Meta-analysis in biological and environmental sciences

yDiv/HIGRADE course 23-26 October 2017

Dr. Dylan Craven, Dr. Katharina Gerstner

iDiv is a research centre of the

DFG Deutsche Forschungsgemeinschaft

Day 1	**Introduction to meta-analysis** • What is a meta-analysis? • Examples of meta-analyses • Why performing a meta-analysis? • Procedure of meta-analysis in a nutshell **Searching the literature** **Effect sizes and moderators** **Data extraction/Coding**
Day 2	**Meta-analytic models** • Fixed effects model • Random effects model • Mixed effects/hierarchical model **Quantifying and explaining heterogeneity**
Day 3	**Assumptions, biases and confounding effects** • Variance homogeneity and normality of residuals • Publication bias • Sensitivity analysis **Interpretation and presentation of results** • Format for meta-analysis report • PRISMA flow diagram • Forest plots
Day 4	**Methodological issues, advances, and common mistakes** • Non-independence among effect sizes • Non-independence of moderators • Missing data **Criticism of meta-analysis**

Meta-analysis in biological and environmental sciences

Meta-analytic models

Procedure of meta-analysis

1. Transformation of data or test statistics from individual studies into effect sizes
2. Combining effect sizes from individual studies into a common estimate of the magnitude of the effect
3. Estimating the significance of the overall effect
4. Estimating the degree of heterogeneity of the effect between studies
5. Exploring the causes of heterogeneity in effect

Considerations when choosing a model

Meta-analysis typically involves

- within-study variation (i.e. sampling variance)
- between-study variation (e.g. due to differential study design or covariables)

Meta-analysis models depend on

- aim of the meta-analysis
- assumptions about the influence of within- and between-study variation on effect size estimates

Why not to use standard statistical procedures on effect size estimates?

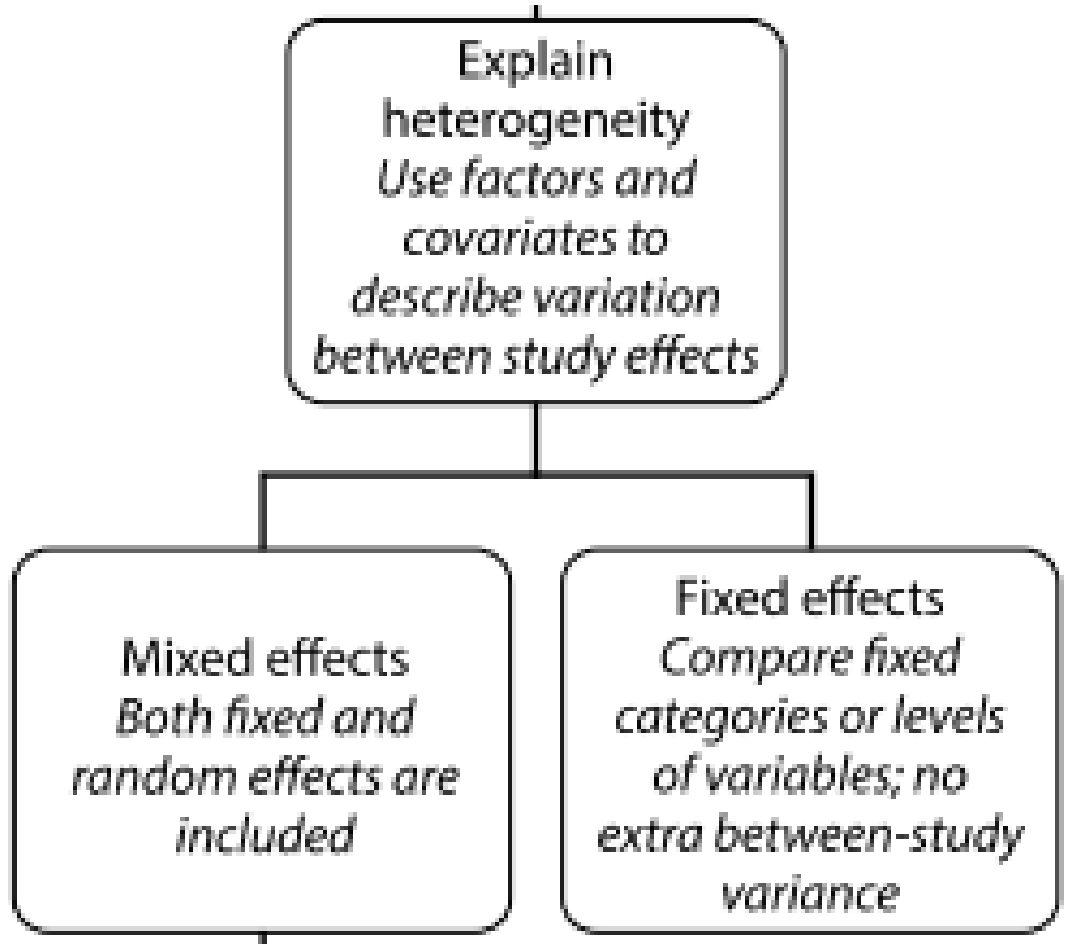
- These procedures do not test whether the variability in effect sizes is due solely to sampling error
- The assumption of homogeneity of variances is often violated in meta-data and transformations usually do not help

Approaches to inference in meta-analysis

1. Moment and least squares approaches
   - Estimate parameters by matching moments or minimizing the squared residuals
   - Less amenable to more complex modelling
2. Maximum likelihood approach
   - Choose parameters that maximize the probability of the data
   - Allows more complex modelling (e.g. based on effects with non-normal distribution, nested models etc)
3. Bayesian approach
   - Makes inferences on the posterior distribution of the model parameters, given the data
   - Allows complex and flexible modelling, non-normal distributions, missing data, incorporation of other info in meta-analysis (e.g. expert opinions)
   - Computationally demanding

Choosing an appropriate model

Fixed, random and mixed effect model



Koricheva et al. (2013), Fig. 8.1

Choosing an appropriate model

Fixed-effects model

Based on the assumption that studies share one common/fixed effect, i.e. no variation between study effects

$$\theta_i = \mu$$

θ_i observed effect in study *i*
μ overall effect

The only variation in effect sizes is due to sampling error *e*

$$T_i = \theta_i + e_i$$

T_i estimated effect size for study *i*

Alternatively: $\theta_i \sim N(\mu, e_i)$

Choosing an appropriate model

Random-effects model

In addition to sampling error e_i there is a true random component of variation in effect sizes between studies ε_i

The true effect size is expected to differ among studies:

$$T_i = \theta_i + e_i$$
$$\theta_i = \mu + \varepsilon_i$$

T_i estimated effect size for study *i*
θ_i observed effect in study *i*
μ overall effect

Alternatively: $\theta_i \sim N(\mu, e_i + \varepsilon_i)$

Choosing an appropriate model

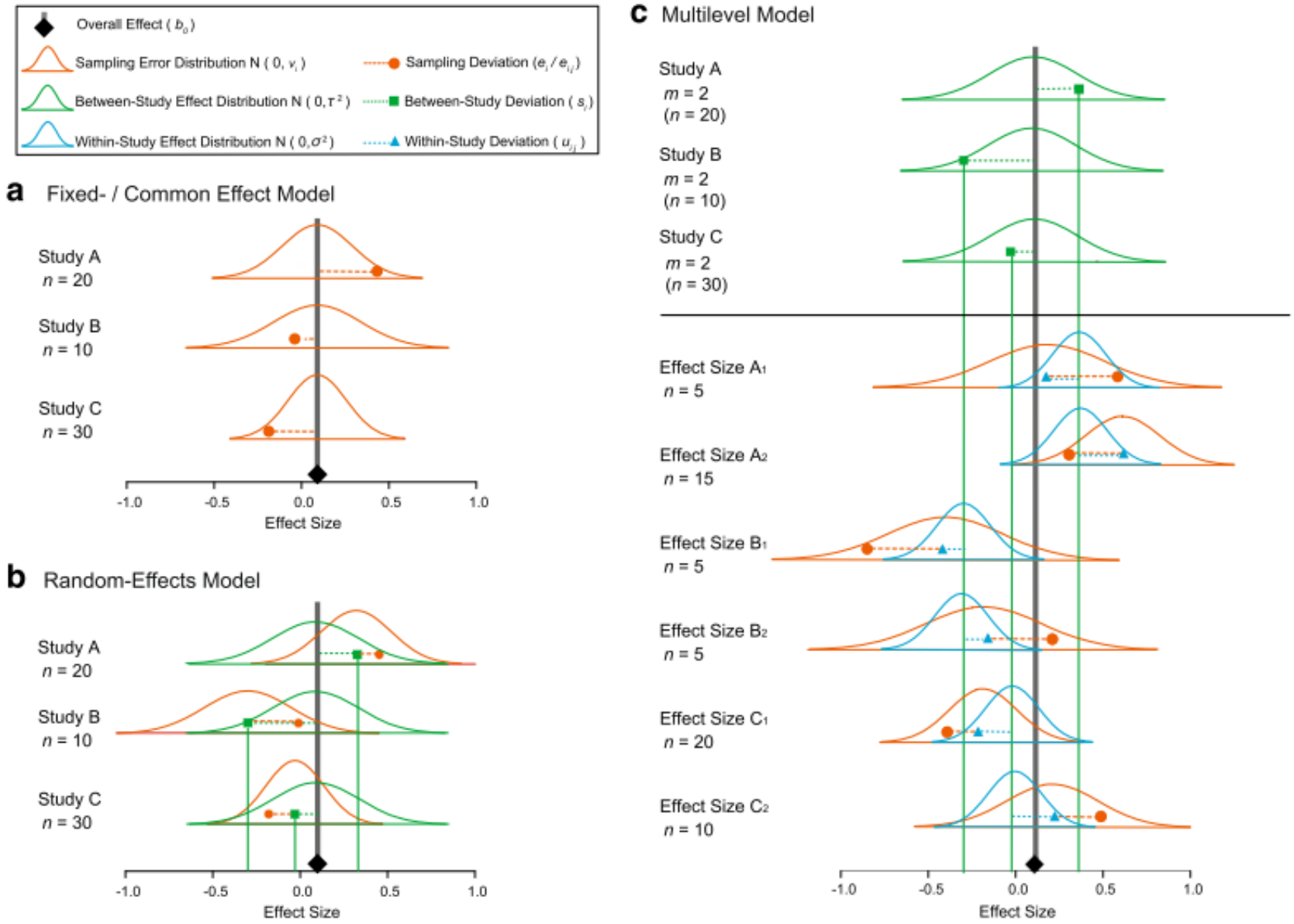
Fixed, random and mixed effect model

- Fixed-effects model is appropriate if:
  - we can assume that variation between studies is negligible
  - our goal is to estimate the common effect size for studies included in the analysis and not to generalize to other studies/populations

- Random-effects model is appropriate if:
  - we expect the effect size to vary between studies due to different experimental conditions, location etc
  - we want our estimate of the mean effect size to be generalizable to a large group of similar studies

Choosing an appropriate model

Fixed, random and mixed effect model



Nakagawa et al. 2017 BMC

Weighting: a fundamental aspect of meta-analysis

- Different studies are not contributing equal amount of information to a meta-analysis
- In meta-analysis, study-specific effect estimates are weighted by the variance terms included in the model

In a **fixed-effects model**, study-specific estimates are measured by study precision (inverse of sampling variance)

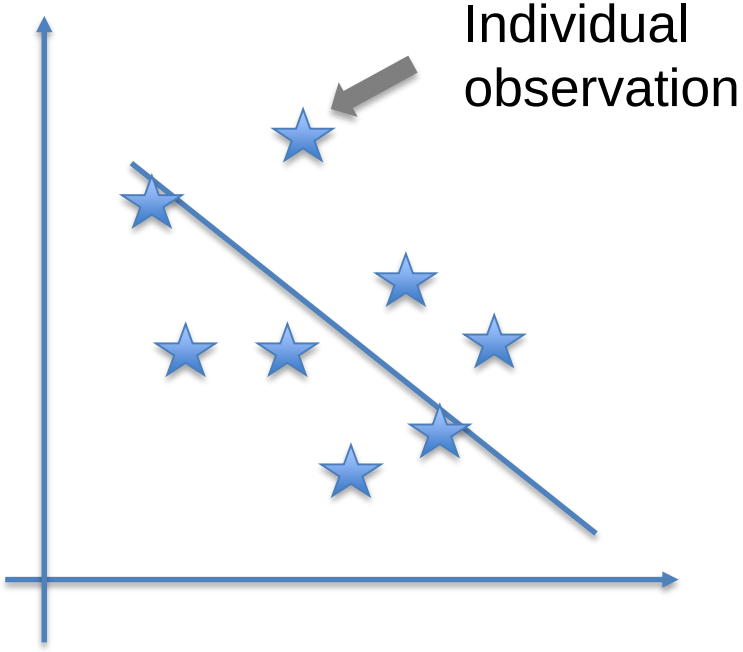
$$w_i = \frac{1}{v_i}$$

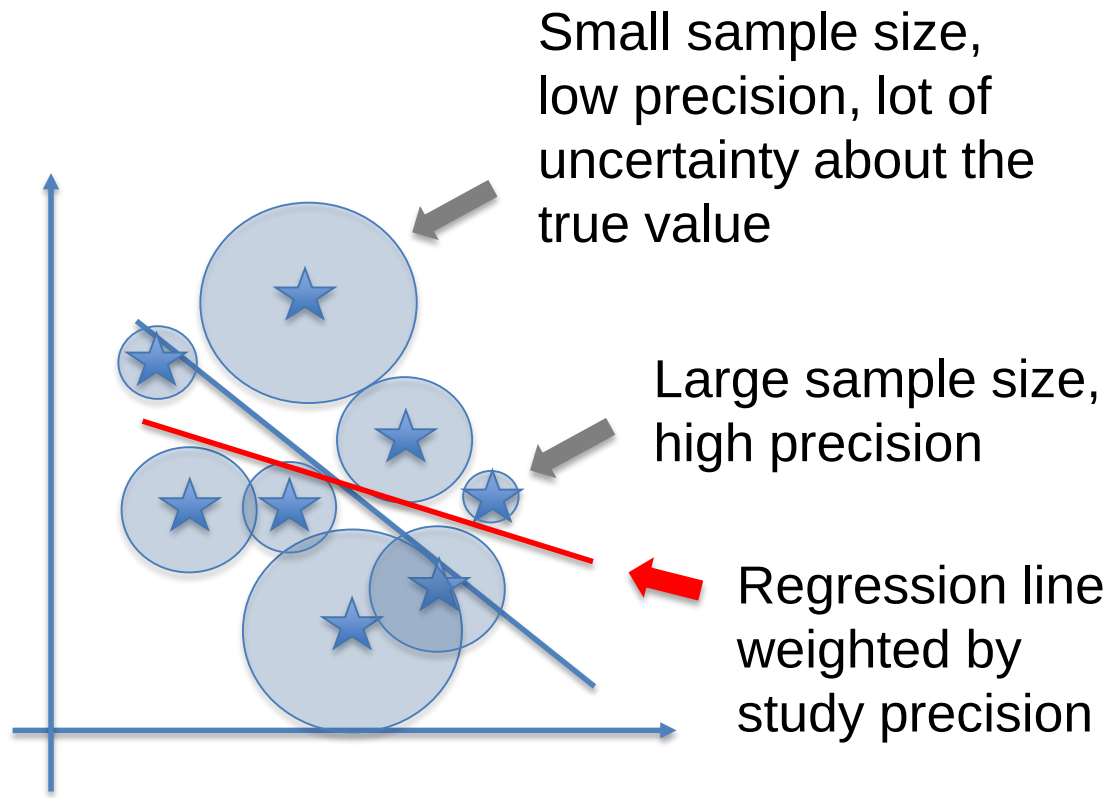
In a **random-effects model**, weighting includes both within- and between-study variance terms

$$w_i = \frac{1}{v_i + \sigma^2}$$

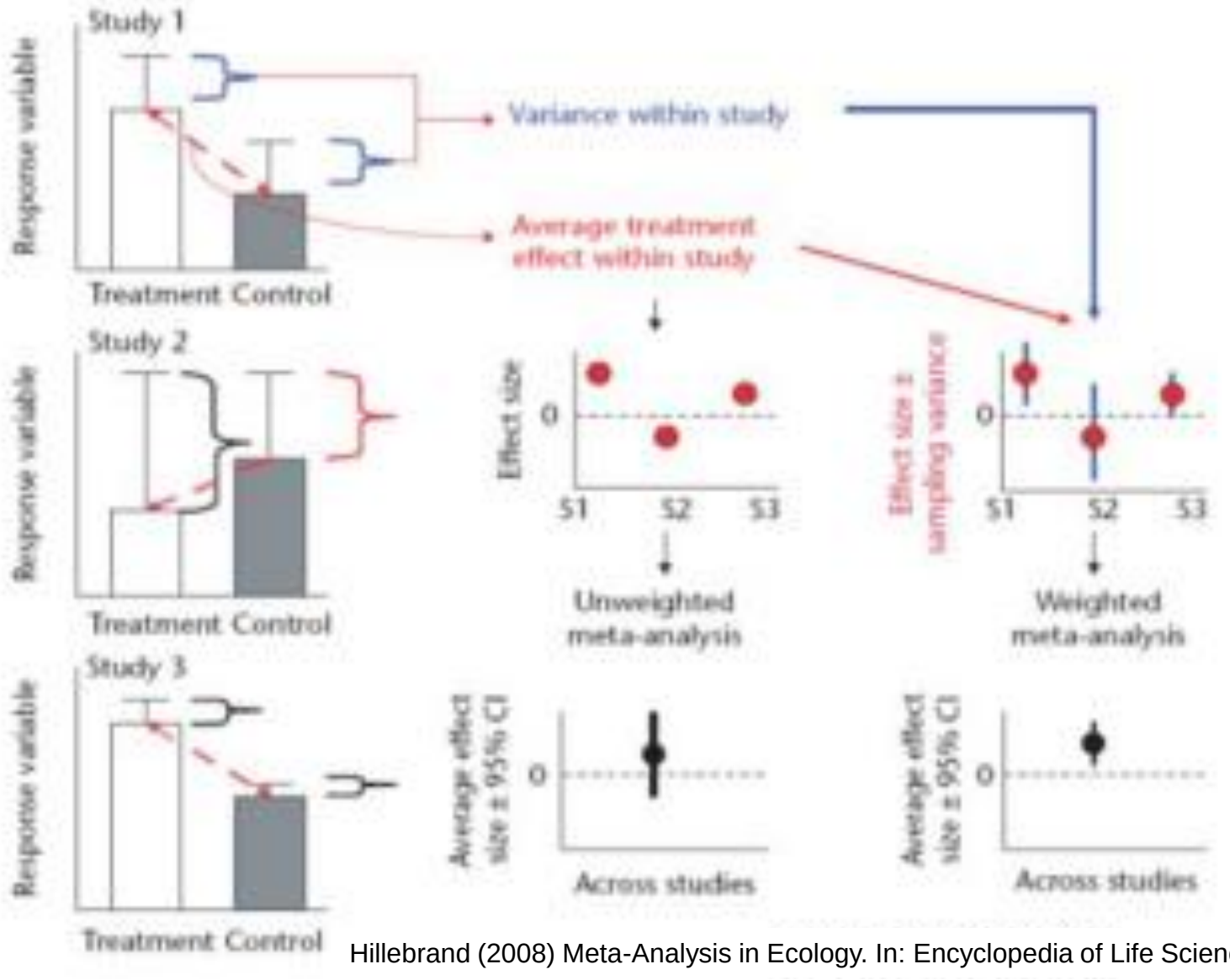
Weight can also incorporate other study-specific measures such as quality scores or phylogenetic information

Why do we need weighting?





Weighted vs. unweighted approach



Hillebrand (2008) Meta-Analysis in Ecology. In: Encyclopedia of Life Sciences (ELS)

Meta-analysis in biological and environmental sciences

Quantifying and explaining heterogeneity

Combining effect sizes across studies

The estimated mean effect size

$$\mu = \frac{\sum w_i \theta_i}{\sum w_i}$$

The variance of the mean effect size

$$s_\mu^2 = \frac{1}{\sum w_i}$$

The 95% confidence interval around the mean

$$95\%\ CI = \mu \pm 1.96 \cdot s_\mu$$

The mean effect is considered significantly different from 0 if its CI does not include 0 (cf. t-test).

Quantifying heterogeneity

Q-test

Is the observed variance in effect sizes significantly different from that expected by chance (i.e. due to sampling error alone)?

Total heterogeneity can be estimated as the weighted sums of squares:

$$Q_T = \sum_{i=1}^{n} w_i(\theta_i - \mu)^2,$$

θ_i observed effect in study *i*
w_i weight of study *i*
μ overall effect

$$Q_T \sim \mathrm{X}^2_{n-1}$$

If Q_T exceeds the critical value, effect sizes are more heterogeneous than would be expected by chance.

Quantifying heterogeneity

Comments on Q-statistic

- Calculation of Q_T is meaningful only in fixed-effects models where we assume that all studies share the same effects
- In random-effects models, we already account for additional between-study variation in effects, hence Q_T cannot be significant
- Choice between fixed- and random-effects models should not be made based on Q_T
- Q_T test has low statistical power if the number of studies in meta-analysis is low (<40)

Quantifying heterogeneity

I^2 statistic

I^2 statistics indicates % of heterogeneity that can be attributed to between-study variance

$$I^2 = \frac{Q_T - df}{Q_T} \cdot 100\%$$

Guidelines for interpretation:

- I^2 = 25% … small heterogeneity
- I^2 = 50% … moderate heterogeneity
- I^2 = 75% … large heterogeneity

Explaining heterogeneity

Meta-regression model

- A meta-regression model allows inclusion of study characteristics in the model to explain variability between studies and thereby reduce between-study variance
- Total heterogeneity Q_T in meta-regression model is partitioned into heterogeneity explained by the model Q_M and unexplained heterogeneity Q_E:

$$Q_T = Q_M + Q_E$$

- Mixed effects model: variation within groups is considered random, and between groups - fixed

Explaining heterogeneity

Meta-regression model

Types of moderators

Categorical

- Type of study organisms
- Type of treatment
- Type of experiment
- Type of response variable

Continuous

- Duration of the experiment
- Intensity of treatment
- Study location (latitude or altitude)
- Year of publication

Single categorical moderator => “ANOVA-like” structure

Single continuous factor => weighted linear regression model

Partitioning heterogeneity

Simple categorical model

Source of heterogeneity			df
Model (between-group heterogeneity)	Q_M	$Q_M = \sum_{m=1}^{M} w_m(\mu_m - \mu)^2,$	M-1
Error (within-group heterogeneity)	Q_E	$Q_T = \sum_{m=1}^{M} \sum_{k=1}^{K_m} w_{mk}(\theta_{mk} - \mu_m)^2,$	n-M
Total	Q_T	$Q_T = \sum_{i=1}^{n} w_i(\theta_i - \mu)^2,$	n-1

M=number of groups, K_m=number of studies in the mth group, n=total number of studies

Partitioning heterogeneity

Linear regression model

Source of heterogeneity			df
Model	Q_M	$Q_M = \frac{\beta^2}{s_{\beta^2}}$ Square of slope of the regression divided by its standard error	1
Error	Q_E	$Q_T - Q_M$	n-2
Total	Q_T	$Q_T = \sum_{i=1}^{n} w_i(\theta_i - \mu)^2,$	n-1

The R-package `metafor`

- *rma* returns an object of the class `rma`.
  - This object behaves like a list.
  - You can use the function `names` to see available elements.
- Frequently Used Elements

Name	Description
b	Summary effect
ci.lb, ci.ub	lower and upper bound of the 95% confidence interval
vb	variance-covariance of summary effects
fit.stats	model fit statistics log-likelihood, deviance, AIC, BIC, and AICc values
QE, QEp	test statistic for the test of (residual) heterogeneity and corresponding p-value
QM, QMp	test statistic for the heterogeneity explained by the model (called omnibus test of coefficients) and corresponding p-value
I2	value of I^2
yi, vi	Vectors of study effect sizes and corresponding variances

The R-package `metafor`

- Functions to extract informations from a `rma` object

Name	Description
coef	Summary effect
confint	confidence interval
summary	summary table of meta-analytic model

- Specifying the Model
  - The function can be used to fit fixed- and random/mixed-effects models

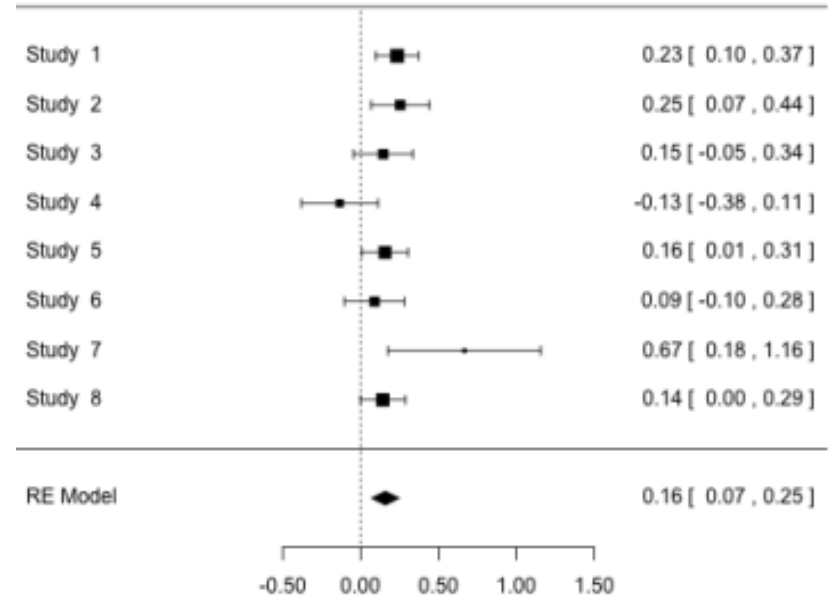
method	Description
FE	Fixed-effects model
DL	Random-effects model using DerSimonian-Laird estimator (Methods-of-Moments)
ML	Random-effects model using Maximum-Likelihood estimator
REML	Random-effects model using Restricted maximum-likelihood estimator

  - as well as meta-regression models including moderators using the mods=~mods1 + mods2

Visualizing heterogeneity
The Forest Plot

“Seeing the forest through the trees... “

- Plots effect sizes and their precisions
- The most common way to report the results of a meta-analysis
- Can help identify patterns across effects
- Can help spot large variation in effects or possible outliers
- `forest` is the function to plot forest plots of rma-objects
- Custimizing forest plots
  - order: Sort by "obs", "fit", "prec", etc.
  - slab: Change study labels
  - ilab: Add study information
  - transf: Apply function to effects
  - refline: Location to plot vertical ‘reference’ line
  - psize: Symbol sizes



EXAMPLE 2.1

1. Fitting and comparing a fixed-effects and random-effects model and obtaining heterogeneity statistics in metafor
2. Fitting a mixed-effects model (also called hierarchical, multi-level model, meta-regression) and obtaining heterogeneity statistics in metafor.

EXERCISE 2.1

Data

Stewart, G.B. A database on windfarm impacts on birds.

1. Determine the percentage each study contributed to the overall effect size variation. Which study contributes the most? How much? Use a `barplot` to show the percentages graphically.
2. Fit a fixed-effects model, a random-effects model
3. Obtain the Q-tests, and I^2 statistics. How can they be interpreted?
4. Fit a meta-regression. Determine to what extent the mods1 explains the remaining heterogeneity in the data. How can the estimated coefficients be interpreted? What is the percentage change in I^2 as compared to the random-effects model?
5. Perform multiple comparison analysis for a categorical trait.
6. Make pretty forest plots for modelled effects.