (or test hypothesis), set appropriate exclusion/inclusion criteria and choose the search terms (used in a search for the relevant data). This is followed by the data search. In evolutionary ecology, datasets are usually scattered across various repositories (for example, Dryad, Figshare, Zenodo) or published in the supplementary materials associated with a paper. Thus, an effective search should be conducted using data-harvesting platforms that crawl through many different research data repositories that host research data (like Web of Science crawls through journals in a search for articles); some also explore supplementary materials of published papers for additional information. A complete overview of how to navigate the data landscape by using data search platforms can be found in ref. [15]. We suggest using DataCite, BASE search engine and DataONE (see Box 1). The original

Discussion paper next week:

Multivariate analysis:

Download from “**Handouts**” tab on course web site.

Presenters: Tahnee & Clark

Moderators: Javad & Alieu