School of Biological Sciences

Introduction to Meta-Analysis in Environmental Sciences

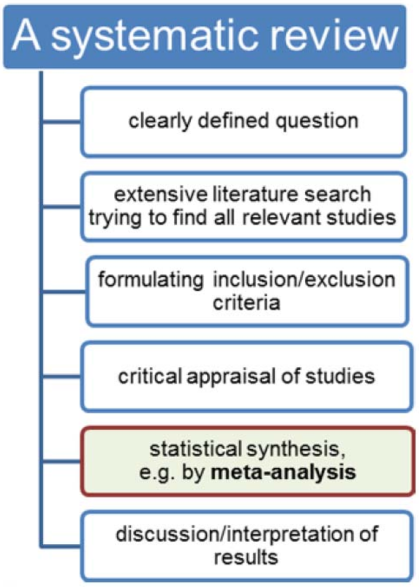
Julia Koricheva

30 December 2020

ROYAL HOLLOWAY UNIVERSITY OF LONDON

Glossary

- Research synthesis = review of primary research on a given topic with a purpose of integrating the findings (creating generalizations, conflict resolution)
- Meta-analysis = a set of statistical methods for combining outcomes (effect sizes) across different data sets addressing the same research question to examine patterns of response across these data sets and sources of heterogeneity in outcomes
- Systematic review = research synthesis on a precisely defined topic using explicit methods to identify, select, critically appraise, and analyse relevant research



Meta analysis is the use of statistical methods within a systematic review

What is wrong with narrative reviews?

- No strict criteria for selection of studies for review or for judging study quality
  - ✓ high degree of subjectivity
  - ✓ low repeatability
- Low efficiency in handling a large number of studies
- Limited ability to deal with variation in study outcomes
  - ✓ the results of studies are often found to be “inconsistent”, “inconclusive” or “conflicting”
  - ✓ little help in conflict resolution and decision making

Meta-analysis: objectives

- Meta analysis is a set of statistical methods for amalgamating, summarizing, and reviewing previous quantitative research.
- By combining information from all relevant studies, MA can provide more precise estimates of the effects of interest than those derived from the individual studies.
- Methods of MA also facilitate investigations of the consistency of evidence across studies, and the exploration of differences across studies.

History of meta-analysis

1976 - the term "meta-analysis" is coined

Glass GV (1976) Primary, secondary, and meta-analysis. Educational Researcher 5: 3-8

At the same time, two other psychologists were doing similar work: Robert Rosenthal and Frank Schmidt. All three are considered the founders of modern meta-analysis.

More history of MA

1977 - first large-scale application of meta-analysis:

- Smith ML & Glass GV (1977) Meta-analysis of psychotherapy outcome studies. American Psychologist 32: 752-760
  - 833 tests of the treatment

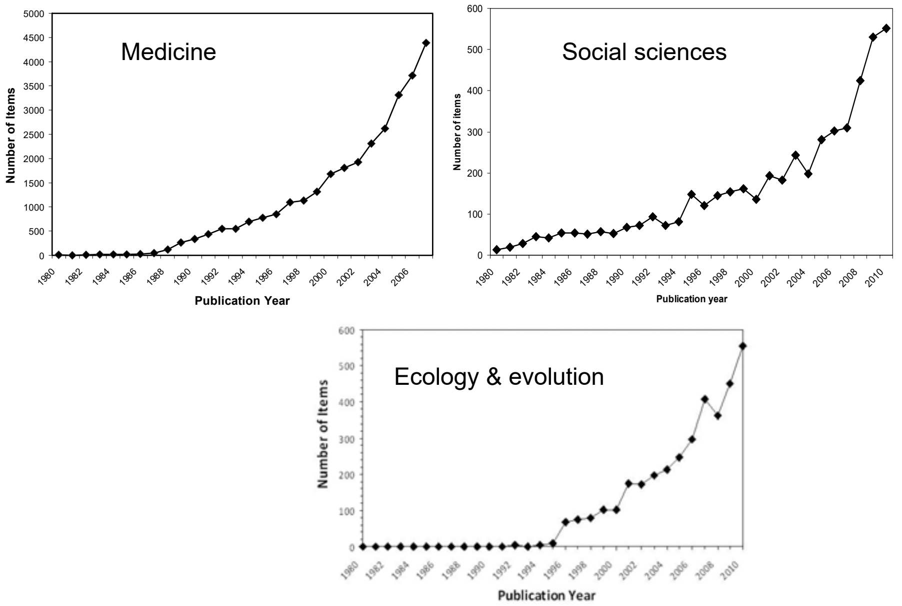
1978 - first critique of the method:

- Eysenck HJ (1978) An exercise in mega-silliness. American Psychologist 33: 517

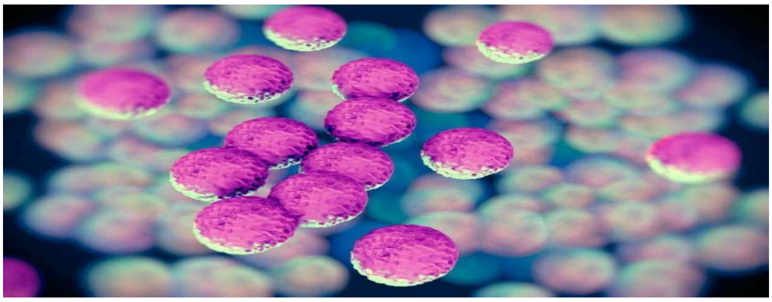
Statistical methods for MA

Proper statistical foundations of meta-analysis were developed by Larry Hedges and Ingram Olkin in early 1980s and published in their seminal book Statistical Methods for Meta Analysis, Academic Press, 1985.

Use of meta-analysis in different disciplines



The Cochrane Collaboration: www.cochrane.org



- 0.5 million trial reports   3,500 Cochrane reviews
- 10,000 other reviews
- 45,000 reviews needed to cover existing trials? (Mallett & Clarke, 2003)

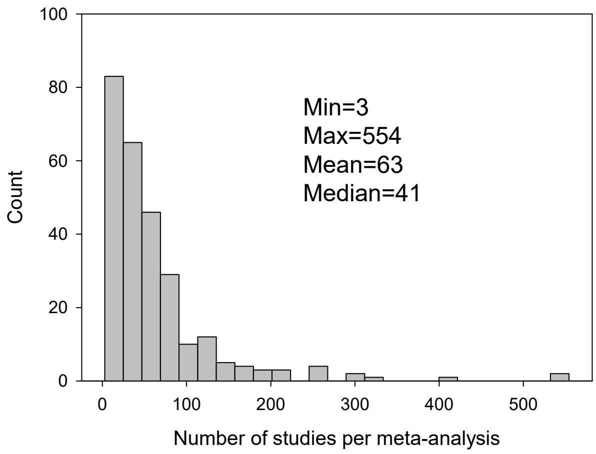
Other collaborations

- The Campbell Collaboration investigates the effects of social interventions
- Collaboration for Environmental Evidence: issues of greatest concern to environmental policy and practice

When is meta-analysis most useful?

- There is a moderate to large amount of empirical work available
- The results are variable across studies
- The expected magnitude of the effect is relatively weak
- The sample sizes of individual studies are limited for some reason

How many studies are needed for a meta-analysis?



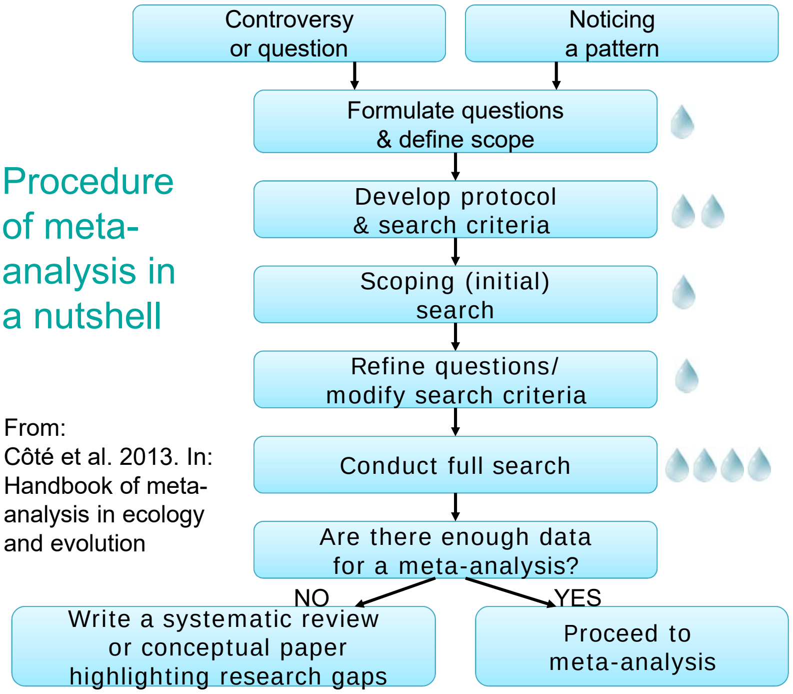
Review of 322 meta-analyses in plant ecology published in 1996-2013 (Koricheva & Gurevitch 2014)

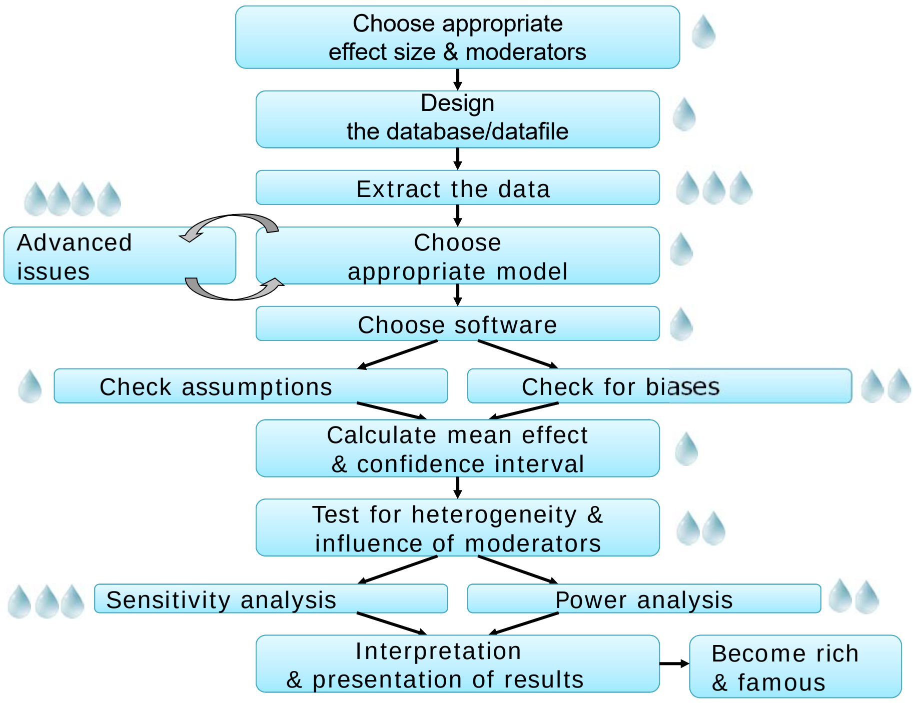
Applications of meta-analysis in ecology

- Synthesizing results of empirical tests of predictions of various hypotheses and theories -> scope & applicability limits of the current hypotheses, nature of mediating processes
- Generating new hypotheses and testing hypotheses which are difficult to test within single study (cross-ecosystem comparisons, latitudinal patterns, cross-taxa comparisons)
- Assessing impact of major environmental drivers (climate change, invasive species, biodiversity loss)
- Assessing effectiveness of conservation and management strategies -> evidence-based conservation and management
- Combining results of multi-site experiments or long-term experiments (within-study meta-analysis)
- Identifying knowledge gaps

Procedure of meta-analysis in a nutshell

From:
Côté et al. 2013. In: Handbook of meta-analysis in ecology and evolution





From Cote and Jennions 2013:

…"the hardest part of a meta-analysis is not mastering complex statistics or software but the labor involved in gathering, evaluating and assimilating papers. In particular, meta-analysis requires the additional work of extracting data, which is often frustrating and challenging, and is not typically part of an old-fashioned narrative review."

Types of ecological data and their conversion to effect sizes

Meta-analysis vs null hypothesis significance testing

- Most statistical approaches commonly used in biological sciences are based on **null hypothesis significance testing** (*P* values)
- **In contrast**, meta-analysis is focused on assessment of the **magnitude** of an effect of interest (=effect size), the precision of its estimate and causes of variation in effect size among studies

What is an effect size?

- Effect size expresses the magnitude of an effect of interest (e.g. effect of elevated CO_2 on photosynthesis or strength of the relationship between two life history traits)
- Effect sizes are the common currency which allows combination of results from different studies into a single, standardized and comparable metric of outcome

Requirements for the effect size metrics

- Effect size metrics should not depend on aspects of study design which may vary from study to study (e.g. sample size)
- Estimate of effect size should be computable from the summary information provided in a primary study (i.e. no re-analysis of raw data should be necessary)
- Sampling distribution of the metric should be known to allow calculation of variances and confidence intervals
- Effect size metric should be easily interpretable in terms of its biological/practical significance

Types of primary data -> effect sizes

- Comparison of two groups (e.g. control and experimental) in terms of continuous response variables

  -> effect sizes based on means

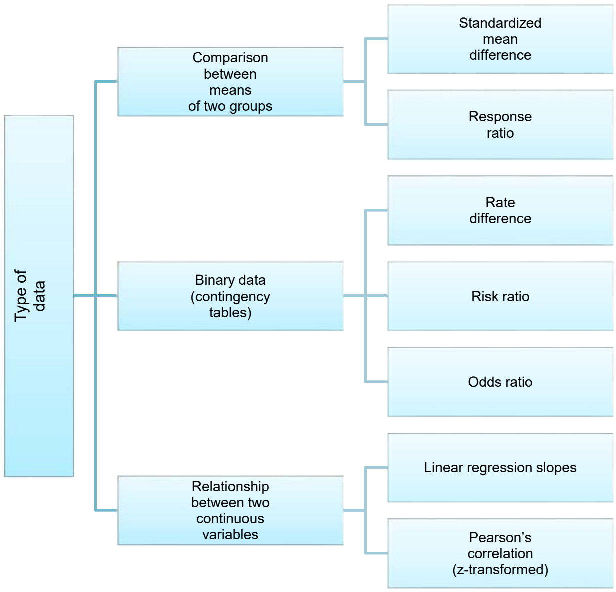
- Comparison of two groups in terms of categorical response variable (2 x 2 contingency table)

  -> effect sizes based on binary data

- Relationship between two continuous variables

  -> effect sizes based on correlations

Choosing effect size measure



Effect sizes based on means:

standardized mean difference (SMD, Hedges' d)

$$d = \frac{(\overline{X}e - \overline{X}c)}{s} J$$

Xe – mean of the experimental group

Xc – mean of the control group

s – pooled standard deviation

J – correction term that removes small-sample-size bias

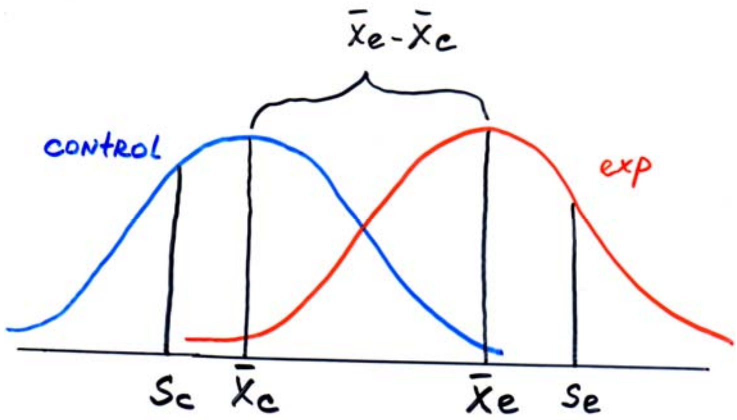
Pooled standard deviation:

$$s = \sqrt{\frac{(N_e - 1)s_e^2 + (N_c - 1)s_c^2}{N_e + N_c - 2}}$$

Correction term:

$$J = 1 - \frac{3}{4(N_e + N_c - 2) - 1}$$

Standardized difference between means



The variance of Hedges' *d*

$$v_d = \frac{N_e + N_c}{N_e N_c} + \frac{d^2}{2(N_e + N_c)}$$

Ne – sample size of the experimental group

Nc – sample size of the control group

Hedges' d: example

January 2012 | Volume 7 | Issue 1 | e28595

OPEN ACCESS Freely available online PLoS one

A Meta-Analysis of Seaweed Impacts on Seagrasses: Generalities and Knowledge Gaps

Mads S. Thomsen[1,2*], Thomas Wernberg[1,2,3], Aschwin H. Engelen[4], Fernando Tuya[2,5], Mat A. Vanderklift[6], Marianne Holmer[7], Karen J. McGlathery[8], Francisco Arenas[9], Jonne Kotta[10], Brian R. Silliman[11]

Xc = abundance of seagrasses on plots without seaweeds

Xe = abundance of seagrasses on plots with seaweeds

Positive effect = higher abundance of seagrasses on plots with seaweeds

Hedges' d

Advantages:

- Scale-free

Problems

- Interpretation of the magnitude of the effect
- Difference in *d* may reflect either differences in the magnitude of the effect or in variance among studies
- Some data needed for calculation of *d* (most commonly SD or sample sizes) are often missing
- Variance of d depends on d

Interpretation of magnitude of *d*

- Cohen's benchmarks:

|d | = 0.2 – small effects

|d | = 0.5 – moderate effects

|d | = 0.8 – large effects

|d | = 0.63 (0.48-0.78) for 21 ecological MA (Møller & Jennions 2002)

- Comparison with mean effects reported in other meta-analyses on similar topics

Effect sizes based on means: response ratio

$$\ln R = \ln\left(\frac{\overline{X}_e}{\overline{X}_c}\right) = \ln(\overline{X}_e) - \ln(\overline{X}_c)$$

Variance:

$$v_{\ln R} = \frac{(s_e)^2}{N_e(\overline{X}_e)^2} + \frac{(s_c)^2}{N_c(\overline{X}_c)^2}$$

Hedges et al. (1999) Ecology 80: 1150-1156

Log response ratio: example

American Journal of Botany

American Journal of Botany 98(3): 539–548. 2011.

A SYNTHESIS OF PLANT INVASION EFFECTS ON BIODIVERSITY ACROSS SPATIAL SCALES[1]

KRISTIN I. POWELL[2], JONATHAN M. CHASE, AND TIFFANY M. KNIGHT

Department of Biology, Washington University in Saint Louis, 1 Brookings Drive, Campus Box 1137, Saint Louis, Missouri 63130 USA

- Ln *R* =ln (uninvaded species richness) – ln (invaded species richness)

Advantages of response ratio

- Easily interpretable: $R = \exp\ (\ln R)$
- Effect sizes are not affected by different variance in control and experimental groups
- SD/SE are not needed for calculation of the effect size (but needed to calculate variance)

Response ratio: limitations

- Response ratio can only be used if:

✓ Neither of the means is equal to 0

✓ Both means have the same sign

✓ $\frac{\sqrt{N_c} \times X_c}{SD_c}$ <3 (Hedges et al. 1999)

Response ratio: limitations

Ecology, 96(8), 2015, pp. 2056–2063

Bias and correction for the log response ratio in ecological meta-analysis

MARC J. LAJEUNESSE[1]

- log response ratio is biased when quantifying the outcome of studies with small sample sizes
- two new estimators that help correct this small-sample bias
- modified diagnostics for assessing when the response ratio is appropriate for ecological meta-analysis:

$$\frac{\bar{X}}{\mathrm{SD}}\left(\frac{4N^{3/2}}{1+4N}\right) \geq 3$$

Binary data

	Treatment	Control	Total
Response	X_T	X_C	$X_T + X_C$
No Response	$n_T - X_T$	$n_C - X_C$	$n_C + n_T - X_T - X_C$
Total	n_T	n_c	$n_T + n_C$

	Treatment	Control	Total
Dead	5	10	15
Alive	95	90	185
Total	100	100	200

Risk difference (*RD*)

	Treatment	Control
Response	X_T	X_C
No Response	$n_T - X_T$	$n_C - X_C$
Total	n_T	n_C

- Probabilities of response: p_T and p_C with estimates

$$\hat{p}_T = \frac{X_T + g}{n_T + 2g} \qquad \hat{p}_C = \frac{X_C + g}{n_C + 2g}$$ where g is a small constant (typically 0.50)

- Risk (rate) difference: $RD = p_T - p_C$
- Variance of sample *RD*:

$$v_{RD} = \frac{p_T(1-p_T)}{n_T} + \frac{p_C(1-p_C)}{n_C}$$

Risk ratio (*RR*)

	Treatment	Control
Response	X_T	X_C
No Response	$n_T - X_T$	$n_C - X_C$
Total	n_T	n_C

- Probabilities of response: p_T and p_C
- Risk ratio or relative risk: $RR = \frac{p_T}{p_C}$
- Usually ln *RR* is used in the analysis
- Variance of sample *RR*: $v_{\ln RR} = \frac{(1-p_T)}{n_T p_T} + \frac{(1-p_C)}{n_C p_C}$

Risk ratio: example

Biological Conservation 142 (2009) 935–947

Impacts of grazing on lowland heathland in north-west Europe

Adrian C. Newton[a,], Gavin B. Stewart[b], Gillian Myers[a], Anita Diaz[a], Sophie Lake[c], James M. Bullock[d], Andrew S. Pullin[b]*

	Grazed	Ungrazed
Graminoid cover	X_T	X_C
Ericoid cover	$n_T - X_T$	$n_C - X_C$

Graminoid cover = response; risk ratio > 1 means that graminoid:ericoid cover increases as a result of grazing

Odds ratio (*OR*)

	Treatment	Control
Response	X_T	X_C
No response	$n_T - X_T$	$n_C - X_C$

OR= $[p_T/(1-p_T)] : [p_C/(1-p_C)]$

= the odds of an event happening in the treatment group relative to the odds of the same event in the control group

Estimated by $[X_T/(n_T - X_T)] : [X_C/(n_C - X_C)]$

Properties of Odds Ratio

- OR= $p_T/(1-p_T) : p_C/(1-p_C)$
- LOR=ln (OR) has unlimited range.
- Sample LOR is approximately normally distributed for large sample sizes.
- Var(LOR) =

$$[n_T\, p_T(1-p_T)]^{-1} + [n_C\, p_C(1-p_C)]^{-1}$$

41

Odds ratio: example

Journal of Ecology 2005
93, 748–757

Is the change of plant–plant interactions with abiotic stress predictable? A meta-analysis of field results in arid environments

FERNANDO T. MAESTRE*, FERNANDO VALLADARES† and JAMES F. REYNOLDS*‡

An odds ratio metric was calculated as the ratio of the odds of survival in the presence of neighbours (neighbour treatment) to the odds of survival in their absence (control)

	Control	Neighbours present
Died	X_T	X_C
Survived	$n_T - X_T$	$n_C - X_C$

Effect sizes based on correlations:
Pearson's *r*

- varies from –1 to +1
- Cohen's "rules-of-thumb":
  - | *r* | = 0.10 – small
  - | *r* | = 0.25 – medium
  - | *r* | = 0.40 - large
- coefficient of determination (r^2)
- mean | *r* | in ecology = 0.19 (95% CI 0.14-0.25) (Møller & Jennions 2002)

Pearson's product–moment correlation coefficient (r)

- The distribution of *r* is complicated and its variance depends on *r*
- Solution: Fisher's *z*-transformation

$$z = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right)$$

variance: $$v_z = \frac{1}{N-3}$$

back-transformation: $$r = \frac{e^{2z}-1}{e^{2z}+1}$$

Pearson’s correlations: example

Ecology, 83(1), 2002, pp. 176–190

META-ANALYSIS OF SOURCES OF VARIATION IN FITNESS COSTS OF PLANT ANTIHERBIVORE DEFENSES

JULIA KORICHEVA[1]

Effect size = correlation between measures of plant allocation to defence and plant growth/reproduction

Negative correlation -> indication of fitness costs of defences, growth/defence trade-offs

Slopes from linear regressions

- Slopes from linear regression can be used as effect sizes providing that:
  - ✓ The slope in every study is measured in the same units
  - ✓ Measure of variance of the estimate of the slope is available (or can be estimated from other statistics, see Rosenberg et al. 2013 In: Handbook of Meta-Analysis in Ecology and Evolution)
  - ✓ Overall not straightforward, see Becker&Wu (2007), Statistical Science, 22, 414-429.

Regression slopes: example

Journal of Ecology 2013, **101**, 784–795 doi: 10.1111/1365-2745.12074

Latitudinal gradients as natural laboratories to infer species' responses to temperature

Pieter De Frenne[1,2*], Bente J. Graae[3], Francisco Rodríguez-Sánchez[2], Annette Kolb[4], Olivier Chabrerie[5], Guillaume Decocq[5], Hanne De Kort[6], An De Schrijver[1], Martin Diekmann[4], Ove Eriksson[7], Robert Gruwez[1], Martin Hermy[8], Jonathan Lenoir[5], Jan Plue[9], David A. Coomes[2] and Kris Verheyen[1]

- Slopes of regression between several plant traits (height, SLA, seed mass, N:P ratio) and latitude
- A negative slope coefficient indicates higher trait values in low-latitude than in high-latitude populations, and vice versa.

Using non-standard metrics of effect size

- If effect of interest has known distribution, variances are relatively easy to derive:
  - ✓ Mean and variance are equal for Poisson distribution (rate of occurrence)
- Three requirements if using non-standard metrics:
  - ✓ The measure of effect is comparable across studies
  - ✓ Sampling variance can be obtained
  - ✓ Do not do it without consulting a statistician!

Retrieval of data from primary studies

- Choosing a metric of effect size
- Obtaining effect sizes or data needed for their calculation from text, tables or graphs
  - ✓ scanning and digitizing graphs (e.g. ImageJ, WebPlotDigitizer)
- Calculating effect sizes from raw data
- Obtaining effect sizes from test statistics
- Converting one measure of effect size into another
- Converting other measures of variance into SD

Converting other measures of variance into SD

- SE -> SD:

$$SD = SE \times \sqrt{N}$$

- 95% CI ->SD:

$$SD = \frac{UCL - \bar{X}}{1.96}$$

where UCL is upper confidence limit and X is the mean effect

Converting Pearson's correlation coefficient (r) into SMD (d)

- Pearson's r can be converted into Hedges *d*:

$$d = \sqrt{\frac{N(N-2)}{n_T n_C}} \frac{r}{\sqrt{1-r^2}}$$

$$r = \frac{d}{\sqrt{d^2 + \frac{N(N-2)}{n_T n_C}}}$$

- Correlation r here is between continuous observations and the binary sample indicator (T or C)

Pearson's correlation coefficient (*r*)

- could be obtained from various test statistics:

Statistic	Conversion equation	Degrees of freedom
Z (standard normal deviate)	$r = \frac{Z}{\sqrt{N}}$	n/a
t	$r = \sqrt{\frac{t^2}{t^2 + df}}$	d.f. = $N_1 + N_2 - 2$
$F_{[1,df]}$	$r = \sqrt{\frac{F}{F + df}}$	Numerator = 1 d.f. Denominator = $N - 1$ df
$\chi^2_{(1)}$	$r = \sqrt{\frac{\chi^2_{(1)}}{N}}$	1

Introduction to mini-meta-analysis exercise

What to do?

- Refine the question for meta-analysis
- Decide on inclusion criteria
- Decide on metric of the effect size
- Decide on moderators (explanatory variables)
- Design extraction spreadsheet (include study ID columns, data needed for effect size calculation, moderator variables)
- Extract data from your allocated papers
- Run analyses in metafor
- Prepare Power Point presentation for Friday (15 min per group)

Question formulation: PICO method

P=population

I=intervention

C=comparator

O=outcome

”What is a 5-year overall mortality (*outcome*) in adults with high levels of cholesterol (*population*) taking statins (*intervention*) compared with those treated with low fat diet (*comparator*)?”

PICO: ecological example

- Effects of wind farms on birds

- What is the abundance (*outcome*) of breeding birds (*population*) on wind farms (*intervention)* compared to nearby reference sites (*comparator*)?

Example of extraction spreadsheet

Study and data ID		Potential moderators			Data for calculation of the effect measure		Data source
Study	Species	Taxon	Habitat	Status	r	N	Source
Adler et al. 1986	Peromyscus leucopus	mammal	interior	resident	0.33	21	Table 1
Ambuel & Temple 1983	Pheuticus ludovicanus	bird	edge	migrant	0.44	26	Figure 3B
Bach 1984	Acalymma vittatum	Insect	interior	resident	-0.04	21	t test p. 355
Bach 1984	Acalymma innubom	insect	interior	resident	0.66	18	t test p. 356